“A-ha, I hadn’t thought of that”: the Bayesian Problem of Awareness Growth

MSc Thesis *(Afstudeerscriptie)*

written by

Vasily Romanovskiy
(born March 22nd, 1996 in Moscow, Russia)

under the supervision of **Dr. Aybüke Özgün and Prof. Dr. Robert van Rooij**, and submitted to the Examinations Board in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
July 6th 2022	Prof. Dr. Yde Venema (Chair)
	Dr. Aybüke Özgün (Co-supervisor)
	Prof. Dr. Robert van Rooij (Co-supervisor)
	Prof. Dr. Sonja Smets
	Dr. Tom Sterkenburg

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

The Bayesian Problem of Awareness Growth is Bayesianism's apparent inability to account for the way our beliefs change in situations of awareness growth, those "a-ha, I hadn't thought of that" moments familiar to all of us. This thesis is concerned with attempts to modify Bayesianism so as to solve this problem and its chief contribution is threefold. First, it clarifies the Bayesian Problem of Awareness Growth by distinguishing two different versions of the problem, arguing against one of these versions on the grounds that it is at odds with the normative ambitions of Bayesianism and finally, delineating for whom and when my preferred version of the Bayesian Problem of Awareness Growth is a problem. The remaining two-part contribution is a contribution to two key debates around the Bayesian Problem of Awareness Growth. Some argue that a "none-of-the-above" option should be included to make space for future awareness growth. As far as the "none-of-the-above" option is concerned, I argue that if one wants to include such an option there is a way to proceed but arguments for always including such an option fail. I then turn to a putative norm that governs a change in beliefs in light of awareness growth: Reverse Bayesianism. I argue that it is not an exceptionless norm of rationality in the case of awareness growth, however, even when it fails it can be used as part of a general procedure for revising credences in light of awareness growth.

Contents

1 Introduction

Our limited awareness of the world around us is a common feature of both science and ordinary life. Physicists working in the 19th century never entertained Einstein's General Theory of Relativity. For a more relatable example imagine you decided to treat yourself to a dinner out and are now faced with a decision as to where to go. You are thinking about whether to get Surinamese, Indonesian or Italian. However, what has slipped your mind is that a Mexican restaurant is also an option - it's right by your place, tasty and above all reasonably priced. In fact, you had been to the restaurant before but thought that it had gone bankrupt due to the pandemic and you never came around to checking whether that was true. Regardless, while thinking about where to go for dinner you are not aware of the Mexican restaurant option. As a personal example of limited awareness, I had spent almost a decade living in the United Kingdom prior to having realised that parents were free to choose their children's middle names. Where I am from, one's middle name is fully determined by one's father's first name - I never even considered the possibility of a free reign, subject only to aesthetic judgement, in matters of choice of middle names.

Having seen a few examples, I hope the reader has come to appreciate how widespread the infusion of unawareness in our epistemic lives is and agrees that this phenomenon is a worthy topic of investigation; a normative theory of rationality should have something to say about those epistemic predicaments.[1] It should tell us how, if at all, our beliefs ought to change when our awareness grows. For instance, it should tell scientists how to reconsider the plausibility of existing theories in light of Einstein's breakthrough. Herein we will consider a theory of rationality: Bayesianism. Bayesianism, at its core, amounts to three claims. The first claim is that we have graded beliefs, called credences, that we assign to propositions. Second, these credences ought to obey the Axioms of Probability at each point in time, so a rational agent's credences just are probabilities. Third, these credences (which should be probabilities), across time, are governed as follows: upon having learnt a piece of evidence, an agent needs to update their probabilistic credences via a norm called Conditionalisation. Note that at least for some, Bayesianism isn't just *a* theory of rationality, it is *the* theory of rationality.

While achievements of Bayesianism might be impressive and widespread, it appears that Bayesianism has nothing to say about how agents ought to behave in cases of awareness growth. There are two possible views on the precise reason as to why that is. On one view, the Bayesian agent, in their omniscient way, assigns (probabilistic) credences to all possible propositions, even those pertaining to scientific theories yet to

[1]Herein, whenever I talk about rationality I mean it in a normative sense.

be conceived.[2] On this interpretation we cannot even make sense of awareness growth and new propositions coming to light; nothing is new and there's nowhere for awareness to grow. On another view, the Bayesian agent assigns (probabilistic) credences to a much more limited number of propositions. Here, genuinely new propositions can be encountered. However, there is then a new problem - we cannot assign this new proposition a non-zero credence because, as per the Axioms of Probability, the Bayesian agent attached maximal credence of 1 to the previous space of propositions. Thus, the agent has no reserve credence for the new proposition and they are forced to assign this new proposition credence 0. That also means that no matter what evidence the agent receives, Conditionalisation would nonetheless keep the updated credence at 0. Consider what that means for, say, Einstein's General Theory of Relativity. A Bayesian would have us assign credence 0 to the theory upon its introduction. That is the lowest possible credence. Then, no matter what evidence the theory receives the credence assigned to it stays that way. This thesis is concerned with Bayesianism's apparent inability to give prescriptions in cases where the awareness of the agent grows, that is, it is concerned with the Bayesian Problem of Awareness Growth.

The chief contribution of this thesis is threefold. First, it contributes to the literature on the Bayesian Problem of Awareness Growth by clarifying the problem. In particular, I distinguish two different versions of the problem, argue against one of these versions on the grounds that it is at odds with the normative ambitions of Bayesianism and finally, delineate for whom and when my preferred version of the Bayesian Problem of Awareness Growth is a *problem*. This clarificatory contribution is the subject of Chapter 3.

The remaining two-part contribution is a contribution to two key debates around the Bayesian Problem of Awareness Growth: the debate as to whether the agent's space of propositions should include a "none-of-the-above" option and a debate around Reverse Bayesianism, a putative norm for revising credences in light of awareness growth. The "none-of-the-above" option, called the catch-all proposition in the literature, is the subject of Chapter 4. In that chapter I consider both the kind of problems one encounters if a catch-all proposition is included in the agent's space of propositions and the extant arguments we have for a catch-all to always be included in the agent's space of propositions. In particular, I first argue that any successful attempt to include a catch-all must not assign what is called a prior and likelihood to the catch-all. While many share my scepticism about the prior and the likelihood almost no one goes be-

[2] Note that whenever I speak of a Bayesian agent I mean an agent who obeys the two core norms of Bayesianism, having probabilistic credences and updating via Conditionalisation. Thus, one should interpret descriptive looking claims as behaviour of an agent that is line with the normative prescriptions of Bayesianism.

yond the intuitive problems with assigning a catch-all a prior and likelihood.[3] I fill in this gap by arguing forcefully against assigning a catch-all either a prior or a likelihood. Having argued for this particular datum I privilege a catch-all proposal due to Wenmackers and Romeijn (2016) from amongst extant proposals that utilise a catch-all. I do so on the grounds that their proposal has an indefinite prior and likelihood in the catch-all and moreover, solves some additional problem with the catch-all. I shore up this result by first raising two problems against the proposal of Wenmackers and Romeijn (2016) and then defusing them. At that point things are looking rather good for the catch-all. I temper this optimism about the catch-all by showing that the two extant arguments for always including the catch-all fail. Thus, if one wants to include a catch-all proposition there is a way to proceed but there's no general requirement to always include it.

Reverse Bayesianism is the subject of Chapter 5. First, I present two counterexamples to Reverse Bayesianism that build on a counterexample due to Steele and Stefánsson (2021a). This apparent failure leads me to take a step back and consider why Reverse Bayesianism is a plausible norms in the first place. In short, it is because it provides a norm for changing credences in light of awareness growth that is most closely aligned with the main Bayesian norm - Conditionalisation. Reverse Bayesianism does so because it offer a conservative norm of credal update in light of awareness growth - the minimal change in credence brought about by awareness growth. Conditionalisation does the same in the case of evidential learning. I then use this insight to point to the role that Reverse Bayesianism can play even if it fails. This point is particularly noteworthy because it acknowledges that Reverse Bayesianism is not an exceptionless norm, and even that no single excpetionless norm is forthcoming, but there is still a systematic way to proceed. This goes against the grain of the literature where large-scale scepticism about an exceptionless norm leads people to conclude that there is no systematic way to proceed when it comes to changing credences in light of awareness growth. My proposal acknowledges the scepticism while denying the pessimistic conclusion.

Before we go onto the main contribution of the thesis some stage-setting is in order. This stage-setting is undertaken in Chapter 2. In it I clarify three notions hitherto encountered: Bayesianism, awareness as found in the Bayesian literature and rationality. Concerning the former, I will present the core tenets of Bayesianism more formally than I have done thus far, alongside arguments frequently offered in favour of it, as well as notable departures from the core of Bayesian claims. Concerning awareness, I present a typology of awareness, discuss how that typology maps onto

[3]A notable exception in the case of the likelihood on the catch-all is Roush (2005).

the Bayesian literature on awareness and crucially, outline and defend the notion of awareness that I will operate on in this thesis. Regarding rationality, while in ordinary speech it is a nebulous concept, enough can be said about rationality that we can safely proceed with the remainder of thesis. Without further ado, let us turn to the stage-setting.

2 Bayesianism, awareness and rationality

Since we will be concerned with the Problem of Awareness Growth for Bayesianism as a theory of rationality, it is only fitting that Bayesianism, awareness (as found in the Bayesian literature) and rationality are clarified.

2.1 Bayesianism

Bayesianism, at its core, is committed to three claims. The first of which is as follows.

Credal Realism. Agents have graded beliefs, called credences, that are attached to propositions.

Unlike the other two claims to follow, this is not a normative claim. A question one might naturally ask is: what sort of things credences are? Put roughly, a credence in a proposition is your expectation of, or the degree of commitment to, the truth of said proposition.

The second Bayesian claim and its first normative one is as follows.

Probability. An agent's credences *ought* to obey the Axioms of Probability.

In effect, a Bayesian agent's doxastic state is represented by a single probabilistic credence function over the space of propositions. Now, in order to formulate this more formally we first need to represent the things over which the probabilistic credence function is to be defined. We need to somehow represent propositions and the logical relationship in which propositions stand with each other. To that end, I will utilise the standard manoeuvre of representing the propositions in which the agent has a credence as a Boolean Algebra of propositions (Bradley 2017; Roussos 2020; Steele and Stefánsson 2021a).

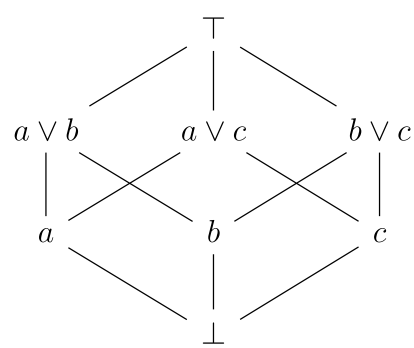
Boolean Algebra.[4] Given a non-empty set of basic propositions $\mathcal{X}$, denote the corresponding closure of $\mathcal{X}$ under $\wedge, \vee$ and $\neg$ as $\mathcal{X}^{cl}$. The structure $\mathbb{B}_{\mathcal{X}} = \langle \mathcal{X}^{cl}, \vee, \wedge, \neg, \top, \bot \rangle$ is a Boolean Algebra of propositions iff for all $a, b, c \in \mathcal{X}^{cl}$ the following laws hold:

[4] The definition is taken from Davey and Priestley (2002).

Commutativity	$a \vee b = b \vee a$	$a \wedge b = b \wedge a$
Associativity	$a \vee (b \vee c) = (a \vee b) \vee c$	$a \wedge (b \wedge c) = (a \wedge b) \wedge c$
Idempotence	$a \vee a = a$	$a \wedge a = a$
Absorption	$a = a \vee (a \wedge b)$	$a = a \wedge (a \vee b)$
Distributivity	$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$	$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$
Join Bound	$a \vee \neg a = \top$	$a \vee \bot = a$
Meet Bound	$a \wedge \neg a = \bot$	$a \wedge \top = a$

This all may seem rather abstract, to that end lets consider a set of basic propositions $\mathcal{X} = \{a, b, c\}$. The Boolean algebra, $\mathbb{B}_{\mathcal{X}}$ is depicted in Figure 1.

Figure 1: Boolean algebra spanned by $\mathcal{X} = \{a, b, c\}$.



Herein, whenever I speak of "an algebra spanned by set $\mathcal{X}$" for some set of basic propositions $\mathcal{X}$, I mean just the corresponding Boolean algebra, $\mathbb{B}_{\mathcal{X}}$. For future purposes, also define $a \models b$ iff $b = a \vee b$.

Now, as per the second Bayesian claim, an agent's credence function $Cr(\cdot)$ is a map from $\mathbb{B}_{\mathcal{X}}$ into the real numbers such that for all $a, b \in \mathcal{X}^{cl}$:

(1) $Cr(a) \geq 0$.

(2) $Cr(\top) = 1$.

(3) If $a \wedge b = \bot$, then $Cr(a \vee b) = Cr(a) + Cr(b)$.

From these 3 principles we can then further infer, amongst others, that for all $a, b \in \mathcal{X}^{cl}$ ought to be such that:

(4) $Cr(a) + Cr(\neg a) = 1$.

(5) $Cr(a) \leq 1$.

(6) If $a \models b$ and $b \models a$ then $Cr(a) = Cr(b)$.

(7) $Cr(a) = Cr(a \wedge b) + Cr(a \wedge \neg b)$.

(8) If $a \models b$, then $Cr(a) \leq Cr(b)$.

(9) If $Cr(a) = 1$ then $Cr(a \wedge b) = Cr(b)$.

(10) $Cr(\perp) = 0$.

Note that (8) says that if a logically entails b then the credence in a ought to be less than or equal to b. I previously said that I will take the three core Bayesian claims for granted. Thus, I take for granted a great deal of powers on part of our Bayesian agents. When it comes to closing credal assignment under logical consequence, the agents I consider will be boundless.[5]

At this point a few words are in order about the set of basic propositions $\mathcal{X}$. First, I will largely take this set to be finite unless I am presenting views of people who do not take it to be finite. Second, you may find it curious that I call this a set of basic as opposed to atomic propositions. After all the atoms of the Boolean algebra are all and only the elements of $\mathcal{X}$. The reason is that sometimes $\mathcal{X}$ will contain propositions that are not atomic because their content is derived from the content of other propositions in $\mathcal{X}$. Specifically, the catch-all proposition has meaning "none-of-the-above" where the "above" are all the other elements of $\mathcal{X}$.

The following point will also be important throughout the thesis. The elements of the set of basic propositions, $\mathcal{X}$, are mutually exclusive. That is, for any $a, b \in \mathcal{X}$ it is the case that $a \wedge b = \perp$ and thus $Cr(a \wedge b) = 0$. Moreover, they are collectively exhaustive, so $\bigvee \mathcal{X} = \top$. Hence $Cr(\bigvee \mathcal{X}) = 1$.

Turning back to Bayesianism, the conjunction of the first two Bayesian claims, Credal Realism and Probability, is a position called Probabilism.

Probabilism. A rational agent has credences and those credences obey the Axioms of Probability.

[5]For a classic paper introducing the Bayesian Problem of Logical Omniscience see Savage (1967). See Skipper and Bjerring (2020) for both a detailed philosophical discussion of logical omniscience in Bayesianism as well as an attempted solution that draws on dynamic logic.

Note how much leniency Probabilism gives in assigning credences. For illustrative purposes, consider the Boolean algebra spanned by $\mathcal{X} = \{\mathsf{H}, \mathsf{T}\}$, where H is the proposition that the coin will land heads, similarly for T. The following credence assignment is fully rational as per Probabilism: $Cr(\mathsf{H}) = Cr(\mathsf{T}) = \frac{1}{2}$, $Cr(\mathsf{H} \wedge \mathsf{T}) = Cr(\bot) = 0$, $Cr(\mathsf{H} \vee \mathsf{T}) = Cr(\top) = 1$. This is the sort of credal assignment we would expect in this case: heads and tails are equiprobable; the coin landing heads and tails simultaneously is deemed subjectively impossible and that either heads or tails comes up is a subjective certainty. However, from the perspective of Probabilism and hence Bayesianism, the following is an equally rational credence assignment: $Cr(\mathsf{T}) = Cr(\mathsf{H} \wedge \mathsf{T}) = Cr(\bot) = 0$, $Cr(\mathsf{H}) = Cr(\mathsf{H} \vee \mathsf{T}) = Cr(\top) = 1$. This assignment deems heads coming up a subjective certainty! The reason for this leniency is that Probabilism demands that credences cohere with each other as mandated by the Axioms of Probability. But as long as one's credences do that, and they do in both of the cases above, Probabilism doesn't tell you anything more substantively about what credences to have. Leonard Savage, one of the founders of (Bayesian) Decision Theory, states that Bayesianism "does not censure the neighbor whom we find superstitious or paranoid" (Savage 1967, p.597). The final Bayesian claim is as follows.

Conditionalisation. An agent with a credence function $Cr(\cdot)$, upon learning proposition $b \in \mathcal{X}^{cl}$ (such that $Cr(b) > 0$) with certainty and nothing else, ought to have a new credence function $Cr^{+}(\cdot)$ such that for all $a \in \mathcal{X}^{cl}$:

$$Cr^{+}(a) = Cr(a|b) := \frac{Cr(a \wedge b)}{Cr(b)} = \frac{Cr(b|a)}{Cr(b)} \times Cr(a).$$

Your credence after the learning experience is called your posterior credence and the one before the learning experience is your prior credence. The term $Cr(b|a)$ is called the likelihood of b on a. This terminology is especially prevalent in Bayesian Confirmation Theory where a is a hypothesis and b a piece of evidence. Note that by (7), we can further expand the expression above to:

$$Cr^{+}(a) = Cr(a|b) := \frac{Cr(b|a)}{Cr(b|a)Cr(a) + Cr(b|\neg a)Cr(\neg a)} \times Cr(a).$$

More generally, let $\{a_i\}_{i \in I}$ be a partition of $\mathcal{X}$. Which is to say that for two distinct elements of $\{a_i\}_{i \in I}$, a_i and a_j, $a_i \wedge a_j = \bot$ and $\bigvee_{i \in I} a_i = \bigvee \mathcal{X}$. We can then obtain the following:

$$Cr^{+}(a_i) = Cr(a_i|b) = \frac{Cr(b|a_i)}{\sum_{i \in I} Cr(b|a_i)Cr(a_i)} \times Cr(a_i).$$

The last formula will be particularly useful in the context of Bayesian Confirmation Theory. There, each a_i is thought of as a single hypothesis from a set of mutually exclusive and collectively exhaustive hypotheses and b is thought of as a piece of evidence.

Recall that in the introduction I stated that a proposition which is assigned a credence of 0 at a given time cannot then be assigned a non-zero credence via Conditionalisation whatever evidence is learnt. This can now easily be proven. Take propositions a and b such that $Cr(a) = 0$ and $Cr(b) > 0$. Then $Cr^{+}(a) = Cr(a|b) = \frac{Cr(a \wedge b)}{Cr(b)} = 0$ since $a \wedge b \models a$ and by (8) we obtain $Cr(a \wedge b) = 0$ from $Cr(a) = 0$ and (1). Note how nothing rested on the credence in b, other than it was not equal to 0. However, that is a pre-condition to applying Conditionalisation anyway.

Now, I tried to push through the presentation of Bayesianism so as to not provoke too many probing questions. However, I am sure the reader has plenty. Can we say more precisely what credences are? Why ought credences be probabilities? Suppose that credences ought to be probabilities, why ought we change our probabilistic credences via Conditionalisation? Granting the general thrust of Conditionalisation, do we have an alternative that doesn't rely on learning a proposition with certainty? I will take these in turn.

Naturally, attempts to characterise what *exactly* credences are have been made. For instance, de Finetti (1990) defines credence in terms of *actual* betting behaviour; to have a credence of $\frac{1}{2}$ in the proposition "it will rain today" is just to accept a bet up to an including even odds on the proposition "it will rain today". A bet with odds of even or better is said to be fair, if you assign a credence of $\frac{1}{2}$ to "it will rain today". One can also weaken this from *actual* betting behaviour to *dispositional* betting behaviour. So for instance consider someone who wouldn't bet on a proposition "it will rain today" if odds were even or better, because one thinks gambling is immoral. However, safe for this immorality, that person could well be disposed to bet if odds were even or better.

What reasons do we have to think that credences ought to be probabilities? Supposing we have probabilistic credences, why ought we change our credences via Conditionalisation? It is important here to distinguish two schools of thought. The most fervent and zealous supporters of Bayesianism offer a priori arguments in support of the two normative Bayesian principles: so-called Dutch Book arguments[6], Representation

[6]Bruno de Finetti (1992) offers a Dutch Book argument in favour of the view that credences ought to be probabilities. In turn, Lewis, David, et al. (1999) offers one in favour of Conditionalisation.

Theorem arguments[7] and Accuracy arguments.[8] They all share a common structure in that they take the Axioms of Probability and/or Conditionalisation and show how something "bad" happens to an agent whose credences violate those two norms. What differs between these arguments is what the "bad" thing is. Dutch Book Arguments show how agents with credences that aren't probabilities or with probabilistic credences that aren't changed by Conditionalisation will accept a set of bets which they regard as fair, while being subject to a sure loss of money. Arguments based on Representation Theorems feature a theorem which shows that an agent whose preferences over certain objects, like different courses of action, meet certain intuitive constraints of rationality, must have a probabilistic credence function.[9] Thus, the "bad" thing is a violation of a constraint on rationality of preferences, say transitivity. One version of Accuracy arguments show that a credence function that fails to be probabilistic or isn't updated by Conditionalisation is accuracy-dominated by one that is probabilistic. That is, no matter how the world turns out, the probabilistic credence function will be in a more accurate correspondence with the world, for a given measure of accuracy. Another version shows that if an agent's credences aren't probabilities or they aren't updated by Conditionalisation, there's a failure to minimise the expected inaccuracy of one's credences, for a given measure of inaccuracy.

The point of these arguments is to show that Bayesianism is necessary for rational agents. For a different school of thought, Bayesianism (and by extension Probabilism) is the best thing currently going, but that is contingent. It is the achievements of Bayesianism as shown in practice that persuades those people, but if a bigger and better thing were to turn up, they would turn away from Bayesianism. Two illustrative quotes that embody this more pragmatic tone can be found below.

> "My approach to BCT [Bayesian Confirmation Theory] is more pragmatic than a priori...Bayesianism is offered to the reader as a superior (though far from perfect) choice, rather than as the only alternative to gross stupidity." (Strevens, n.d., p. 5-6)

> "once we get to probabilism, it provides us with many fruits. Above all, it forms the basis of a unified theory of decision and confirmation - it combines

[7] Savage (1954) gives an argument for the two normative claims of Bayesianism based on a representation theorem of his. Before Savage, Ramsey (1926) used a representation theorem to argue for Probabilism.

[8] Examples of Accuracy arguments are Joyce (1998), Leitgeb and Pettigrew (2010a), Leitgeb and Pettigrew (2010b), Greaves and Wallace (2006) and Briggs and Pettigrew (2020). The latter two in favour of Conditionalisation, the former three in favour of credences being probabilities.

[9] Technically, this only establishes that they must be representable as having probabilistic credences. But usually, such arguments have a supplementary premise that aims to bridge the gap between "being representable as" to "actually having".

seamlessly with utility theory to provide a fully general theory of rational action, and it illuminates or even resolves various hoary paradoxes in confirmation theory. I consider that to be the best argument for probabilism." (Hájek 2009, p. 816)

To give a stylised and admittedly implausible example, suppose all you care about in a theory of rationality is that it gives an adequate response to the Raven Paradox (Hempel, 1945). You would then count yourself in the club of Bayesians represented above if you accepted Bayesianism as the current best going theory of rationality based on the fact you thought it gave the only adequate response to the Raven Paradox.[10] If another adequate, non-Bayesian solution were to show up then Bayesianism would lose its privileged status. Perhaps you even turn away from Bayesianism because the non-Bayesian solution is deemed by you to be superior. Realistically, for any given person there are many more things to care about in a theory of rationality than what it has to say about the Raven Paradox. However, I hope the point being made has been illustrated. Namely, the commitment of people whose sentiment is captured above is obviously defeasible and their choice in favour of Bayesianism is comparative. Thus, they look at how Bayesianism performs in practice and based on this performance they judge it to out-compete its rivals.

Moving on, what about learning experiences that fall short of learning a proposition for certain? Does Bayesianism have something to say? The usual answer is the so-called Jeffrey Conditionalisation (Jeffrey, 1990). It is best introduced by considering a certain perspective on Conditionalisation simpliciter. Essentially, when you learn a propositions b with certainty, your credences in the partition of $\{b, \neg b\}$ change in such a way that the first element is assigned credence 1 and the latter credence 0. To characterise uncertain learning, we consider possibly finer partitions and credence changes that are possibly less dramatic towards any side of the partition.

Given a set of propositions $\mathcal{X}$ as before, let $\{b_i\}_{i\in I}$ be a partition of $\mathcal{X}$ defined as before.

Jeffrey Conditionalsiation. An agent with a credence function $Cr(\cdot)$, who undergoes a (possibly) uncertain learning experience that has the effect of shifting, for each elements of the partition $\{b_i\}_{i\in I}$, credences from $Cr(b_i)$ to $Cr^+(b_i)$, ought to have credence $Cr^+(\cdot)$ such that for all $a \in \mathcal{X}^{cl}$:

[10] Attempted Bayesian solutions to the paradox include Hosiasson-Lindenbaum (1940), Mackie (1963), Howson and Urbach (2006) and more recently Rinard (2014).

$$Cr^+(a) = \sum_{i \in I} Cr(a|b_i) \cdot Cr^+(b_i).$$

It is important to note that Conditionalisation is just a special case of Jeffrey Conditionalisation.[11] The contrast between two Conditionalisations can also be brought out by considering the two principles below.

Rigidity. For all $a, b \in \mathcal{X}^{cl}$ such that $Cr(b) > 0$, $Cr^+(a|b) = Cr(a|b)$, where b and only b was learnt between $Cr(\cdot)$ and $Cr^+(\cdot)$.

Certainty. For all $b \in \mathcal{X}^{cl}$ such that $Cr(b) > 0$, $Cr^+(b) = 1$, where b and only b was learnt between $Cr(\cdot)$ and $Cr^+(\cdot)$.

The first point worth making is that the conjunction of these two principles is equivalent to Conditionalisation.[12] This equivalence is important because Jeffrey Conditionalisation still respects Rigidity but not Certainty.[13] This should come as no surprise since the very point of Jeffrey Conditionalisation was to capture uncertain learning experiences. It is worth pausing on Rigidity for a second, and not only because it is the common core of Conditionalisation and Jeffrey Conditionalisation.

Rigidity is tantamount to saying that the experience of learning b for certain shouldn't change how probable the agent takes b to make a. In other words, the relative credence of a given b doesn't change upon learning b because learning b does not give evidence or reason to change one's relative credence of *a given b*. This thought underlying Rigidity will be used as an argument in favour of a particular rule for changing credences in light of awareness growth, called Awareness Rigidity (Bradley 2017, p.258). Awareness Rigidity is closely related to Reverse Bayesianism, a norm that was mentioned in Chapter 1.[14]

In closing I would like to consider a somewhat radical departure from Bayesianism that nonetheless preserves some part thereof. This will serve a future purpose, this departure will be instrumentally valuable in setting the boundaries of my preferred version of the Problem of Awareness Growth in Chapter 3. Consider, time-slice epistemology (Hedden, 2015a; Hedden, 2015b). Within the realm of epistemology most

[11]Proof Under Claim 2, Appendix A2.

[12]Proof under Claim 1 in Appendix A, section A2.

[13]See Diaconis and Zabell (1982) for a detailed discussion of the properties of Jeffrey Conditionalisation, including Rigidity and Certainty.

[14]This relation will be made formally explicit in Chapter 5.

relevant to us, the position is as follows.[15]

Time-Slice Epistemology. What credences one ought to have at a time does not directly depend on one's credence at other times. When determining what credences you ought to have at a given time, your credences at other times play the same role as credences of other people.

So time-slicers retain Probabilism but they give up Conditionalisation. They give up Conditionalisation because Conditionalisation makes for a direct dependence between one's prior and posterior credences. What isn't ruled out is the uninteresting case of indirect dependence between credences at different times. For instance, your credences at one time lead to a certain course of action that in turn results in you acquiring some evidence, evidence that then factors into your credences at some later time.

Now that we have seen enough of Bayesianism and related positions, it is time to turn to awareness as understood in the Bayesian literature.

2.2 Bayesian awareness

It has to be acknowledged that awareness is a multifaceted concept. Below I present a typology adapted from (Elliott, n.d.).[16] To be aware of proposition a can mean any of the following.

Entertainability: Being able to entertain the possibility that a.

Attentional: Consciously attending to the possibility that a.

Deliberative: Having considered the possibility that a in one's reasoning.

While doubtlessly one can cut this typology more finely, we now have the bare bones structure needed to understand the different notions of awareness as found in the Bayesian literature. It is now time to do just that. We (debatably) find each of the three remaining notions of awareness - entertainability, attentional and deliberative -

[15] My statement of the position is an adapted version of statements found in Hedden (2015a), Hedden (2015b) and Builes (2020).

[16] I adapted this typology by dispensing with Elliott's derivability notion of awareness. This is because the agents I will be considering are logically omniscient in the sense that they close credal assignments under logical consequence. Thus, they can't be unaware in the sense of not being able to derive something instantaneously.

in said literature.[17]

A key source of motivation for the Problem of Awareness Growth, especially early on, were some of the grandest of conceptual breakthroughs in science - Einstein's Theory of General Relativity being the main culprit/victim. In fact, originally the problem was called The Problem of Novel Theories, not the Problem of Awareness Growth.[18] The commonly taken departure point for the Problem of Novel Theories is Glymour (1980).[19] The Problem of Awareness Growth, as such, seems to have gained prominence with Bradley (2017).[20]

At first blush, an introduction of a novel theory, especially one as monumental as Einstein's General Theory of Relativity, involves the entertainability notion of awareness. After all, it is fair to say that the General Theory of Relativity was not only not entertained by anyone before Einstein, but that it was not even entertainable pre-Einstein. Even if one doubts the latter point, one can agree that this holds for people living a given time before Einstein.

However, even we confine our attention to new theories more broadly; it needn't be the case that the entertainability notion is the most faithful notion of awareness at play. To draw a distinction due to Kuhn (1962), there is revolutionary science and normal science. The introduction of General Theory of Relativity amounts to a scientific revolution and therefore a very significant alteration to the conceptual space within the field of physics and indeed beyond. However, we can refine the notion of revolutionary science, as suggested by McMullin (1993). There are shallow revolutions, intermediate revolutions and deep revolutions. The discovery of the X-Ray, at least according to McMullin, is an instance of a shallow revolution. For our purposes, the discovery of the X-Ray nonetheless constitutes a discovery of a novel scientific theory or hypothesis and it is thus an instance of awareness growth. I also leave open the possibility that some instances of normal science, in Kuhn's terminology, would nonetheless constitute the

[17]We can also find cases of each of these notions of awareness being used in the broader formal literature on awareness. The attentional sense of awareness can be found in Fagin and Halpern (1987), Dekel, Lipman, and Rustichini (1998b) and Modica and Rustichini (1999). The deliberative sense can be found in Dekel, Lipman, and Rustichini (1998a). Finally, the entertainability notion of awareness can be found in Heifetz, Meier, and Schipper (2008) and Walker (2014).

[18]For the interested reader, the General Relativity example is used in the literature on Novel Theories by the following authors: Glymour (1980), Earman (1992) and Strevens (2010).

[19]Notable contributions to the Problem of New Theories include: Chihara (1987), Earman (1992), Otte (1994), Maher (1995), Strevens (2010), Henderson, Goodman, Tenenbaum, and Woodward (2010),Wenmackers and Romeijn (2016) and Canson (2020). Work touching related themes includes: Shimony (1970), Teller (1975), Salmon (1990) and Gillies (2001). I will engage with some of those in time. A noteworthy point is that two works mentioned predate Glymour's article.

[20]Notable work that explicitly puts things in awareness terms includes Mahtani (2020), Canson (2020), Steele and Stefánsson (2021a), Steele and Stefánsson (2021b) and Pettigrew (n.d.).

kind of cases which are aptly characterised as cases where new theories are introduced and our awareness grows.

The question is what notion of awareness best captures the entire spectrum of scientific discovery from the discovery of the X-Ray through to Einstein's General Theory of Relativity. To me, it's doubtful that awareness as entertainability does the best job. This is because the discovery of the X-Ray didn't feature an uprooting of extant theories or their replacement: "what changed were the experimental procedures used in working with cathode-ray equipment and the expected outcomes of such work." (McMullin 1993, p. 59). I am thus not sure that prior to the discovery of X-Rays scientists couldn't entertain their workings. The broader point is that it is plausible to think of some novel scientific theories as having been in the realm of conceptual possibility at the time of their introduction. While they lay in that realm they simply went unrecognised.[21]

I do not pretend that these considerations have *decisively* ruled against the entertainability definition of awareness, as far as the introduction of novel scientific theories is concerned. However, I do think strong enough doubt has been established about its viability that we can resort to a broader - has been considered - notion of awareness, at least as far as awareness pertains to novel scientific theories. This is also harmless since its entirely accurate to say that physicists in the 19th century hadn't considered the possibility of Einstein's General Theory of Relativity. It is just the reason for them having failed to consider it was rather strong.

For another notion of awareness in the Bayesian literature, consider the quote below.

> "there are at least three different senses or grades of unawareness that should be distinguished. Firstly, one can be unaware of a prospect when one has not encountered it or heard anything about it before, such as when one is unaware of the fact that there is a bus that goes to the town one wants to visit because one has never been there before. Secondly, one might have heard about something but have since forgotten about it or fail to recall it at a particular time because it slips one's attention. And, thirdly, one may deliberately exclude possibilities from consideration by 'blinding' oneself to them, by removing them or by having them removed from one's attention" (Bradley 2017, p. 253)

As evidenced by the fact Bradley includes cases of forgetting and even deliberately purging possibilities from one's consciousness as cases of unawareness, his notion of

[21] I am not alone in making this point. For instance, it can be found in Earman (1992, p.196).

awareness isn't one of entertainability. Indeed, he seems to confirm the attentional interpretation of awareness as the binding aspect of the three cases he considered.

> "What these situations of unawareness have in common is that certain contingencies or prospects are not available to the agent's consciousness at the time at which she is deliberating on some question" (Bradley 2017, p. 253).

Note that Bradley considers quite a diverse range of reasons a proposition (or prospect, as he calls it) might be unavailable to an agent's consciousness. A similarly impressive scale of reasons for unawareness is given by Steele and Stefánsson (2021b).

> "The examples we will appeal to include dramatic cases of limited awareness and subsequent growth that involve novel combinations of concepts (e.g., a "greenhouse effect" at the global level prior to the 20th century) or even novel concepts simpliciter (e.g., an "electron" prior to the late 19th century). But we will also appeal to more mundane cases of limited awareness and subsequent growth due to temporary shifts in attention or imaginative ability." (Steele and Stefánsson 2021b, p.8)

I think Steele and Stefánsson are operating on a deliberative notion of awareness. It is this notion which best unifies the particular cases of awareness in the quote above - even the case of a lack of awareness due to an attention shift. The reason for this is that Steele and Stefánsson spend quite a lot of time setting up and detailing the decision problem facing the agent and thus, they are analysing the agent's reasoning with respect to a given problem facing the agent. In such a case lack of awareness due to an attentional slip can be faithfully thought of as not having considered a possibility in the salient reasoning episode.

Having said all of this, now is the time to make my commitments to a definition of awareness known. To that end, a number of points are worth making. I take it that to be aware of a proposition is synonymous with having considered the proposition. Thus, I think, my understanding of awareness is most closely aligned with that of Steele and Stefánsson (2021a).

I choose this notion for two reasons. First, I take the Problem of Novel Theories and the Problem of Awareness Growth to be one and the same problem and the having-considered notion of awareness unites Problem of Novel Theories and the Problem of Awareness Growth under a single notion of awareness. Second, I believe the having-considered notion of awareness is an interesting-enough, if not the most interesting, notion of awareness. This notion of awareness is the broadest notion of awareness that captures the idea that to become aware of something is to encounter it for the first time. Now it might not be the very first time you encounter the possibility, but that the possibility is novel in the particular reasoning episode at hand. Thus, my notion of awareness is better put as having-considered in the reasoning episode at hand.

One might feel the urge to add further stipulations to the having-considered notion of awareness by, say, adding a further cognitive component to the failure to consider the proposition, for instance that one has failed to consider the proposition but not due to a lack of care or diligence but due to some exculpating circumstance. In response, a quote by Frank Ramsey (1926) is apt: "this would, I think, be rather like working out to seven places of decimals a result only valid to two" (Ramsey 1926, p.76). My definition of awareness, as well as the refinement sketched above are both readily subject to counterexamples insofar as their status as the one true definition of awareness is concerned. That does not concern me. It should not concern you either. It is more fruitful to get a working notion of awareness and then investigate its properties, in this case as far as they interact with Bayesianism. That way, we might illuminate both awareness *and* Bayesianism. Having done so, we then might have to go back to the awareness drawing board, but we will never get this sort of illuminating discussion if we are forever stuck on definitions. This holds not just in philosophy but the sciences as well. Development economists would get absolutely nowhere if all their time was spent defining inequality. Whereas having done some economics armed with imperfect notions of inequality, one can then go back to conceptual foundations being far better equipped.[22]

At any rate, all of the cases of awareness that I will consider are those that are usefully and faithfully represented as additions, in some sense, to the algebra of propositions. By this, I mean additions to the algebra, as opposed to leaving the algebra unchanged or just re-working its existing elements.[23] To my mind, linking the definition of awareness to certain transformations in the algebra is as good of a methodological rule as one can get when dealing with such nebulous concepts as awareness. It is now time to discuss another nebulous concept that lurks in the shadows of this thesis, and sometimes even comes out - rationality.

[22]For an illuminating discussion of how science can push on against the background of value-laden terms and concepts see Alexandrova (2018). For a philosophically sensitive discussion of economic inequality that is informed by (at least at the time) cutting edge economics see Sen et al. (1997).

[23]The "addition" part will be made much more precise in Chapter 5.

2.3 Rationality

The terms rational and irrational are ubiquitous in every day life. Those terms apply to a wide variety of things, amongst them beliefs, people, actions and taxation policies. Moreover, different people might mean quite different things when applying "rational" to the same thing, and the same person might mean quite different things applying "rational" to two different things. To draw an analogy with a concept that is (hopefully) more familiar than rationality, consider love. One and the same person means quite different things when professing love for a football team and love for a family member.[24] Even when we restrict application of love to people, love for a friend and love for a romantic partner are quite different things.

However, I think the case of love points a way out of the muddle around rationality - we can make our lives easier by delineating the scope of application of the terms rational and irrational. This is easily done. This thesis squarely concerns the rationality of credences.

Can anything more substantive be said about what rationality means, even once we have delineated the scope of application of the term? A way to do it would be to survey the populace at large and ask: "What do you think rationality of belief means?" I will not do that. This is because rationality is a concept that plays a certain role in epistemology, ethics and philosophy of action, and it's by considering this role that I intend to sketch the appropriate notion of rationality. To continue the love analogy, love has a certain normative profile. For instance, love is a thing worth having. Even this single datum can help rule out certain usages of the term "love", or at least allow them to be rendered as somehow degenerate uses of the term. For instance not all claims of love are consistent with that love being a thing worth having, but we wouldn't arrive at this from merely surveying usage.

Turning back to rationality, intuitions of laymen are at best secondary in getting a grasp on a philosophically interesting notion of rationality. This is because intuitions of laymen are either insensitive to philosophical concerns or to the extent that they are, they are nonetheless derivative from said concerns. It is by considering the place of rationality in the wider system in which it plays a role that allows us to sketch its contours.

Let us begin by noting that rationality is not solely in the purview of philosophers. Psychologists and economists have been noted to describe beliefs, actions and preferences as either rational or irrational. Since psychology and economics are empirical sciences, this kind of work on their part is a contribution to a *descriptive* theory of ra-

[24] Exactly one of them is unconditional but I will not say which.

tionality. A descriptive theory of rationality aims to describe, systematise and explain the extent to which humans, non-human animals and even institutions are rational. Moreover, rationality surely plays a role in explanation and prediction of human, animal and institutional behaviour. For instance, explaining individual behaviour in the stock market proceeds with recourse to what is rational for a given trader to do given her beliefs about the situation, propensity for risk, capital at hand and so on.

Of course, such descriptive work cannot proceed against some background theory of what is rational in a given situation - that is what desires, beliefs or actions one *ought* to have or perform in a given situation. As noted previously, my concern in this thesis is about this normative sense of rationality. More specifically, it is about the normative sense of rationality of credences.

We can now reflect upon cases where rational is used to denote something someone *ought* to have done. I think there are at least three different roles which a normative theory of rationality can play. Note that while there's more to rationality than rationality of credence (e.g. rationality of desire and action) since this thesis is solely concerned with rationality of credences I put the different roles of rationality of credences alone, for brevity.

> Evaluative: a normative theory of rationality can be used to judge credences as rational or irrational.
>
> Action-guiding: a normative theory of rationality can be used to guide agents in their credence assigning and credence updating policies.
>
> Attributive: a normative theory of rationality can be used to attribute praise or blame for having particular credences.

It will be useful to consider an example to appreciate the three-fold distinction. For illustrative purposes consider a moral norm: it is wrong, always, everywhere, and for anyone, to attack non-combatant civilians.

Take for granted that this norm is evaluative and that it is a correct norm. All of this doesn't imply that the norm above works as either an action-guiding norm or a norm for attributing praise or blame. For instance, a soldier might not only know of this norm but wholeheartedly subscribe to it, and yet it may fail to be a suitable action-guiding norm in war time because in such circumstances it might be hard or even impossible to tell a non-combatant civilian from a foe. Similarly, we can think of circumstances when the above norm was tragically violated, and yet the violator was rendered blameless, e.g. it was impossible to tell that the victim was a non-combatant and the violator was acting in what they thought was self-defence based on a credible threat and so on.

I will not peg down Bayesianism as fulfilling exclusively one of the roles above, but the distinction should be kept in mind as it will come in handy later. Specifically, in the next section I will argue against one of the two possible versions of the Problem of Awareness Growth because the kind of views it presupposes are at odds with all three of the normative roles above. This also illustrates a more general point that when I offer negative arguments against this or that view, I aim to make it independent of any given normative role that one might take Bayesianism to fulfil.

That said, Chapter 5 is to be read as analysing Bayesianism as offering action-guidance in the face of awareness growth. This is because action-guidance is the normative role explicitly assumed by other parties in the debate.[25] Moreover, my own positive proposal from that chapter is best interpreted as a proposal for how the agent themselves ought to respond to awareness growth.

[25]For instance, when motivating his book-length project, Bradley says: "A normative decision theory, adequate to such circumstances, would provide guidance on how bounded agents should represent the uncertainty they face, how they should revise their opinions as a result of experience and how they should make decisions when lacking full awareness or precise opinions (that they have confidence in) on relevant contingencies. The book tries to provide such a theory" (Bradley 2017, p. xiii). Similarly, Steele and Stefánsson say "This book is about the limits of internal consistency, in particular due to an agent's (limited) awareness, or what she perceives to be the possible contingencies or ways the world might be. But we need an understanding of the guidance that internal consistency can provide in order to see what are the shortcomings of this guidance " (Steele and Stefánsson 2021b, p. 11).

3 The Bayesian Problem of Awareness Growth

In this Chapter I turn squarely to the Bayesian Problem of Awareness Growth.[26] First, I use some examples of awareness growth to motivate and introduce my preferred version of the Problem of Awareness Growth. Second, I present another version of the problem and argue against it on the grounds that it is at odds with the normative ambitions of Bayesianism. Finally, I delineate for whom and when my preferred version of the Bayesian Problem of Awareness Growth is a *problem*.

3.1 The Problem

Consider the scenario below.

> DinnerDate1: Ruslana and Liudmila are approaching their 5th anniversary together and Liudmila is thinking about where Ruslana will take her out for the anniversary dinner. By this point she knows Ruslana pretty well, they have been together for 5 years after all, and she thus knows the likely restaurant options Ruslana is considering - an Indonesian restaurant, a Surinamese restaurant and an Italian restaurant. Liudmila takes all options to be equiprobable and considers only those three possible. As she is thinking out loud about these possibilities her flatmate, Valentina, being the culturally sensitive soul that she is, points out that there are at least two prominent types of Indonesian cuisine, Balinese and Javanese. Once this is brought to Luidmila's attention she continues to consider the Italian and Surinamese restaurants as possibilities, but she now splits the possibility of an Indonesian restaurant into two - Javanese and Balinese.

This situation can be captured more formally. Set $\mathcal{X} = \{\mathsf{Ind}, \mathsf{Ita}, \mathsf{Sur}\}$ of basic propositions forms the bedrock of the Boolean algebra. After Liudmila becomes aware of the two strands of Indonesian cuisine, the new set of basic propositions becomes $\mathcal{X}^+ = \{\mathsf{Bal}, \mathsf{Jav}, \mathsf{Ita}, \mathsf{Sur}\}$ and hence she is now aware of the Boolean Algebra spanned by $\mathcal{X}^+$. In the literature on awareness growth, this case of awareness growth is called a case of *refinement*. The term is apt because some possibility was cut finer. In this case it was the Indonesian restaurant option. To consider another type of awareness growth turn to the case below.

> DinnerDate2: Ruslana and Liudmila are approaching their 5th anniversary together and Liudmila is thinking about where Ruslana will take her out

[26] Herein just Problem of Awareness Growth.

for the anniversary dinner. By this point she knows Ruslana pretty well, they have been together for 5 years after all, and she thus knows the likely restaurant options Ruslana is considering - an Indonesian restaurant, a Surinamese restaurant and an Italian restaurant. Liudmila thinks she will definitely be taken to one of these restaurants but for all she knows any one of them is as likely as the other. As she is thinking out loud about these possibilities her flatmate, Valentina, tells her that she overheard Ruslana talking about a Mexican restaurant recently. Thus, Liudmila begins considering an additional option to the previous three, that of a Mexican restaurant.

The initial set of basic propositions $\mathcal{X}$ is as before. After becoming aware of the Mexican restaurant option her new bedrock of basic propositions is given by $\mathcal{X}^+ = \{\mathsf{Ind}, \mathsf{Ita}, \mathsf{Sur}, \mathsf{Mex}\}$ and hence she is now aware of the Boolean Algebra spanned by $\mathcal{X}^+$. In the literature on awareness growth, this case of awareness growth is called a case of *expansion*. This is an apt term since her space of possibilities expanded to include new ones. I follow the remainder of the literature in taking that every case of awareness growth is a case of either refinement or expansion or the result of a mix of the two. Let us now consider what credences Liudmila assigns in the second dinner date case.

In line with the story, Liudmila's credences at the start of the situation are such that $Cr(\mathsf{Ind}) = Cr(\mathsf{Ita}) = Cr(\mathsf{Sur}) = \frac{1}{3}$, $Cr(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}) = 1$ and for any two distinct elements of $\mathcal{X}$ their conjunction is assigned credence 0. So for instance $Cr(\mathsf{Ind} \wedge \mathsf{Ita}) = 0$. Note that herein I will denote the top elements of the algebra, in this case it is $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$, not as $\top$ but as $\bigvee \mathcal{X}$, where $\mathcal{X}$ is the set of basic propositions that generate the algebra. The purpose is to make this set perspicuous. This is necessary because we need to talk about different algebras - pre and post growth - thus the top element of the old algebra isn't the top element of the new algebra and representing them both with $\top$ would be misleading.

Now, once Liudmila becomes aware of the Mexican restaurant option she needs to assign credences to the Boolean algebra spanned by $\mathcal{X}^+$. Her credences post awareness growth will be denoted as $Cr^+(\cdot)$. Some things are easy enough. For instance, $Cr^+(\bigvee \mathcal{X}^+) = 1$. This much follows from the Axioms of Probability. Moreover, for all distinct elements of $\mathcal{X}^+$ their conjunction is assigned credence 0. So for instance $Cr(\mathsf{Ita} \wedge \mathsf{Mex}) = 0$. But what credence do we assign to Mex? If $Cr^+(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}) = Cr(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}) = 1$ then $Cr^+(\mathsf{Mex}) = 0$.[27] This is at odds with the description of

[27]Suppose $Cr^+(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}) = 1$. By (2) we know that $Cr^+(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur} \vee \mathsf{Mex}) = 1$ and we also know that $(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}) \wedge \mathsf{Mex} = \bot$. Thus, by (3), $Cr^+(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur} \vee \mathsf{Mex}) = Cr^+(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}) + Cr^+(\mathsf{Mex})$. Thus, $Cr^+(\mathsf{Mex}) = 0$.

the case because Liudmila considers the Mexican restaurant option a live possibility. Things get worse. To produce non-zero posteriors, Conditionalisation requires non-zero priors, but Mex is simply not in the domain of $Cr(\cdot)$. Thus the transition from $Cr(\cdot)$ to $Cr^{+}(\cdot)$ cannot happen purely by Conditionalisation. Thus, the core of Bayesianism is unable to capture the change in credences brought about by awareness growth. It is this inability which I take to be the Problem of Awareness Growth.

As a result of the problem, a fundamental question arises. What rule should govern the switch from $Cr(\cdot)$ to $Cr^{+}(\cdot)$? A successful rule needs to account for two key desiderata. First, that a new possibility can be assigned a non-zero credence. Second, it should also account for our intuitions in how credences in old possibilities change off the back of awareness growth.

Note that because I take the Problem of Awareness Growth to apply to both the formulation of new scientific theories as well as more mundane cases of awareness growth I take that a solution to the problem has to apply quite generally. Its application shouldn't be restricted to just, say, novel scientific theories. Perhaps that is too much to ask. But if we fail in our quest for a general solution, we can always retreat to more modest goals.

So what are the candidates rules for changing credences in light of awareness growth? Most of the attention has centred around two related rules: Reverse Bayesianism and Awareness Rigidity. A detailed exposition of both will be seen in Chapter 5, where the relationship between them is made precise. However, a first-pass statement of Reverse Bayesianism will be given to close this section.

Let $\mathcal{X}$ be a set of basic propositions of which the agent is initially aware and let $\mathcal{X}^{+}$ be the set of basic propositions of which the agent is aware post awareness growth. Furthermore, let $Cr(\cdot)$ be the agent's initial credence function and $Cr^{+}(\cdot)$ her credence function post awareness growth.

Reverse Bayesianism.[28] For all $a, b \in \mathcal{X}$ such that $Cr(a) > 0$ and $Cr(b) > 0$, $Cr^{+}(\cdot)$ over $\mathbb{B}_{\mathcal{X}^{+}}$ ought to be such that:

$$\frac{Cr(a)}{Cr(b)} = \frac{Cr^{+}(a)}{Cr^{+}(b)}.$$

Informally, this just amounts to keeping the ratios of credences in basic propositions that were there pre awareness growth constant. In the second dinner date case this amounts to continuing to take Indonesian, Italian and Surinamese restaurants as

[28] This formulation is akin to one found in Mahtani (2020) and Steele and Stefánsson (2021a).

equiprobable anniversary date venues. However, it opens up the possibility of assigning a non-zero credence to the Mexican restaurant option. If $Cr^+(\mathsf{Mex}) = Cr^+(\mathsf{Ita}) = Cr^+(\mathsf{Sur}) = Cr^+(\mathsf{Ind}) = \frac{1}{4}$ then Reverse Bayesianism is respected in the second dinner date case.

3.2 A different take on the problem

In the introduction, I noted that there are two versions of the Problem of Awareness Growth, each corresponding to two different views on why Bayesianism fails to appropriately account for awareness growth. Before moving on, it will be important to deal with the version that I will not be taking in this thesis. Notably, it is a distinction that, with a few notable exceptions, is not made - usually people just assume one of the versions and proceed from there.[29] By foregrounding the distinction in this section I aim to do my part in overcoming this neglect.

Even though I set it aside, it will be useful to give the other version due care for two reasons. The first reason is that some statements of the Problem of Awareness Growth take this alternative view of the issue. The second and more important reason is that distinguishing two versions of the problem will be of dialectic use further along in the thesis. In Chapter 4, I will object to an argument for the claim that a catch-all should always be included in the agent's algebra on the grounds that it subtly equivocates between the two versions of the problem. Thus, paying attention to the two different versions is not an idle pursuit, hence why I emphasise the distinction.

In his seminal "Why I am not a Bayesian" Glymour presses a host of objections to Bayesianism. While pressing a particular objection, the Problem of Old Evidence, he presents the following quoted view as a non-starter: "an ideal Bayesian would never suffer the embarrassment of a novel theory" (Glymour 1980, p.87). The reason why an ideal Bayesian would not suffer such an embarrassment is because they already have prior credences in all the possible propositions, at least that is what Glymour takes Bayesianism to be committed to. The view is a non-starter because scientists discover new theories all the time, thus becoming aware of new possibilities. Glymour thus concludes that a theory of rationality unable to account for this kind of awareness growth is in trouble.

He is not alone in thinking Bayesians might subscribe to the view that prior credences are assigned to all possible propositions. He is joined by, for instance, Earman (1992) and Strevens (n.d.). Moreover Patrick Maher notes that: "Some people seem

[29] In the philosophical literature Strevens (n.d.) is the notable exception. Another notable exception is Gillies (2001) who uses the distinction to force a dilemma for Bayesians, along the way to arguing that the classical statistics view of hypothesis testing is superior to the Bayesian one.

to think that Bayesian theory requires a person to give probabilities to every proposition that could ever be formulated" (Maher 1995, p.103). To be fair, there is some motivation for the view that Bayesians need to have priors in all possible propositions. If Conditionalisation is the only learning mechanism, then having credences in all the possible propositions seemingly allows the rest of inquiry to proceed by merely observing evidence and then applying Conditionalisation.

Having offered this view on the Problem of Awareness Growth, I now wish to set it aside. On a purely practical level, none of the recent literature on the Problem of Awareness Growth takes this view. But there are substantive reasons as well. A number of Bayesian proponents take Bayesianism to only be applicable in what Savage (1954) called "small worlds".

Savage's purpose was to offer Bayesianism not only as a theory of rationality of credences, but as a theory of rationality of decision-making and action. Credences play a seminal role but so do the agent's desires. Speaking in this context, Savage recommends that we structure decision problems by "artificially confining attention to a small world" (1954, p.16). A small world is the kind of thing one sees represented in a state-outcome matrix depicted in Figure 2. Say you are going for a stroll with the possibility of rain looming large on your mind. The matrix below represents the possible states of the world (e.g. it's raining vs it's not raining), possible actions (e.g. taking an umbrella vs not taking an umbrella) and possible consequences of actions at states (e.g. being dry vs not being dry). Realistically, in this situation the consequences will involve more details, e.g. you might be annoyed at taking the umbrella and not using it in the case it doesn't rain. However, for illustrative purposes the matrix in Figure 2 suffices.

Figure 2: State-outcome matrix

	Rain	No Rain
Umbrella	Dry	Dry
No Umbrella	Wet	Dry

Savage (1954, p.16) has some choice words calling the proposal to "envisage every conceivable policy for the government of his whole life", "utterly ridiculous" on the grounds that it is simply impossible for humans to conceive of the space of all possibilities. To the extent that Savage's point applies to the normative theory of decision-making it also applies to a theory of epistemic rationality.

To put Savage's point more forcefully, recall the three possible roles of a normative theory of rationality: evaluation, action-guidance and attribution of praise or blame. It is doubtful that Bayesianism can play either of the three roles if we take agents to have credences in all possible propositions. From an action-guiding perspective, the cognitive load of such a number of propositions would overwhelm anyone. This is even before we start considering updating credences based on incoming evidence. Thus, any action-guidance on part of Bayesianism so conceived will, to the contrary, result in paralysis. Relatedly, because the reality of real agents is so divorced from assigning all possible propositions credences - even implicitly or dispositionally - it is hard to see how Bayesianism, construed as a norm for assigning doxastic praise or blame, would have anything to say. Surely we are blameless for not assigning all possible propositions credences, for it is impossible to do so. Even taking the evaluative perspective, it at best can be said to apply to ideal agents. But ideal in what sense? Requiring that priors are attached to all possible propositions amounts to a gross kind of omniscience. To be clear, Bayesianism isn't without its omniscience issues - even I take for granted logical omniscience. But having priors in all the possible propositions is something quite different. It amounts to a kind of foreknowledge. How are scientists of today to know the content of future theories, let alone assign those theories a credence? How am I to assign a credence to the proposition that so and so climatic event happens in so and so location on some far flung date? I am no climate scientist after all.

Now it has to be acknowledged that there are more tempered positions than thinking credences are assigned to all possible propositions. For instance, one can take the space of possibilities of a given agent to include just the propositions whose content can be be presently formulated, or to only include propositions that will plausibly become relevant through the agent's lifetime. Thus, the critique stated above can plausibly be avoided, in whole or in part, for some such tempered positions that nonetheless fall short of taking credences to be defined in a small number of mundane propositions that play a part in a very localised scenario. I am prepared to admit as much. However, once we start taking the agential perspective seriously - as we must in order to fulfil the normative roles that a theory of rationality has to fulfil - I do not see how anything short of a very localised investigation that we saw in the dinner date cases is at all motivated.

One might further object that I was too quick in jumping the gap between assigning a proposition some credence and knowing its content. That charge is to some extent correct. In the next chapter we will see a concrete proposal about how to include propositions the content of which we don't know in our algebra and assign them credence. However, this proposal is akin to using a catch-all proposition, which will be the subject of scrutiny in the next chapter.

In closing, even though I set the credences-in-all-propositions view aside, the view of the problem I do take can be read as the problem into which we run if we try to dispense with having agents assigning prior credences to all possible propositions, and thus having them engage in more localised investigation. It's during the course of this investigation that they run into new possibilities.

3.3 The Bayesian *Problem* of Awareness Growth

Before we go on to tackle the Problem of Awareness Growth head on, it will be useful to consider the boundaries of the problem. Can we somehow diffuse it, or at least make it much less of a problem? This sort of stage-setting rarely goes beyond a sentence-length remark about what happens if the Bayesian has no solution to the problem. However, as we will see, a spectrum of views is available on the extent to which the Problem of Awareness Growth is a problem. Sketching out this spectrum allows us to see the problem in its broader context, and furthermore, offers inspiration for my own positive proposal for how solve the problem, which we will see in Chapter 5.

Two strategies for diffusing the Problem of Awareness Growth naturally suggest themselves. Awareness growth occurs across time and thus the Problem of Awareness Growth demands a diachronic solution. However, not everyone is convinced in the need for diachronic norms of rationality. According to time-slice epistemologists there are no diachronic norms of rationality, so the need to link credences pre and post awareness growth isn't there; you deal with credal assignments at those points in time quite separately. Of course, time-slice epistemology solves the Problem of Awareness Growth for Bayesianism at the cost of Bayesianism. After all, Conditionalisation is dispensed with as well and so we don't have a solution that can properly be called Bayesian.

There is another strategy. I argued at length that we shouldn't take the space of propositions over which credences are defined to include all propositions, and that we should conceive of a very localised space of all possible propositions, as seen in the dinner date cases. I thus conceive the Bayesian agent as engaged in a localised investigation. Perhaps awareness growth is the kind of event that prompts a re-working of our localised investigation. The idea is as follows. We started out with the original space of propositions to which priors are attached, possibly learnt some pieces of evidence and updated credences by Conditionalisation. We then become aware of a new possibility.

As a result, we just re-work our space of propositions, assign prior credences to this space, and then proceed as before: evidence learning, Conditionalisation and perhaps further re-working in light of future awareness growth. To my knowledge, the possibility of doing this was first broached by Paul Teller: "to the best of my knowledge no personalist has tried to describe conditions which determine when an agent should continue conditionalization on his old probability function and when he should start from scratch, restructuring his problem and making new subjective evaluations of a prior probability function" (Teller 1975, p.170).

If one finds this proposal unpalatably anarchic it can be made less so. Strevens (2010) for instance suggests a proposal along the same lines as above but with a key difference: we are to keep track of all the evidence learnt and when we re-work the space of propositions we should then conditionalise on all the evidence learnt before the re-working. I have a lot of sympathy for Strevens' proposal and I think we should proceed along lines outlined by him, in the absence of a satisfactory proposal that is more "smooth" or "continuous". Reverse Bayesianism, if successful, provides just this kind of smooth, or at least smoother, transition from credences pre awareness growth to credences post awareness growth.

One can now look at the debate around Reverse Bayesianism as a debate about the nature of awareness growth and whether, at least sometimes, awareness growth upends enough to force us to abandon not only our previous space of possibilities but drastically re-work our credal assignment.

To sum up, the Problem of Awareness Growth isn't a problem for those who boil down credal assignment to assigning credences at discrete points in time with no link between those times, but that results in giving up Conditionalisation and Bayesianism with it. There is also a way out of the problem that just re-works the space of propositions at every instance of awareness growth with new priors being re-assigned each time awareness growth takes place.

I think this latter point should assuage some anxieties one might have had about the Problem of Awareness Growth. A proposal like Strevens' delineates the bottom line of how bad the problem can get - Bayesianism will not be left paralysed by cases of awareness growth. It is from this island of relative security that we can now tackle the problem itself. Our next port of call is the catch-all proposition.

4 The Catch-All Proposition

In light of the fact that awareness growth is ubiquitously present in our lives, it seems prudent that we factor in future awareness growth. But how does one expect the unexpected? It might have struck you during the discussion of the dinner date cases that Liudmila should have been more cautious; she shouldn't have taken the explicit possibilities she formulated (Sur, Ind and Ita) to exhaust all the possibilities. That is, Lidumila should have included a proposition corresponding to "none of the above". This proposition is called the catch-all proposition, which I denote as CA. The catch-all proposition serves the function of acknowledging that the space of explicitly formulated possibilities is not actually exhaustive. In our context, the catch-all proposition also serves the purpose of being a store of reserve credence, so that upon the formulation of a new explicit possibility we have some left-over credence to give it. Thus, the proposal is that we have a new putative norm for Bayesianism.

Catch-all. An agent's set of basic propositions $\mathcal{X}$ ought to include a catch-all proposition CA and we should assign it a positive credence.[30]

Now, how does the catch-all help us with the Problem of Awareness Growth? First, note how it changes cases of awareness growth by expansion into cases of awareness growth by refinement. Proponents of the catch-all contend that Liudmila's basic propositions at the start of deliberation should be $\{\mathsf{CA}, \mathsf{Ind}, \mathsf{Sur}, \mathsf{Ita}\}$ and not just $\{\mathsf{Ind}, \mathsf{Sur}, \mathsf{Ita}\}$. In DinnerDate2 Liudmila then becomes aware of Mex. Since Mex is included in the catch-all, Liudmila's $\mathcal{X}^+$ is given by $\{\mathsf{CA}', \mathsf{Mex}, \mathsf{Ind}, \mathsf{Sur}, \mathsf{Ita}\}$. That is, CA fragmented into Mex and CA'. Much like in DinnerDate1 the proposition Ind fragmented into propositions Jav and Bal. Thus, if we always include a catch-all, all cases of awareness growth are cases of refinement.

To be clear, changing all cases of awareness growth into cases of awareness growth by refinement doesn't answer the Problem of Awareness Growth by itself, but it helps answer the problem in conjunction with a suitable rule for revising credences in light of awareness growth. The intuitive idea is to shave off the reserve credence stored in the catch-all proposition and give it to the proposition the agent has become aware of (Earman 1992, p. 196). The dynamics of this shaving-off are well captured by

[30] I have kept this purposefully vague. For one, this formulation isn't committed to saying who exactly should include the catch-all proposition - the agent themselves or the person analysing the agent. Perhaps more crucially, I have left out whether the catch-all should *always* be included, or only *sometimes*. Until we have considered the catch-all proposition in detail, it is premature to commit to more precise formulations.

Reverse Bayesianism on the proviso that a proposition which fragments is somehow equivalent to, or can be identified with, the propositions into which it fragments. To illustrate, take a set of scientific hypotheses at some time $\mathcal{X} = \{\mathsf{Hyp}_1, \mathsf{Hyp}_2, \mathsf{CA}\}$. The agent has a credence function $Cr(\cdot)$ over the algebra spanned by $\{\mathsf{Hyp}_1, \mathsf{Hyp}_2, \mathsf{CA}\}$. Suppose then the agent becomes aware of a new possibility and so CA splits into Hyp_3 and CA' resulting in $\mathcal{X}^+ = \{\mathsf{Hyp}_1, \mathsf{Hyp}_2, \mathsf{Hyp}_3, \mathsf{CA}'\}$. It can be shown, under the assumption that $\mathsf{CA}' \vee \mathsf{Hyp_3}$ just is the same proposition as CA, that Reverse Bayesianism mandates that $Cr^+(\mathsf{Hyp}_i) = Cr(\mathsf{Hyp}_i)$ for $i = 1, 2$, $Cr^+(\mathsf{CA}) = Cr(\mathsf{CA})$ and that $Cr^+(\mathsf{CA}') = Cr(\mathsf{CA}) - Cr^+(\mathsf{Hyp}_3)$, which is exactly Earman's shaving-off proposal.[31]

The intuitively appealing idea of including the catch-all in the set of basic propositions nonetheless has a curious status in the context of the Problem of Awareness Growth. Many proposals use a catch-all one way or another (Shimony, 1970; Salmon, 1990; Maher, 1995; Henderson et al., 2010; Wenmackers and Romeijn, 2016). Yet, the problems of including a catch-all are well known: shaving off credence from the catch-all leads to successively less credence being available for future instances of awareness growth, resulting in new propositions being assigned too low of a credence to stand a chance to become the most probable possibility in light of evidence (Earman, 1992); shaving off fails to account for the radical changes in credence brought about by scientific revolutions (Earman, 1992); it is especially hard to assign a prior credence to the catch-all (Shimony, 1970; Henderson et al., 2010; Bradley, 2017; Steele and Stefansson, 2021a); and it is especially hard to say what the likelihood of the evidence on the catch-all is (Salmon, 1990; Roush 2005, Strevens n.d.). On the other hand, some of those who don't use a catch-all are apologetic for so doing (Strevens, 2010). More radically, some argue that the inclusion of a catch-all is rationally mandated at all times (Canson, 2020). In sum, we have a philosophical puzzle on our hands.

In this chapter, I concern myself both with arguments against the catch-all and arguments in favour. I will initially assume that the catch-all is included in the agent's algebra and consider what problems we thereby run in. On the negative side, as far as the catch-all is concerned, I will argue that assigning a prior to the catch-all, while not demonstrably impossible, looks very difficult to do in general and the likelihood on the catch-all fares even worse and is at least sometimes utterly intractable. While a lot of authors share my conclusion, their reasoning is either scant or it does not consider the full depth of considerations that proponents of the catch-all can marshal in its defence. I fill in this gap by arguing forcefully against assigning a catch-all either a prior or a likelihood.

[31] Proofs under Claims 1 and 2, Appendix B.

More positively, it will be shown that there is an implementation of a catch-all proposition in conjunction with a suitable procedure for revising credences, due to Wenmackers and Romeijn (2016), that answers two criticisms of the catch-all due to Earman (1992). Since their implementation does not assign a definite prior or likelihood to the catch-all, we have a result indicating that the inclusion of a catch-all might not be so problematic after all. I shore up the implementation of the catch-all by Wenmackers and Romeijn, by defending it from possible objections.

Finally, on the side of arguments for the inclusion of the catch-all I will contend that the two extant arguments for always including the catch-all fail and thus, there's currently no argument for the claim that the catch-all should always be included.

But before we move onto the arguments against and for the catch-all, let us consider two different proposals on how to implement the catch-all proposition in a Bayesian setting, at least when the catch-all proposition is understood functionally - as a way to acknowledge that our minds are parochial and as a way to break free of the negative consequences of this parochialism. The two proposals we will see will subsequently act as both targets of criticism and ways for proponents of a catch-all to defend themselves.[32]

[32] In the context of the problem of Awareness Growth there are three more authors who utilise a catch-all. Shimony (1970) and Salmon (1990) simply denote the catch-all proposition in a way analogous to my denoting it with CA. Thus, for them it is just a proposition with special qualities. Henderson et al. (2010) are the third catch-all proposal, at least on my functional understanding of a catch-all. These authors invite us to construe the space of scientific hypotheses as hierarchical, or levelled. That is, instead of a single $\mathcal{H}$ we have $\mathcal{H}^i$ for some $i \in I$. The guiding idea is to capture the post-positivist view of science as occurring within paradigms Kuhn (1962) or research programmes Lakatos (1978) or research traditions Laudan (1977). The highest level corresponds roughly to the most general theories of the domain: paradigms, research programmes, research traditions, whatever one's heart desires. The higher level generates the theories in a level one down and this continues all the way to the bottom level. This approach is called the Hierarchical Bayesian model (HBM) approach. This proposal meets my functional definition of a catch-all because: "there is no need to enumerate the lower-level theories" and "this provides a way of effectively modeling the introduction of theories that are 'new' in the sense that they may be regarded as implicit in assumptions about how the lower-level theories are generated, although not explicitly enumerated or recognized as possible hypotheses."(Henderson et al. 2010, p.191). However, there are serious limitations of this kind of catch-all proposal, as far as a general catch-all proposition to-be-included in all algebras is concerned. Hence why I do not include it in the body of the next sub-chapter. The authors admit as much when they say: "HBM Bayesianism can be validly applied only if we are in a situation in which there is a fixed and known hierarchy that it is reasonable to suppose will not be altered in the course of the investigation" (Henderson et al. 2010, p.191). Two points standout from this quote as far as limiting the scope of application of this approach. First, that there is a hierarchy. Suppose it is plausible in case of scientific theories. It is not plausible for more mundane cases, like the dinner date cases, where no plausible hierarchy is available let alone known. Second, supposing that there is a hierarchy, the fact it is always fixed is implausible.

4.1 Varieties of catch-all

Because the proposals we will encounter use different notation amongst themselves and indeed different notation from me, a little bit of stage setting is in order. Hitherto I have denoted the set of basic propositions as $\mathcal{X}$. The basic propositions are the building blocks of the algebra which make up the domain of the credence function. The proposals we will encounter in this section all have slightly different building blocks. However, some common ground can be established. Because all proposals in this section concern themselves with scientific hypotheses, the set $\mathcal{H} = \{\mathsf{Hyp}_1, \mathsf{Hyp}_2, ...\}$ of (mutually exclusive) scientific hypotheses will play a unifying role. Another important piece of notation is $\mathcal{H}_n$, the set of all scientific hypotheses that have been explicitly formulated up to and including time n. As will be seen, each of the proposals below get creative with $\mathcal{H}$ one way or another, in order to construct the domain of the credence function in such a way so as to be able to have credences in hypotheses that haven't yet been explicitly formulated.

The first proposal that we will consider is due to Maher (1995).[33] The domain of the credence function in Maher's system is built from the Cartesian product of two sets of propositions $\mathcal{H}$ and $\mathcal{V}$. To put things in terms of the set of basic propositions, we have $\mathcal{X} = \mathcal{H} \times \mathcal{V}$. Thus, the domain of the credence function is $\mathbb{B}_{\mathcal{H}\times\mathcal{V}}$. $\mathcal{H}$ is as stated above. $\mathcal{V}$ is the set of propositions of the form $m(\mathsf{Hyp}_i) = r$ for each hypothesis Hyp_i, where r is a real number. This mysterious $m(\cdot)$ is a measure of theoretical virtue for a given hypothesis Hyp_i. Theoretical virtues are the sort of thing that are supposed to guide choice amongst competing hypotheses: simplicity, scope, accuracy, consistency, fruitfulness (Kuhn, 1977) or what have you.

Within this framework, we can designate future theories in such a way as to divorce them of their content: by reference to the temporal order in which they will enter the space of possibilities. That is, for a given time n, we make $\mathcal{H}$ such that $\mathcal{H} \setminus \mathcal{H}_n \neq \emptyset$. That way we have $\mathcal{H}$ include explicitly formulated theories and those that will be formulated in the future: Hyp_{n+1}, Hyp_{n+2} and so on. Since all such elements of $\mathcal{H}$, whether explicitly formulated or not, are elements of the algebra in Maher's system, they can be assigned a non-zero credence. Thus, we have left open the possibility that there are possible hypotheses beyond those explicitly formulated and they can be assigned a non-zero credence. Hence, Maher's proposal meets my definition of a

[33]Maher's proposed solution to the Problem of Awareness Growth is noteworthy not just for the fact he has a creative implementation of the catch-all hypothesis. First, at least to my knowledge, he is the first to explicitly start thinking in terms of the algebra of propositions. Second, he attempts to solve the Problem of Awareness Growth with the combination of a catch-all and Conditionalisation. Note that this doesn't contradict my earlier claim that the Problem of Awareness Growth cannot be solved *purely* by Conditionalisation.

proposal that includes a catch-all proposition. That is enough for now, but we will return to Maher's system when discussing one of the key objections to including a catch-all: that assigning said catch-all a prior credence is so difficult that it seems like it should not be assigned a prior at all. That is where the theoretical virtue measure comes in - they are ways to assign credences to as-yet unformulated theories.

The second catch-all proposal is due to Wenmackers and Romeijn (2016). Their framework is as follows. All the possible pieces of evidence are captured as an algebra, $\mathbb{B}_{\mathcal{E}}$, spanned by the set of basic pieces of evidence $\mathcal{E} = \{\mathsf{Evi}_1, \mathsf{Evi}_2 ...\}$. Intuitively $\mathcal{E}$ can be a set of possible outcomes of an experiment, e.g. a lateral flow test coming out positive or negative. Now, the set of hypotheses $\mathcal{H}$ is as follows:

$$\mathcal{H} = \{Pr : \mathbb{B}_{\mathcal{E}} \to [0, 1] |\ Pr \text{ is a probability function}\}.$$ [34]

Thus, on Wenmackers and Romeijn's proposal, a hypothesis is specified extensionally in terms of its empirical content: the probability it assigns to all the possible pieces of evidence. Note that rational credence function, Cr, is a subjective probability function and above, you see a reference to a probability function, Pr. It is crucial to stress that I use Pr to denote some other *kind* of probability function than a credence function. Namely, Pr is used to denote a physical probability function. Thus, hypotheses describe chances of events in the evidence space. The domain of the credence function is the algebra, $\mathbb{B}_{\mathcal{H}\times\mathcal{E}}$, spanned by $\mathcal{H} \times \mathcal{E}$. When we speak of some hypothesis Hyp as an element of the domain of the credence function $\mathbb{B}_{\mathcal{H}\times\mathcal{E}}$ that is a shorthand way of denoting the element $\bigvee_{i\in I}(\mathsf{Hyp}, \mathsf{Evi}_i)$. Similarly, when we speak of some piece of evidence Evi this is a shorthand way of saying $\bigvee_{i\in I}(\mathsf{Hyp}_i, \mathsf{Evi})$.

The framework enables Wenmackers and Romeijn to formalise those hypotheses that have not been explicitly formulated yet as a single catch-all. We can take $\mathcal{H}_n$ which is a set of hypotheses that have been explicitly formulated, then the catch-all proposition is the unique element, $\neg \bigvee \mathcal{H}_n$.

Having seen two concrete and sophisticated implementations of a catch-all proposition, it is now time to turn to arguments against the catch-all.

4.2 Arguments against the catch-all

The catch-all seems like an easy way to factor in future awareness growth and reflect this in one's credal assignment. The aim of this section is to evaluate whether appearances

[34]The actual proposal from Wenmackers and Romeijn is a little more complicated. What I call the set of hypotheses they call the set of *elementary* hypotheses, denoted by $\mathcal{H}_e$. The set of all hypotheses simpliciter, $\mathcal{H}$, is the algebra spanned by $\mathcal{H}_e$, $\mathbb{B}_{\mathcal{H}_e}$. Therefore, hypotheses are probability functions or disjunctions thereof. This added sophistication is unnecessary for our purposes and thus I simplify this in the interest of the forthcoming discussion. Nothing crucially depends on this simplification.

are illusory and whether the catch-all actually introduces more problems than it solves. So, let us suppose that the catch-all is included in the domain of the agent's credence function and see where that leads us.

4.2.1 Prior in the catch-all

Let us first focus on the prior in the catch-all. The sentiment that assigning the catch-all a credence is deeply problematic appears to be widely shared. Shimony (1970), who pioneered the use of the catch-all in the Bayesian framework, suggested we should not assign the catch-all a prior credence at all. Wenmackers and Romeijn (2016) say:

> "it is not sensible to assign any definite value to the prior of the catch-all" (p. 1234), because "the catch-all is not based on a scientific theory, the usual 'arational' considerations [...] for assigning it a prior, namely by comparing it to hypotheses produced by other theories, do not come into play here" (p. 1234).

Bradley (2017) asks:

> "given that we don't know anything about the prospects that we are potentially unaware of, on what basis are we to determine [...] what probability we should assign to the catch-all prospect?" (p. 255).

In a similar vein, Steele and Stefánsson (2021a) note that: "it is hard to see how the catch-all could be well-defined for the agent" (p.1226). This presumably implies that they find the content of the catch-all so inscrutable that this inscrutability invalidates attempts to assign it a prior credence.

The quotes above are largely not only the first but the last words the authors above spend on the catch-all. While these authors are right to be sceptical of the prior in the catch-all, it is far from trivial to *show* that we shouldn't assign a prior to the catch-all. Correspondingly, this section contributes to the literature on awareness growth by providing sustained and extensive argument for the claim that we shouldn't assign a prior to the catch-all. The next section does the same for the likelihood.

Note that the putative issue with assigning the catch-all a prior (and likelihood) isn't that it is just hard to do, but rather that it is categorically harder to do than for propositions explicitly considered. It is harder to the extent that it seems best not to assign it a definite prior (and likelihood) at all.

I have previously noted that if one is just sticking to the core of Bayesianism, the range of what's permissible to do with respect to prior assignment is rather large. Still, there needs to be some plausible way for the agent to assign prior credences in the first place. There needs to be some way to say that the agent has credence x in a given proposition as opposed to y.[35] However, one might note that we are never in a position to say that the agent's credence in a given proposition is 0.7500000001 and not 0.7500000002.[36] I will take this impossibility for granted. However, taking the impossibility for granted is consistent with being able to say that the agent's credence is in the vicinity of 0.8 and not in the vicinity of 0.2. The point I want to make about the catch-all is that even this is not always possible. That is, sometimes the attitude the agent has towards the catch-all can't even be put into qualitative terms like "more probable than not", "unlikely" and so on. Thus, when I say that perhaps a catch-all should not be assigned a definite prior, I mean that the prior does not have a numerical value and is instead a variable in the interval $(0, 1)$.[37]

Let us for the moment focus on assigning credences to scientific hypotheses, as opposed to more mundane propositions. Perhaps we can assign a prior credence to the catch-all by considering how we assign credences to ordinary hypotheses and then emulating that for the catch-all. One line of thought takes credences in scientific hypotheses to be judgements of plausibility based on considerations of theoretic virtue (Shimony, 1970; Salmon, 1990). For instance, Salmon (1990) notes that:

> "Plausibility considerations are pervasive in the sciences; they play a significant - indeed, *indispensable* - role [...] Plausibility arguments serve to enhance or diminish the probability of a given hypothesis prior to - i.e., without reference to - the outcome of a particular observation or experiment" (p. 182)

That is good news, since considerations that enter in the assessment of the initial

[35] One might ask, who does the assigning - the agent themselves or the person analysing the agent? When I argued for a particular view of what the Problem of Awareness Growth is, I committed myself to taking the agential perspective seriously. Ultimately, if Bayesianism is to fulfil either of the three normative roles it can play, it should represent aspects of the agent that are of normative relevance faithfully. Since we are analysing the rationality of credences, we should thus match the agent's assignment. Hence, it is the agent doing the assigning, but the person analysing the agent follows the credal assignment of the agent.

[36] Anti-luminosity considerations (Williamson, 2000) would speak in favour of this view as far as the agent themselves knowing their credence. Moreover, it is doubtful that from a third-person perspective that we can discern the agent's credence so precisely either. Can a person analysing the agent really be able to say that that the agent's credence in a given proposition is 0.7500000001 and not 0.7500000002? I think that is implausible.

[37] Credence 0 is excluded by my definition of a catch-all; it needs to be a store of credence. Credence 1 is excluded since that would imply that all explicit hypotheses are assigned credence 0. While this can happen, this is not the case in situations that interest us, where agents have some explicit possibilities which they take to be live possibilities.

plausibility of hypotheses or theories - the so-called theoretic virtues - are well studied and known. A prominent list of such virtues includes accuracy, consistency, scope, simplicity and fruitfulness (Kuhn 1977, p. 322).

Of course the list itself is disputed. For instance, Laudan (1990) offers the following alternative list of theoretic virtues: internal consistency, surprising and correct predictions given background assumptions, testing against a diverse range of kinds of phenomena (p. 285). Even if the list of virtues was not disputed, there are still two outstanding issues, as Kuhn (1977) himself points out. First, the virtues are somewhat vague, so there is scope for interpreting one and the same virtue differently. This potentially leads to scoring one and the same hypothesis quite differently with respect to a given theoretic virtue, thus potentially resulting in a difference in how plausible it is judged to be. Simplicity is a good example of a vague virtue, as the application is ultimately a judgement call by the scientists. Suppose that our hypotheses of a given domain are just polynomial equations that we are trying to fit to the data. One way to understand the simplicity of polynomials is by number of unknown terms. Another way is to understand in terms of the highest degree of the polynomial.[38] Second, different virtues might pull in different directions for a given hypothesis. For instance, one hypothesis might have greater scope and be less simple, while the other one is simpler but narrower in scope. There is therefore room for weighing different virtues differently, prioritising some and downgrading others, thus resulting in a different plausibility judgement with respect to a given hypothesis.

Still, Bayesianism, as I have presented it, is permissive in assignment of prior credences. We don't need a procedure that yields a unique credence for a given hypothesis. Considering the theoretic virtues of a theory provides a plausible way to assign credences to hypotheses, which may nonetheless be different for different scientists. That is enough for our purposes.

Here is where the fact that the catch-all is not a definite scientific theory raises problems. Its content is entirely inscrutable, and it seems impossible to assess the theoretical virtues of something so inscrutable. To appreciate the extent of this inscrutability, think of the catch-all as a container with names of as-yet-unconceived hypotheses of a given domain. The first problem is that we do not know how big the container is, since we do not know how far into the future the list of explicitly formulated hypotheses will keep growing. The second problem is that we do not know how many names are in the container, though it is highly plausible that the container is not empty. The third problem is that even if we found out the number of names in the container, we still don't know the names. Fourth, even if we found out the names, we

[38]This example is taken from Nguyen and Marcoci (2019).

have absolutely no idea what stands behind those names.

Note that it is only in the fourth sense that the content of complex scientific hypotheses is inscrutable to a non-specialist. Even then, while the non-specialist does not know what the Standard Model of particle physics is, someone does. With the catch-all, no one knows the content of the next hypothesis, even if one were to somehow know the label. In sum, we don't know the make-up of the catch-all hypothesis in the sense of being able to even faithfully enumerate the sub-hypotheses in the catch-all; all the while the content of each sub-hypothesis remains inscrutable.

One might reply that while the content of each individual hypothesis making up the catch-all is inscrutable, that needn't prevent us from assigning them a credence with reference to theoretical virtues, or so contends Maher (1995). Recall that Maher has a real-valued theoretical virtue measure over possible hypotheses in a given domain, some of them explicitly formulated and some of them as-yet-unformulated. Those hypotheses that have been explicitly formulated have just a single theoretical virtue score. Those that have not been explicitly formulated have different possible theoretical virtue scores. Maher points out that we can have credences with respect to virtues of as-yet-unformulated theories: "I might think it unlikely that the next theory to be formulated will be simpler than any existing theory" (Maher 1995, p. 105).

Maher is surely right that we can have opinions about future theories that haven't yet been formulated. We might even suppose that for a given list of theoretic virtues, we can have opinions about future hypotheses with respect to all those virtues. However, some reflection on how we apply theoretical virtues to assess the plausibility of explicit hypotheses shows why we can't do the same for as-yet-unformulated theories, even if we can say some uncertain things about their theoretical virtues. Scientists exercise judgement when applying a given virtue, say simplicity, to a theory. It thus appears impossible to evaluate the simplicity of a future hypothesis without knowing its content because in such a case, it's undetermined just in what way a hypothesis is simple or not. Moreover, scientists exercise judgement when weighing different virtues against each other when considering a concrete hypothesis. This weighting, inescapable in the case of concrete hypotheses, is simply not possible with respect to a hypothesis we do not know the content of. Sure, we can assign different numbers to the different virtues and even implement some clever algorithm to aggregate them, but the potential sophistication of the formal proposal obscures the fact that when we assess the plausibility of hypotheses with reference to their theoretic virtues, we start from the content of the hypothesis and then consider the virtues against this content - something that is not possible in the case of the catch-all because we do not know its content.

The point to emphasise is not just that it's hard to assess the virtues of the catch-all

- that much can be said for an ordinary hypothesis as well. The point is that there is a significant additional layer of difficulty in assessing virtues of the catch-all that jeopardises an attempt to do so to any degree of plausibility. Thus it undermines the attempt to assign the catch-all a credence with reference to theoretic virtues.

But perhaps our method of assigning credences wasn't quite right. After all, I said that *one* line of thought takes credences in scientific hypotheses to be judgements of plausibility based on considerations of theoretic virtue. Another line of thought is that credences in scientific hypotheses are just dispositions to bet on the truth of scientific theories (Maher, 1993). Indeed, in light of the popularity of the betting interpretation of credences, this is surely the majority Bayesian view on how to assign credences to scientific hypotheses, since the interpretation is supposed to apply quite generally. Maher is (again) correct to point out that we can bet on the truth of scientific theories: "I might prefer the option of receiving a dollar if the next theory formulated is false, to the option of receiving the dollar if that theory is true" (Maher 1995, p. 105). This also plausibly sidesteps the issue of content of individual sub-hypotheses in the catch-all. We designate the next hypothesis in a given domain as Hyp_{n+1} and simply consider how we would bet on its truth, with no need to explicitly consider its content.

However, this proposal also faces problems. First, it strains credulity to think one has any opinions about the truth of Hyp_{n+7} or that one would be disposed to bet on the truth of Hyp_{n+7}. Surely, even if presented with a bet on Hyp_{n+7} one can say: "I have no idea". That is, one might be stably disposed to refuse to partake in any bets regarding the truth of Hyp_{n+7}. Second, this proposal runs into the other kind of trouble with the catch-all. We really don't know much about the space of as-yet-unformulated hypotheses. We don't know how many as-yet-unformulated hypotheses there are and we need to assign them all a credence to assign a credence to the entire catch-all. One might retort that the catch-all is subjective and doesn't need to get the number of future hypotheses right. But then at some point we will exhaust the catch-all and would have to resort to awareness growth by expansion. Thus, it is not clear why we bothered with the catch-all in the first place - we could have just expanded our algebra when a new hypothesis turned up.

But perhaps because the catch-all is so different, it needs a different approach - a meta-approach for a meta-hypothesis. Stanford (2006) has argued at length that we should not be convinced in the truth of our current scientific theories because the historic record gives us strong grounds to believe that at any given point in the history of science there are superior alternatives to current scientific theories that go unconceived. At last the solution appears in our grasp. Once we launch into the realism vs anti-realism debate, we will be sceptical of the exhaustiveness of our current hypotheses and will be disposed to bet heavily on the truth of the catch-all.[39]

The point I want to make in response is that while assigning the catch-all a high credence based on the history of science is plausible, it is doubtful that this serves as a basis for determinate betting dispositions in other cases. One thing to note at this point is that surely not all domains of scientific theorising are such that the realism vs anti-realism debate provides much guidance. After all, the realism vs anti-realism debate usually involves high-level scientific theories which feature unobservables, and it's the epistemic status of unobservables that is commonly taken to shape the warrant, or lack thereof, to believe in the truth of hypotheses that feature them. But this is not a feature of science more broadly.[40] For instance, consider hypotheses about the origins of syphilis. Syphilis itself is very well-studied at the level of diagnosis, treatment and microbiology. However, we are still unsure as to whether it was native to Europe in the 15th century (time of first outbreak) or it was brought over from the Americas, or some syphilis-antecedent was brought over from the Americas but evolved into syphilis in Europe. So we know a lot about syphilis but not everything. It doesn't appear that the same basis for assigning a high credence to the catch-all is available as compared to the basis available for so doing when betting on the catch-all, with respect to our theories of what, fundamentally, matter is. There's just less uncertainty about syphilis as a whole, than matter. The point isn't that we will be disposed to bet strongly on the falsity of the catch-all in the syphilis case, but rather that grounds for any determinate disposition are absent. Still, you might insist that some other grounds exist for determinate betting dispositions in the catch-all with respect to hypotheses about the origins of syphilis, other than thinking about the realism vs anti-realism debate. Let us even grant that such grounds exist in science quite generally, although what those grounds are possibly changes from case to case.

Let us take a step back from scientific hypotheses and consider more mundane scenarios. I am even more sceptical that such a basis always exists in non-scientific cases. Consider a variation on the dinner date case where Liudmila is convinced that she'll be taken to some really unexpected place for the anniversary date precisely because it is such a special occasion. The prior in the catch-all is high. But absent that description of the case, there are no grounds to assign any given credence to the catch-all. Liudmila could very well reply "I honestly have no idea" to the question of

[39] The realism vs anti-realism debate has been rendered of practical significance for science, a significant outcome that leaves ornithologists comparatively worse off.

[40] Stanford of course disputes the importance of the observable/non-observable distinction insofar as it affects the epistemic standing of scientific theories, but the broader point stands. Namely, that higher-level scientific theories have a different epistemic standing, be it due to the observable/non-observable distinction or something else.

whether she thinks she'll end up going somewhere she hasn't thought about. To put the point even more forcefully, Liudmila might be stably disposed to refuse all bets on the truth of the catch-all.

In sum, assigning a prior to the catch-all is indeed intractable if we think of prior credences as assessments of initial plausibility of propositions. We have more luck if we think of priors as dispositions to bet. Specifically, there appears to be a principled way to assign a (high) prior credence to catch-alls with respect to high-level scientific hypotheses because the history of science gives us reason to think that we haven't exhausted the space of possibilities with our explicit hypotheses. I am more sceptical that such a basis always exists for lower-level scientific hypotheses and for more mundane, non-scientific cases, though I admittedly can't rule out the possibility that it always exists. It is now time to turn to another problem with the catch-all: coming up with the likelihood of the evidence on the catch-all.

4.2.2 Likelihood on the catch-all

Another key term that involves the catch-all if it is to be included is the likelihood of the evidence on the catch-all. Recall, given a hypothesis Hyp and piece of evidence Evi, the likelihood is the conditional credence of the evidence given the hypothesis: $Cr(\mathsf{Evi}|\mathsf{Hyp})$.

The likelihood on the catch-all, while perhaps a more neglected worry, is treated with similar suspicion as the prior. Salmon (1990) at one point notes that: "the point to be emphasized right now is the utter intractability of the likelihood on the catchall" (p. 191). Salmon's reasons for this intractability are similar to those of why Wenmackers and Romeijn are sceptical of assigning the catch-all a determinate prior: "the seriously considered candidates are bona fide hypotheses, the catchall is a hypothesis only in a Pickwickian sense. It refers to all of the hypotheses we are not taking seriously, including all those that have not been thought of as yet" (p. 191). Roush (2005) states: "the probability of the evidence on the catch-all [...] we know how to evaluate this term for some modest hypotheses, it is extremely difficult to evaluate for a high-level theory" (p. 210).

Again let us begin by considering how one assigns likelihoods to scientific hypotheses. There are two cases where it is extremely easy. Suppose hypothesis Hyp logically entails, for some piece of evidence Evi, its negation. In such a case $Cr(\mathsf{Evi}|\mathsf{Hyp}) = 0$. Similarly, $Cr(\mathsf{Evi}|\mathsf{Hyp}) = 1$ if Hyp logically entails Evi.

Those are not the only cases where assigning likelihood on the catch-all is straightforward. Recall the physical probability function Pr. While a credence function Cr is distinct from a physical probability function Pr, I will assume, throughout this section, that they are related. Specifically, for a given hypothesis Hyp and piece of evidence Evi, the likelihood of the evidence on the hypothesis, $Cr(\mathsf{Evi}|\mathsf{Hyp})$, is set equal to the physical probability the hypothesis assigns to the evidence $Pr(\mathsf{Evi})$, where such a physical probability is available.[41]

To illustrate, suppose the piece of evidence is the number of larvae in a plot of land and further suppose that our hypothesis is that the distribution of larvae in the plot of land follows a Poisson distribution with bias parameter, say 0.2. This means that the probability of finding n larvae in a plot is given by $\frac{e^{-0.2}\times 0.2^n}{n!}$.[42] In cases like this, assigning the likelihood is straightforward: we just set the likelihood to match the probability assigned by the statistical hypothesis to the evidence.

But not all scientific hypotheses make direct pronouncements about the probability of the evidence. For instance, the hypothesis that a tenth of the patients in the hospital have Covid does not by itself imply anything about the probability of the next tested patient testing positive. Maybe you are testing a patient in the burns unit and the Covid patients are all in the pulmonology and infectious diseases wards. In such a case, some work is needed to extract a probability that the hypothesis assigns to the evidence. We go to assign a likelihood in conjunction with auxiliary hypotheses. In this case an auxiliary hypothesis is readily available; we can assume that Covid is evenly distributed amongst the patients of the hospital and then the main hypothesis, in conjunction with the auxiliary hypothesis, indeed assigns a probability of 0.1 to the next patient being tested testing positive. You then set the likelihood equal to this probability and proceed in the usual Bayesian manner.

Of course not all cases of extracting a probability from a non-statistical hypothesis are this straightforward. The hypotheses can be very high-level and pieces of evidence very localised. For instance what probability does the Standard Model assign to the observation of a B_s meson decaying into two muons in the Large Hadron Collider? That depends, amongst other things, on an inordinate number of auxiliary assumptions about the equipment in the Large Hadron Collider.

I will assume throughout that for explicit hypotheses, we can always extract a definite probability of the evidence from a hypothesis and then assign a likelihood on the hypothesis by setting it equal to that probability. Of course this is not always easy and is plagued with difficulties, perhaps insurmountable ones.[43] However, my point

[41] This is most closely related to Miller's Principle (Miller, 1966), among others.

[42] This example is taken from Gillies (2001).

[43] Note that Wenmackers and Romeijn seem to assume that we can always perform this manoeuvre. They assume that given any hypothesis we can always, perhaps with great ingenuity, assign a likelihood to the evidence on the hypothesis. At least that seems a more plausible interpretation than taking them to restrict the scope of their proposal to just those cases where we can make the hypotheses

here is not to argue that producing likelihoods is easy but to argue that however hard it may be for explicit hypotheses, it is categorically harder for the catch-all, so hard in fact that we should not assign it a definite likelihood at all.

Let us begin from the good case for proponents of the catch-all due to Roush (2005). Suppose our explicit hypothesis is that Mary is pregnant. The catch-all hypothesis is then equivalent to the hypothesis that Mary is not pregnant. The piece of evidence is that Mary's pregnancy test comes out positive. The likelihood on the explicit hypothesis is set to match the sensitivity of that particular test, and the likelihood on the catch-all is set to equal the rate of false positives of that particular test. So far, so good.

But now consider some high-level hypothesis, such as the Standard Model of particle physics. For simplicity suppose that it is the only explicit hypothesis of the domain, and then further suppose that our piece of evidence is the observation of the aformentioned decay of the B_s meson into two muons in the Large Hadron Collider. The catch-all contains as-yet-unformulated hypotheses. In the case of the catch-all to the Standard Model, even if we could enumerate all the different sub-hypotheses included in the catch-all, we don't know anything definite about many of them. If the catch-all includes the true hypothesis, granted the true hypothesis assigns probability 1 to the evidence. But the catch-all also features many other hypotheses that aren't the true one and are not even co-extensional with it with respect to the probability it assigns to the evidence. It is for those latter catch-all sub-hypotheses that we can't possibly determine the probability they assign to the evidence - we don't know anything about them other than that they are not the correct hypothesis. Determining the likelihood on a hypothesis by setting it equal to the probability that the hypothesis assigns to the evidence seems to be the only way to proceed. Thus, I am left to conclude that at best we cannot proceed in the same way with the catch-all as we did with the explicit hypotheses.

Perhaps we again need to abandon all attempts to treat the catch-all in the same manner as the explicit hypotheses. We black-box the content of the catch-all and assign it a likelihood in a way that differs from how we assign likelihoods on explicit hypotheses. One suggestion (without endorsement) is that we just take the average of the likelihoods on the explicit hypotheses and set the likelihood on the catch-all equal to that average (Strevens, n.d.).[44] The formal tractability of this proposal obscures how justifiable it is. Note that newly introduced hypotheses can assign a drastically different likelihood to the evidence than their older competitors. For instance, the Special Theory of Relativity assigned a credence of 1 to the piece of evidence that

statistical.

[44] Strevens himself quickly dismisses this suggestion.

the speed of light in a vacuum is constant in all inertial frames, since this piece of evidence is a postulate of the Special Theory of Relativity. Prior to the introduction of the Special Theory of Relativity, however, the average likelihood for the evidence was less than 1. Are we really to expect that the course of history will be such that it will correct the change in the average brought about by the introduction of the Special Theory of Relativity? Not only is this expectation baseless, it also privileges the average at some point in time. Why the average now, as opposed to when the next hypothesis is introduced? In sum, I do not think that assigning a likelihood on the catch-all is always possible.

To make an intermediate conclusion, an approach to the catch-all that seeks to be fully general should not assign a likelihood and prior to the catch-all, for that is not always possible. Thus, any catch-all proposal that assigns an indefinite prior and likelihood is in a privileged position with respect to other catch-all proposals. Having argued for this particular datum I now intend to further privilege a catch-all proposal due to Wenmackers and Romeijn (2016) from amongst extant proposal that utilise a catch-all. I will do so on the grounds that their proposal has an indefinite prior and likelihood in the catch-all and moreover, solves some additional problem with the catch-all. Let us now turn to these additional problems.

4.2.3 Earman's complaints

Earman has two complaints about the catch-all in conjunction with the shaving-off idea.[45] First, because $Cr^+(\mathsf{CA}') = Cr(\mathsf{CA}) - Cr^+(\mathsf{Hyp}_3)$ this leads to "the assignment of ever smaller initial probabilities to successive waves of new theories until a point is reached where the new theory has such a low initial probability as to stand not much of a fighting chance" (Earman 1992, p.196).[46] Second, the dynamics are a problem because sometimes an introduction of a new theory (radically) reduces credences in older theories. The example Earman cites is that Einstein's Special Theory of Relativity taking away credence from its competitors, theories of Lorentz and Abraham which where set in classical space-time (Earman 1992, p.197). However, as pointed out earlier we have that $Cr^+(\mathsf{Hyp}_i) = Cr(\mathsf{Hyp}_i)$ for $i = 1, 2$.

A proposal due to Wenmackers and Romeijn (2016) offers a way out. In their framework we first assign priors, $Cr(\mathsf{Hyp})$, and likelihoods $Cr(\mathsf{Evi}|\mathsf{Hyp})$, to individual hy-

[45]Recall that the catch-all provides a solution to the Problem of Awareness Growth only in conjunction with a suitable rule for revising credences in light of awareness growth, so we have to consider the catch-all as part of a package deal.

[46]Recall the result from the first section of this chapter. The credence in the new catch-all is just the credence in the original catch-all minus the credence in the new explicit hypothesis Hyp_3.

potheses $\mathsf{Hyp} \in \mathcal{H}_n$ while ignoring the catch-all.[47] The catch-all proposition is assigned an indefinite prior, $Cr(\neg \bigvee \mathcal{H}_n) \in (0,1)$, and an indefinite likelihood, $Cr(\cdot|\neg \bigvee \mathcal{H}_n)$. As argued at length before, this is only right as far as a general proposal of how to treat the catch-all is concerned. Having assigned the indefinite prior and likelihood, we then normalise the priors we set while ignoring the catch-all by multiplying them by $1 - Cr(\neg \bigvee \mathcal{H}_n)$. This ensures that the priors in the explicit hypotheses and the catch-all sum to one.

Wenmackers and Romeijn then apply their proposal to the example of a hygiene inspector coming to a restaurant with a surprise visit. The food inspector goes into a restaurant and orders 5 dishes. She will test these dishes for Salmonella. Her credence in the proposition that a dish tests positive given that the restaurant is taking the health and safety precautions required by law is 0.01. Her credence in the proposition that a dish tests positive given that the restaurant is not taking the health and safety precautions required by law is 0.2. As it happens, all 5 dishes test positive. This leads the food inspector to consider a new hypothesis: the testing kit is contaminated.

Formally, the set-up of the case is as follows. The evidence space is given by $\mathcal{E} = \{\mathsf{Evi}_1, \neg\mathsf{Evi}_1\} \times ... \times \{\mathsf{Evi}_5, \neg\mathsf{Evi}_5\}$ where Evi_i is the proposition that ith dish tests positive for infection with Salmonella. For ease of reference we will denote the proposition that all 5 tests come out positive as Evi. The hypothesis space $\mathcal{H}$ is the vast set of all possible probability functions over $\mathcal{E}$. The inspector is only explicitly considering two hypotheses Hyp_1 and Hyp_2. So $\mathcal{H}_n = \{\mathsf{Hyp}_1, \mathsf{Hyp}_2\}$. Hyp_1 is the probability function corresponding to a binomial distribution on the evidence space with bias parameter 0.2. Informally we can think of Hyp_1 as the hypothesis that the test coming out positive is down to a violation of health standards by the restaurant. Hyp_2 is the probability function corresponding to a binomial distribution on the evidence space with bias parameter 0.01. Informally we can think of Hyp_2 as the hypothesis that the test coming out positive is down to the restaurant in some way, but for some reason other than a violation of health standards by the restaurant. In line with the assumption that the likelihood of the evidence is set to equal the physical probability the hypothesis assigns to the evidence, the likelihoods are $Cr(\mathsf{Evi}|\mathsf{Hyp}_2) = 0.01^5$ and $Cr(\mathsf{Evi}|\mathsf{Hyp}_1) = 0.2^5$ respectively. Suppose that the priors, without the catch-all, are

[47] Wenmackers and Romeijn (2016) actually present two proposals. The one I discuss is called *vocal* open minded Bayesianism. The other one is *silent* open minded Bayesianism. However, the former suffices to solve two problems for the catch-all that Earman raises and neither of the two proposals assigns a definite prior and likelihood to the catch-all. Thus, both proposals (tacitly) admit the other two problems with using a catch-all. Hence, I am not doing a disservice to Wenmackers and Romeijn, vís-a-vís their implementation of a catch-all proposition. Moreover, we will encounter silent open minded Bayesianism under the guise of Awareness Rigidity in Chapter 5.

$Cr(\mathsf{Hyp}_1) = Cr(\mathsf{Hyp}_2) = \frac{1}{2}$.

The catch-all in this case is $\neg(\mathsf{Hyp}_1 \vee \mathsf{Hyp}_2)$, the set of all probability functions over the evidence space other than Hyp_1 and Hyp_2. We set the prior, $Cr(\neg(\mathsf{Hyp}_1 \vee \mathsf{Hyp}_2)) = r_n$, to be some unspecified $r_n \in (0,1)$ and normalise previous priors $Cr(\mathsf{Hyp}_1) = Cr(\mathsf{Hyp}_2) = (1 - r_n)\frac{1}{2}$. Also let the likelihood of the evidence given the catch-all, $Cr(\mathsf{Evi}|\neg(\mathsf{Hyp}_1 \vee \mathsf{Hyp}_2)) = x_n(\cdot)$, be some unspecified $x_n(\cdot)$ Because Hyp_1, Hyp_2 and $\neg(\mathsf{Hyp}_1 \vee \mathsf{Hyp}_2)$ are mutually exclusive and when taken together (and only when taken together) are collectively exhaustive, we can calculate $Cr(\mathsf{Evi})$ using the law of total probability: $Cr(\mathsf{Evi}) = Cr(\mathsf{Evi}|\mathsf{Hyp}_1)Cr(\mathsf{Hyp}_1) + Cr(\mathsf{Evi}|\mathsf{Hyp}_2)Cr(\mathsf{Hyp}_2) + r_n x_n(\mathsf{Evi})$. At this point two things are worth noting. First, we do not have a numerical value for the final term. Second, we have to have a numerical value for both the likelihood of the catch-all and its prior in order to assign a numerical value to the posterior credence of any hypothesis, in particular the explicitly formulated ones. This is the case because for $i = 1, 2$ we have:

$$Cr^+(\mathsf{Hyp}_i) = \frac{Cr(\mathsf{Evi}|\mathsf{Hyp}_i)Cr(\mathsf{Hyp}_i)}{Cr(\mathsf{Evi}|\mathsf{Hyp}_1)Cr(\mathsf{Hyp}_1) + Cr(\mathsf{Evi}|\mathsf{Hyp}_2)Cr(\mathsf{Hyp}_2) + r_n x_n(\mathsf{Evi})}.$$

While we can't exactly evaluate $Cr^+(\mathsf{Hyp}_1)$ and $Cr^+(\mathsf{Hyp}_2)$ numerically, we can evaluate the ratios of the posteriors in the two explicit hypotheses $\frac{Cr^+(\mathsf{Hyp}_1)}{Cr^+(\mathsf{Hyp}_2)} = \frac{Cr(\mathsf{Evi}|\mathsf{Hyp}_1)Cr(\mathsf{Hyp}_1)}{Cr(\mathsf{Evi}|\mathsf{Hyp}_2)Cr(\mathsf{Hyp}_2)} = \frac{0.2^5 \times 0.5}{0.01^5 \times 0.5} = \frac{0.00016}{0.5 \times 10^{-10}} = 3200000$. We can see the hypothesis that there was a violation is judged to be comparatively much more likely. However, in light of the fact the likelihoods of explicitly considered hypotheses are so small, the introduction of Hyp_3 - a new explicit hypothesis - seems warranted.

Hyp_3 is the probability function corresponding to the binomial distribution over the evidence space with bias parameter 1. Intuitively, it is the hypothesis that the tests themselves are contaminated. So $\mathcal{H}_{n+1} = \{\mathsf{Hyp}_1, \mathsf{Hyp}_2, \mathsf{Hyp}_3\}$. We then retroactively assign Hyp_3 a prior, $Cr(\mathsf{Hyp}_3)$. The credence in the original catch-all r_n is now equal to $Cr(\mathsf{Hyp}_3) + r_{n+1}$ where r_{n+1} is the prior in the new catch-all $\neg(\mathsf{Hyp}_1 \vee \mathsf{Hyp}_2 \vee \mathsf{Hyp}_3)$, which was also assigned retroactively. Similarly, the likelihood of the evidence given Hyp_3 is $x_{n+1}(\cdot)$. After having normalised priors for Hyp_1 and Hyp_2 in light of this new catch-all, under some minimal assumption Wenmackers and Romeijn obtain $\frac{Cr^+(\mathsf{Hyp}_3)}{Cr^+(\mathsf{Hyp}_1)} = 3125$.[48] That is, the new hypothesis is comparatively much better confirmed by the evidence than the old hypothesis which was best confirmed by the evidence.

How does this answer Earman's complaints? Earman's first complaint was that

[48] See pages 1238 to 1241 of Wenmackers and Romeijn (2016) for details.

because we are constantly shaving off credence from the catch-all to assign to new hypotheses, there is less reserve credence for each subsequent new hypothesis than before. Eventually, new hypotheses will be assigned too small of a credence to stand a chance to become the best confirmed hypothesis, for we have drained the well of reserve credence. The one crucial part of Earman's complaint that doesn't come true is that new hypotheses don't stand a chance. They do stand a chance because priors in the old hypotheses are normalised at each point a new hypothesis is introduced, and new hypotheses can receive comparatively stronger confirmation than the old hypotheses. It is this comparative dynamic which also assuages Earman's second complaint: novel hypotheses can come to comparatively dominate old hypotheses in a way that's at least qualitatively consistent with the way Einstein's Special Theory of Relativity came to comparatively dominate theories of Abraham and Lorentz.

For all of the negative points that have been made thus far, we now have a positive result - there is a catch-all proposal that sidesteps the prior and likelihood issues by simply assigning an indefinite prior and likelihood to the catch-all. Moreover, it solves the two problems for the catch-all raised by Earman.

That said, independent objections to the proposal of Wenmackers and Romeijn could be raised. It might be objected that the catch-all that Wenmackers and Romeijn use is implausible. Recall, on their proposal the catch-all includes all the probability functions over the evidence space that aren't the probability functions corresponding to the explicit hypotheses. Indeed, an uncountably large set of probability functions in the agent's space of possibilities would look to jeopardise each of the three normative roles that Bayesianism seeks to play. However, there is a simple fix as we needn't think of all hypotheses as statistical and thus, we needn't use the same catch-all as Wenmackers and Romeijn. Instead of taking the set of hypotheses to be the set of all probability functions on the evidence space, we can take $\mathcal{H} = \{\mathsf{Hyp}_1, \mathsf{Hyp}_2, ..., \mathsf{CA}\}$. Thus, the set of basic propositions becomes $\mathcal{H} \times \mathcal{E}$, where $\mathcal{E}$ is a set of possible pieces of evidence. The catch-all is no longer an uncountable set of probability functions but a much simpler proposition. We can then proceed as Wenmackers and Romeijn do: when a new hypothesis gets formulated we shave it off retroactively assign it a prior and likelihood, normalise the priors in all the pre-existing hypotheses and then conditionalise on existing evidence.

Another issue one might take with their proposal is that the end result does not allow for a numerical evaluation of posteriors; instead, we only have the ratio of posteriors. Why might we find this unsatisfying? Because it (only) allows us to say that some hypothesis is better confirmed by the evidence than another, or that a hypothesis is best confirmed amongst a set of hypotheses, but no more. Thus, evaluation of hypotheses becomes comparative, and if one chooses to read credences as assessments of plausibility, this assessment becomes relative to the hypotheses being considered. Notably, one is no longer able to say that the best confirmed hypothesis is even somewhat plausible, for a high ratio of posteriors is consistent with the posteriors themselves being tiny. This is indeed what we saw for the posteriors in the food inspection case if we run it without the catch-all.

To make matters worse for the catch-all, we can derive a qualitatively similar result to Wenmackers and Romeijn without using a catch-all. In what follows I do not outline this alternative proposal but instead, focus on the results it produces. Bradley (2017) applies his own proposal to the food inspector example to show that the food inspector thinks that the hypothesis that her testing kit is contaminated is 20 times more likely than the hypothesis that the restaurant had violated regulatory standards.[49] Moreover, Bradley's proposal allows for calculation of numerical posteriors for all the hypotheses since he doesn't rely on the catch-all. Thus, it appears as though everything that we can do with the catch-all we can do without it, and we can do more since we can produce numerical posteriors.

However, two points can be made in reply. First, in some contexts a comparative assessment of hypotheses reflects the (putatively) comparative nature of the issue at hand. Salmon (1990) argues that the ratio of posteriors is the appropriate metric for hypothesis choice since the post-Kuhnian consensus is that the very nature of hypothesis choice is comparative:

> "Kuhn has often maintained that in actual science the problem is never to evaluate one particular hypothesis or theory in isolation; it is always a matter of choosing from among two or more viable alternatives[...] On this point I think that Kuhn is quite right" (Salmon 1990, p. 191).

Thus, a ratio of posteriors looks like the appropriate metric in the context of hypothesis choice. Still, *pace* Salmon you might say that the nature of hypothesis choice is not comparative or point to some other context in which the ratio of posteriors looks inappropriate.

There is a second point to be made, however. The mere fact that a proposal produces numerical posteriors does not imply that we can interpret those posteriors

[49]Note that Bradley uses different assumptions to those of Wenmackers and Romeijn. For one, he uses different priors. Additionally, in his version of the case the new hypothesis is introduced after the second test, and then it's the evidence from the subsequent three tests against which the three hypotheses are tested, not all five. The central point still stands. We have proposals that don't use the catch-all which can achieve the result that the new hypothesis is judged to be much more likely than the best confirmed old hypothesis.

to mean something like absolute commitment or absolute assessment of plausibility. Bradley says: "[w]hat our current opinions reflect are the relative plausibility and desirability of the prospects that we are aware of" (Bradley 2017, p. 255). Hence, while in his framework we can obtain numerical posteriors, at the level of interpretation they are no different to their ratios on the proposal by Wenmackers and Romeijn. Thus, the face-value deficiency of not having numerical posteriors and only having their ratio, is not a deficiency as such, since it's a feature that lies below the surface of proposals that do give us numerical posteriors. Moreover, the ratio of posteriors is plausibly just what's needed in some contexts. In sum, we have a positive result for the catch-all hypothesis.

However, we have thus far been assuming that the catch-all is included in the agent's algebra. But are there any reasons for including it? I will now turn to arguments that purport to show that it should always be included and argue that they fail.

4.3 Arguments for the catch-all

Our first argument for the view that a catch-all should be included in the agent's algebra comes from a proposal that, ironically, does not use a catch-all (Strevens, 2010). In short, Strevens' idea is to re-work the space of possibilities at each point awareness grows, re-assign priors and then conditionalise on the previously received evidence. He does not include a catch-all and says that this omission is tantamount to a "failure to represent a scientist's (presumably non-zero) subjective probability that some as-yet unknown theory is correct" (Strevens, 2010).[50] If a scientist, or layman for that matter, has a positive credence in the catch-all, then of course their algebra should include the catch-all and it should be assigned the requisite positive credence. However, absent additional reasons, that only amounts to an argument for the following claim.

Sometimes Catch-all. An agent's set of basic propositions $\mathcal{X}$ sometimes ought to include a catch-all proposition CA, and in those cases it should be assigned a positive credence.

The times when the agent has a positive credence in the catch-all are just the times when it should be included. That is the argument that Strevens seems to be proposing. Note that because I take the agential perspective as the main departure point for representing the agent's space of possibilities, I am already committed to including the

[50]He himself responds to this worry by saying that he is using the Bayesian framework to represent only a part of the scientist's epistemic state.

catch-all more often than Strevens seems to suggest we should. Namely, it is enough for the agent to assign the catch-all any credence, including a zero credence, for me to include it in the algebra. Of course that still doesn't amount to an argument that we should *always* include the catch-all in the agent's algebra, because sometimes, indeed often, the agent does not even consider the catch-all. In the dinner date cases Liudmila didn't consider the possibility that none of the explicitly formulated possibilities are the right one. You might of course say that I just gerrymandered the cases to suit my purposes, but ask yourself - do you really always formulate a none-of-the-above option when deliberating? I think not. The very normative force of the proposal to always include the catch-all derives from the fact that we do not always do it.

Elsewhere, Strevens does state a further reason as to why we should include the catch-all in the agent's set of basic propositions. Talking in the context of proposals that do not include a catch-all, Strevens says:

> "If there is a problem with this approach, it is that, because your priors for the competing hypotheses at any time sum to one, you are committed to taking a bet at very unfavorable odds that no plausible new theory will turn up" (Strevens n.d., p. 132)

The argument seems to be as follows: agents should not only include a catch-all proposition in their algebra, they should assign it positive credence. Otherwise we can present them with a bet that gives away their life's possessions if none of the hypotheses they are explicitly considering are true and gives them nothing if one of them is true, and they would accept it. This argument strikes against two kinds of people: those, who did not include a catch-all in their space of possibilities and those who did, but assigned it a zero credence. Since we are interested in the question of whether we should include a catch-all at all, we will look at how the argument works against people who failed to consider the catch-all.

Take Liudmila in the dinner date cases. Suppose we offer her a bet that gives her nothing if she is indeed taken to either an Indonesian or Surinamese or Italian restaurant for the anniversary date, and takes her life's possessions if she's taken elsewhere. I agree that Liudmila would do well to reject the bet. However, it appears that she must accept the bet, as she did not consider the catch-all and therefore did not assign it a non-zero credence. But note that "the proposal of a bet may inevitably alter [...] state of opinion" (Ramsey 1926, 68). The problem with the argument above as an argument that we should always include the catch-all proposition is that offering a catch-all as part of a bet makes Liudmila consider the catch-all; it makes her aware of it. Of course once she has considered it, the bet does not look good and she's free to reject the bet.

Thus, the argument above involves a growth in awareness that occurs across time and the argument cannot be used as an argument to include the catch-all pre awareness growth. Thus, Strevens (n.d.) really offers an argument for the following claim.

Catch-all awareness. An agent's set of basic propositions $\mathcal{X}^+$ post awareness growth sometimes ought to *come to* include a catch-all proposition CA, and CA should be assigned some positive credence.

To reiterate, the argument above is not an argument that says CA should be in $\mathcal{X}$, the original set of basic propositions. Thus, we still do not have an argument that a catch-all proposition should always be included in the agent's algebra.

The second argument in favour of the inclusion of the catch-all is due to de Canson (2020). Her argument for the inclusion of the catch-all comes in two steps. First, she argues for a particular credal norm which says that one shouldn't assign credence 0 or 1 to propositions that are only knowable a posteriori and that are not part of one's evidence. The second step is to use this norm to argue that we should include the catch-all proposition in the algebra of propositions, or so she claims.

Let us focus on the second step. Suppose that we should not assign credence 1 or 0 to propositions that are only knowable a posteriori and that are not part of one's evidence. For Liudmila, at the beginning of the dinner date cases, that entails that her credence function $Cr(\cdot)$ ought to be such that $0 < Cr(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}) = x < 1$. We should then assign credence $1 - x$ to $\neg(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur})$, as that follows by the Axioms of Probability. According to de Canson, $\neg(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur})$ is a catch-all proposition and so it follows that we should assign the catch-all proposition a positive credence. This already should give us pause since de Canson wanted to argue "that a catch-all proposition must be included in an agent's algebra" (de Canson 2020, p. 89) and she managed to go a step further and argued that we should also assign it a positive credence.

From my analysis of the case it follows that $\neg(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur})$ is just $\bot$ and we must assign it a zero credence. Correspondingly, $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$ is just $\top$ and we must assign it credence 1. Thus, de Canson must include some other basic propositions in $\mathcal{X}$ to ones that I do. Note that in order for $\neg(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur})$ to be interpretable as a catch-all proposition, that means that the space of possibilities must have included at least all the propositions that could possibly make up the catch-all. However, since de Canson takes $\neg(\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur})$ to be the catch-all proposition, the catch-all proposition has been in the agent's algebra all along and she just provided an argument for assigning it a positive credence.

Even if de Canson does not provide an argument for *including* a catch-all proposition in the agent's algebra, some of what she says threatens my analysis of the dinner date cases, namely the putative norm of rationality which she argues for: one ought not to assign credence 1 or 0 to propositions that are only knowable a posteriori that are not part of one's evidence. Since the proposition $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$ is not part of Liudmila's evidence and is only knowable a posteriori, if this norm is true then, contrary to what I said $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$ should not be assigned credence one.

The argument that de Canson herself thinks is successful in arguing for the norm takes as premises the conception of epistemology as a means-ends endeavour "on which the agent deploys her means to achieve the ultimate epistemic aim, that of determining what is the case" (de Canson 2020, p. 95). It is also assumed by de Canson that the agent has two means of inquiry - observation and reason - and moreover, the agent is faultless with respect to those two means of inquiry. This is to say she has determined all that was determinable a priori and knows that she has. Furthermore, she is in the position to know what can be settled by observation, and once she observes a proposition that can be so settled, she is certain of the proposition. Why is assigning a credence of 1 to $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$ unacceptable on this line of thought? Because, says de Canson, Liudmila did not settle on $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$ on the basis of her means of inquiry. But, I don't think she has settled as to whether $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$, she simply assumed it for now. At the beginning of the dinner date cases those are all and the only possibilities she has considered. Even if assigning credence of 1 to $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$ was the end of investigation as opposed to the beginning, recall that I conceive of investigation as localised. For instance, once we performed a series of experiments and updated credences in the different hypotheses accordingly, that is far from the end. Just because some hypothesis is now lent a high credence, does not mean it will always stay that way. Similarly, there's a lot more that can happen after we assign a credence of 1 to $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$. We can become aware of new salient possibilities and revise our beliefs accordingly. For instance, as we saw, Reverse Bayesianism allows a credence of less than 1 to be assigned to $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$ after Liudmila's awareness grows to include Mex.

Thus, it is far from the case that upon assigning Liudmila has settled on anything. Even if she settled on whether $\mathsf{Ind} \vee \mathsf{Ita} \vee \mathsf{Sur}$, she didn't settle simpliciter but settled provisionally. The root difference between me and de Canson seems to be how we conceive of the space of possibilities and the corresponding picture of the kind of agent we are analysing. I previously noted that de Canson seems to include all the necessary propositions that make up the catch-all in the agent's algebra. Because she does not relativise the space of possibilities to the issue at hand, and the space of possibilities included all the propositions that make up the catch-all with respect to some specific issue, it is reasonable to infer that the space of possibilities includes a great deal of propositions. Plausibly, all those propositions that are actually relevant to the whole course of an agent's life.[51] Another point in favour of this interpretation is that this allows us to interpret a credence of 1 in a proposition as settling, in absolute terms, as to whether that proposition is true. The norm of rationality which de Canson argues for also looks eminently plausible. If one is to interpret a credence of 1 in a proposition as settling on the truth of said proposition once and for all, we shouldn't do so for propositions that are only knowable a posteriori and that we haven't observed. I simply do not equate a credence of 1 with settling once and for all or with absolute certainty in the truth of a proposition.

4.4 Conclusion

In sum, we have found that assigning the catch-all proposition a prior credence does not look to always be possible, and assigning the likelihood of evidence on the catch-all is at least sometimes intractable. Having argued for this particular datum I privilege a catch-all proposal due to Wenmackers and Romeijn (2016) from amongst extant proposal that utilise a catch-all. I did so on the grounds that their proposal has an indefinite prior and likelihood in the catch-all and moreover, solves some additional problem with the catch-all. I then shored up this result by first raising two problems against the proposal of Wenmackers and Romeijn (2016) and then defusing them.

However, I also showed that the two extant arguments for always including the catch-all fail. Thus, if one wants to include a catch-all proposition there is a way to proceed but there's no general requirement to always include it. I now turn to Reverse Bayesianism, a putative norm for revising credences in light of awareness growth.

[51] I previously argued against this way of conceiving the space of possibilities for Bayesianism quite generally; its at odds with the three normative roles that Bayesianism as a theory of rationality seeks to play. But even if one was not persuaded by me, then the distinction between two different ways of conceiving of the agent's space of possibilities is nonetheless salient.

5 Reverse Bayesianism

Recall the statement of Reverse Bayesianism.

Reverse Bayesianism. For all $a, b \in \mathcal{X}$ such that $Cr(a) > 0$ and $Cr(b) > 0$, $Cr^{+}(\cdot)$ over $\mathbb{B}_{\mathcal{X}^{+}}$ ought[52] to be such that:

$$\frac{Cr(a)}{Cr(b)} = \frac{Cr^{+}(a)}{Cr^{+}(b)}.$$

The informal idea is that the ratio of credences in the basic propositions that were there before awareness growth stays the same. However, as things stand, Reverse Bayesianism is incomplete as it lacks an appropriate criterion for identifying propositions across awareness growth. Recall $\mathsf{DinnerDate1}$ where Ind disappears in favour of Bal and Jav and $\mathsf{Bal}, \mathsf{Jav} \notin \mathcal{X}$. Intuitively, the spirit of Reverse Bayesianism would demand that $\frac{Cr(\mathsf{Ind})}{Cr(\mathsf{Ita})} = \frac{Cr^{+}(\mathsf{Bal} \vee \mathsf{Jav})}{Cr^{+}(\mathsf{Ita})}$. This is because Ind was merely cut into two and intuitively $\mathsf{Bal} \vee \mathsf{Jav}$ is associated with an element of the old algebra, namely Ind.

Another issue is that the negation of one and the same proposition is different between $\mathbb{B}_{\mathcal{X}}$ and $\mathbb{B}_{\mathcal{X}^{+}}$. Recall $\mathsf{DinnerDate2}$. Originally $\neg\mathsf{Ita}$ is equivalent to $\mathsf{Ind} \vee \mathsf{Sur}$. However, after Liudmila becomes aware of Mex, $\neg\mathsf{Ita}$ becomes equivalent to $\mathsf{Ind} \vee \mathsf{Sur} \vee \mathsf{Mex}$. This realisation pushes some to say that the proposition Ita pre awareness growth is a different proposition from Ita post awareness growth (de Canson, 2020; Steele and Stefánsson, 2021b). While the identity of propositions across awareness growth is not my concern per se, to avoid confusion I will use capital letters for propositions in the old algebra (Ita) and lower case letters for propositions in the new algebra (ita). Strictly speaking, that implies that the two algebras do not share any propositions in common, but that is not an issue if we supply a way to identify propositions from the old algebra with those in the new algebra.

Roussos (2020) supplies exactly this sort of identification procedure via the notion of an embedding. Let $\mathbb{B}_{\mathcal{X}} = \langle \mathcal{X}^{cl}, \vee, \wedge, \neg, \top, \bot \rangle$ and $\mathbb{B}_{\overline{\mathcal{X}}} = \langle \overline{\mathcal{X}}^{cl}, \overline{\vee}, \overline{\wedge}, \overline{\neg}, \overline{\top}, \overline{\bot} \rangle$ be two Boolean Algebras and consider the following notion.

Embedding.[53] A function $h : \mathbb{B}_{\mathcal{X}} \to \mathbb{B}_{\overline{\mathcal{X}}}$ is an embedding iff it is total, injective[54] and for all $a, b \in \mathcal{X}^{cl}$:

[52]Recall a point I made at the end of Chapter 2, in Chapter 5 the "ought" is an action-guiding "ought" as opposed to an evaluative "ought", or the "ought" we use to attribute praise and blame.

[53]Definition taken from Davey and Priestley (2002).

[54]A function $f : X \to Y$ is said to be total iff for every $x \in X$ there is a $f(x) \in Y$. A function is said to be injective iff $f(x_1) = f(x_2)$ implies that $x_1 = x_2$.

$$h(a \vee b) = h(a) \overline{\vee} h(b) \quad \text{and} \quad h(a \wedge b) = h(a) \overline{\wedge} h(b).$$

Why, one might ask, is the notion of an embedding an appropriate criterion by which to identify propositions across awareness growth? The fact that an embedding is total and injective makes sure that we can "find" every proposition from the old algebra in the new algebra and we can find, for each old proposition, exactly one new proposition.

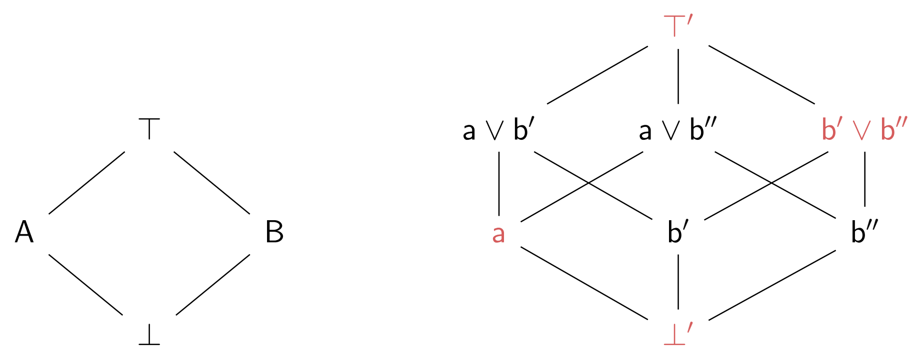
The other two conditions in the definition ensure disjunction and conjunction preservation. But equally important is what the notion of an embedding leaves out: negation and bound preservation. There is no condition which says that $h(\neg a) = \overline{\neg} h(a)$ or that $h(\top) = \overline{\top}$ and $h(\bot) = \overline{\bot}$. The absence of the first disjunct is desirable since, as mentioned, the negation of an element in the old algebra is different in the new algebra. Having no bound preservation just means that the top element of the old algebra need not be the top element of the new algebra, and the same goes for the bottom element. The fact that the top element of the old algebra needn't be mapped to the top element of the new algebra is desirable. For instance, in $\mathsf{DinnerDate2}$ we would want $\mathsf{Ind} \vee \mathsf{Sur} \vee \mathsf{Ita}$ to be mapped to $\mathsf{ind} \vee \mathsf{sur} \vee \mathsf{ita}$ and not to the top element of the new algerba, $\mathsf{ind} \vee \mathsf{sur} \vee \mathsf{ita} \vee \mathsf{mex}$.

Still, the notion of an embedding is a purely formal notion and, as will be seen, there are many ways to embed the pre awareness growth algebra into the post awareness growth algebra, much is left at the discretion of the person analysing the case. However, I would like to illustrate how an embedding is to be specified in the case of refinement and expansion respectively.[55]

Let us consider a case of refinement where the old set of basic propositions $\{\mathsf{A}, \mathsf{B}\}$ turns into $\{\mathsf{a}, \mathsf{b}', \mathsf{b}''\}$, upon the refinement of B. The embedding I choose for now is as follows: $h(\top) = \top'$, $h(\bot) = \bot'$, $h(\mathsf{A}) = \mathsf{a}$ and $h(\mathsf{B}) = \mathsf{b}' \vee \mathsf{b}''$. In Figure 3 you can see the old algebra and the algebra post awareness growth with the elements to which the embedding maps element of the old algebra highlighted appropriately.

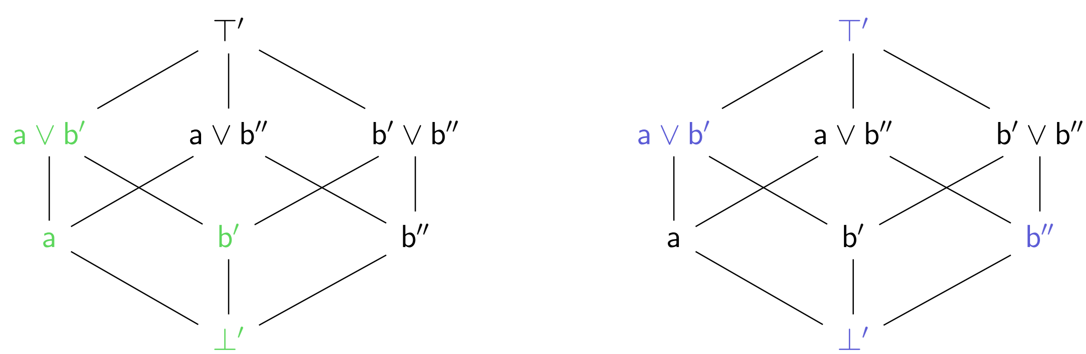
[55]A general theme we have thus far encountered and will continue to encounter in this chapter is that it is hard to predict just how awareness will grow, and as a result I am wary of claiming that we have fully general ways for coming up with an embedding for each instance of awareness growth. Below I provide an embedding for standard cases of awareness growth by refinement and expansion. Note that all cases that have appeared and will appear in this thesis are standard, but I am open to the possibility that there are some non-standard cases.

Figure 3: Hasse diagrams of the old algebra (on the left) and the new algebra with the refinement embedding (on the right).



However, h, as seen above, is not the only possible embedding. Consider h' that is as follows: $h'(\top) = \mathsf{a} \vee \mathsf{b}'$, $h'(\bot) = \bot'$, $h'(\mathsf{A}) = \mathsf{a}$ and $h'(\mathsf{B}) = \mathsf{b}'$. Also consider h'' that is as follows: $h''(\top) = \top'$, $h''(\bot) = \bot'$, $h''(\mathsf{A}) = \mathsf{a} \vee \mathsf{b}'$ and $h''(\mathsf{B}) = \mathsf{b}''$. Below, in Figure 4, you can see the same post awareness growth algebra as in Figure 3, but with the h' embedding highlighted in green and the h'' embedding highlighted in purple.

Figure 4: Hasse diagrams of two new algebras with alternative embeddings.



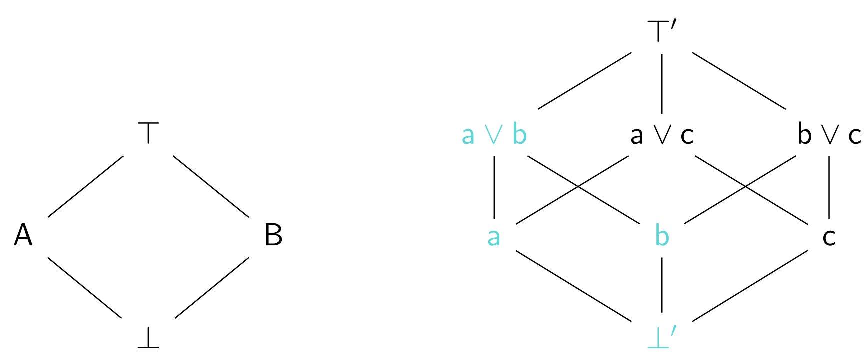
Thus, we can see that there are at least three possible ways to embed the original algebra into a post refinement algebra and indeed, there are more. But which one is the appropriate one?

For one, in the case of awareness growth by refinement, it is reasonable to map the top element of the old algebra to the top element of the new algebra. This is because during refinement we further partition some proposition, but the sum total of all propositions - their disjunction - remains the same. This then rules out the green embedding. Perhaps even more importantly, it is only right that we map the fragmenting proposition to the disjunction of the propositions into which it fragments. Thus B should be mapped to $\mathsf{b}' \vee \mathsf{b}''$. These two desiderata are met only by the red embedding. Of course there are other embeddings not depicted here that meet these two desiderata. For instance, we can map A to $\mathsf{a} \vee \mathsf{b}'$ and $\perp$ to b'. However, both of those moves seem unwarranted. There are no good reasons to map A anywhere other than a, the closest corresponding proposition. This then forces $\perp$ to be mapped to $\perp'$.

To sum up, in the case of refinement, we map the basic propositions that did not fragment to "themselves". The fragmenting proposition is mapped to the disjunction of propositions into which it fragments and the bounds of the old algebra are mapped to the bounds of the new algebra.

Let us now turn to an embedding that would be standard in the case of awareness growth by expansion. Take the same set of basic propositions $\{\mathsf{A}, \mathsf{B}\}$ that now turns into $\{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$. Since we have already seen that there is more than one way to embed a two atom algebra into a three atom algebra, in Figure 5 I merely present the old algebra alongside the new algebra with the intuitively correct embedding highlighted accordingly. The embedding g is as follows: $g(\top) = \mathsf{a} \vee \mathsf{b}$, $g(\perp) = \perp'$, $g(\mathsf{A}) = \mathsf{a}$ and $g(\mathsf{B}) = \mathsf{b}$.

Figure 5: Hasse diagrams of the old algebra and the new algebra with the expansion embedding.



This is the intuitively correct embedding because none of the old basic propositions underwent change, and therefore they should be mapped to the closest corresponding propositions. The rest follows by the definition of an embedding. Armed with the notion of an embedding we can present a new-look Reverse Bayesianism that is as follows.

Updated Reverse Bayesianism.[56] Given an embedding $h : \mathbb{B}_{\mathcal{X}} \to \mathbb{B}_{\mathcal{X}^+}$, for all $a, b \in \mathcal{X}^{cl}$, $Cr^+(\cdot)$ over $\mathbb{B}_{\mathcal{X}^+}$ ought to be such that:

$$\frac{Cr(a)}{Cr(b)} = \frac{Cr^+(h(a))}{Cr^+(h(b))}.$$

The purpose of this chapter is twofold. First, is to argue that even Updated Reverse Bayesianism is subject to counterexamples and is thus not a general requirement of rationality in cases of awareness growth. Second, is to draw conclusions from this apparent failure. While I am not alone in thinking (Updated) Reverse Bayesianism is not a general requirement of rationality (Mahtani, 2020; Steele and Stefánsson, 2021a; Steele and Stefánsson, 2021b; Pettigrew, n.d.), the conclusions I draw from this failure are rather different to the aforementioned authors.[57] In particular, those authors are sceptical of any general requirements of rationality in the case of revising credences in light of awareness growth and thus have a pessimistic outlook on the Problem of Awareness Growth. I, however, suggest that Updated Reverse Bayesianism isn't useless when it fails. Instead, in cases when it fails, it provides a procedure to diagnose when we should reconsider how the problem facing the agent is to be circumscribed. Moreover, the manner of failure of Updated Reverse Bayesianism indicates how exactly the problem facing the agent is to be reconsidered. This fits naturally with the notion of a "localised investigation" introduced in Chapter 3. In sum, we shouldn't be so pessimistic about the Problem of Awareness Growth. There is a way to proceed, with plenty of work still to do.

The remainder of this chapter is as follows. First, building on work by Steele and Stefánsson (2021a), I present my own counterexamples to Updated Reverse Bayesianism. Crucially, I distinguish two different mechanisms by which it fails. Second, I take a step back and consider why Updated Reverse Bayesianism, is desirable in the first place. In short, it is the notion of conservative upgrade - the minimal change in credences warranted by the learning experience at hand. Third, I argue that once we bring into focus the fact Bayesianism is to be applied in well circumscribed situa-

[56] This formulation is my own. Roussos (2020), who introduces the embedding idea to the discussion of awareness growth opts for a different, though related, principle for changing credences in light of awareness growth. We will turn to this principle later.

[57] Sometimes I say (Updated) Reverse Bayesianism instead of Updated Reverse Bayesianism because strictly speaking, the authors did not have the embedding version in mind. Still, the spirit of the arguments leads to Updated Reverse Bayesianism just the same.

tions ("small world" problems in Savage's terminology), the conservativity of Updated Reverse Bayesianism allows it to play a role even when it fails. It can play the role of a diagnostic tool, and what it reveals is that our initial way of circumscribing the situation was defective and we should go back to the drawing board. Moreover, the mechanism of failure points us to the appropriate way to restructure the problem.

5.1 Awareness growth, evidential relevance and the mechanism of credal assignment

Consider the following case.

> GamesShow1: Jill is a contestant on a games show and she is asked the following question. In what historic period did Offa of Mercia rule? She is told that he ruled after 1 AD and before 2000 AD inclusive and presented with 4 options: Antiquity, the Middle Ages, the Early Modern Period and the Modern Period. Jill has no idea. However, she does know something about the historic periods - she knows roughly for how long they stretched. She thus assigns credences to the options based on what proportion of 2000 years they occupy. She then decides to call a historian friend. Unfortunately for Jill, her friend specialises on the Assyrian Empire (14-7th century BC) and has no idea either. However, Jill's friend, while scrambling to recite all the facts she knows about the periods involved, makes Jill aware of the distinction between the early, high and late Middle Ages. This three part distinction makes Jill think that the Middle Ages stretch for longer than she thought before. In particular, the period starts earlier. Thus, becoming aware of the distinction in the Middle Ages period makes Jill comparatively more confident than she was before, that Offa of Mercia ruled in the Middle Ages, than say the Modern Period.

The basic propositions at the start of the case are $\{\mathsf{AN}, \mathsf{MA}, \mathsf{EM}, \mathsf{MP}\}$. The original credal assignment was driven by the length for which the periods stretch in time. Jill thinks the periods are as follows: Antiquity stretches from 1 AD to 750 AD, the Middle Ages stretch from 750 AD to 1400 AD, the Early Modern Period stretches from 1400 AD to 1800 AD and the remainder is the Modern Period. The corresponding credal assignment is given by $Cr(\mathsf{AN}) = \frac{750}{2000}$, $Cr(\mathsf{MA}) = \frac{650}{2000}$, $Cr(\mathsf{EM}) = \frac{400}{20000}$ and $Cr(\mathsf{MP}) = \frac{200}{2000}$. After becoming aware of the tripartite distinction between the early, high and late Middle Ages, the new space of basic propositions is given by

$\{\mathsf{an}, \mathsf{ema}, \mathsf{hma}, \mathsf{lma}, \mathsf{em}, \mathsf{mp}\}$. Thus, prima facie, we have a case of awareness growth by refinement.

Since it is reasonably a case of refinement, we map AN to an, EM to em and MP to mp and MA to $\mathsf{ema} \vee \mathsf{hma} \vee \mathsf{lma}$. Thus, the fragmenting proposition is mapped to the disjunction of the propositions into which it fragments. Thus, we have a counterexample to Updated Reverse Bayesianism since $\frac{Cr(\mathsf{MA})}{Cr(\mathsf{MP})} < \frac{Cr^{+}(\mathsf{ema} \vee \mathsf{hma} \vee \mathsf{lma})}{Cr^{+}(\mathsf{mp})}$.

To diagnose what happens after awareness growth occurs, the refinement alters the relation of the propositions to the mechanism which drove the credal assignment. Jill simply assigned credences based on how long she thought the periods were, but awareness growth led her to adjust the period length. However, that is not the only way that awareness growth can affect the relation between propositions and the mechanism behind credal assignment. In particular, sometimes awareness growth changes the mechanism itself, as the case below shows.

> $\mathsf{GamesShow2}$: Jill is a contestant on a games show and she is asked the following question. In what century was the Treaty of Rapallo signed? Was it the 17th, 18th, 19th or 20th century? Jill has no idea so deems them all equally likely. She decides to phone a friend who is a historian. Unfortunately for Jill, her friend is an ancient historian and doesn't know when the Treaty of Rapallo was signed. However, she makes Jill aware that the League of Nations was founded in 1920 to promote international co-operation. Jill now divides the 20th century option into pre and post 1920. As a result of this refinement, she surmises that perhaps an international body explicitly founded to improve co-operation would lead to more treaties post 1920. Hence, after refining the 20th century option into pre and post 1920, Jill now assigns relatively more credence to the 20th century option than any other option.

At the formal level the situation is similar to before. The set of old basic propositions is $\{\mathsf{17C}, \mathsf{18C}, \mathsf{19C}, \mathsf{20C}\}$. After awareness growth, the set of basic propositions is $\{\mathsf{17c}, \mathsf{18c}, \mathsf{19c}, \mathsf{20c}', \mathsf{20c}''\}$. Since this is also plausibly a case of refinement, the embedding is intuitive; in particular, $\mathsf{20C}$ is mapped to $\mathsf{20c}' \vee \mathsf{20c}''$. Counter Updated Reverse Bayesianism, we have $\frac{Cr(\mathsf{20C})}{Cr(\mathsf{19C})} < \frac{Cr^{+}(\mathsf{20c}' \vee \mathsf{20c}'')}{Cr^{+}(\mathsf{19c})}$.

As mentioned, in this case, the awareness growth prompts the agent to re-consider the mechanism underlying credal assignment itself. Jill doesn't know when the Treaty of Rapallo was signed and thinks it can be at any point in these centuries. Jill then assigns credences based on the equal amount of time that the centuries occupy. After her awareness grows, she uses a different mechanism for credal assignment. She now thinks in terms of an event that would seek to promote international co-operation and assigns credences based on whether the centuries are wholly before 1920, or otherwise.

Before we go on to discuss possible replies from defenders of Updated Reverse Bayesianism, it is important to compare the counterexamples above to an extant counterexample in the literature. Indeed, the cases above have clear parallels with a counterexample featuring awareness growth by refinement due to Steele and Stefánsson (2021a). The central feature of that case, in the words of Steele and Stefánsson themselves, is that: “the original awareness context is partitioned according to some new property [...] that is taken to be evidentially relevant to the comparison of some pair of incompatible basic propositions” (Steele and Stefánsson 2021a, p. 1220). In the $\mathsf{GamesShow1}$ case, the evidential relevance consists in the fact that the distinction prompts Jill to reconsider the relation between propositions and the mechanism which she uses to govern credal assignment. In $\mathsf{GamesShow2}$ the evidential relevance consists in the fact that the distinction prompts Jill to re-consider the mechanism of credal assignment itself. The point to be emphasised is that there are a number (at least two) quite distinct ways for awareness growth by refinement to be evidentially relevant to pairs of pre awareness growth propositions. This will become useful when we come to discuss how Updated Reverse Bayesianism is to be used in the cases when it fails.

But now, what can a defender of Updated Reverse Bayesianism say in the cases above? A natural first response in both cases is to map the fragmenting proposition to less than the disjunction of all the propositions into which it fragments. For instance, suppose in $\mathsf{GamesShow1}$ we map MA to $\mathsf{hma} \vee \mathsf{lma}$ and in $\mathsf{GamesShow2}$ we map $\mathsf{20C}$ to $\mathsf{20c}'$. Then Updated Reverse Bayesianism (merely) mandates that $\frac{Cr(\mathsf{MA})}{Cr(\mathsf{MP})} = \frac{Cr^+(\mathsf{hma} \vee \mathsf{lma})}{Cr^+(\mathsf{mp})}$ and $\frac{Cr(\mathsf{20C})}{Cr(\mathsf{19C})} = \frac{Cr^+(\mathsf{20c}')}{Cr^+(\mathsf{19c})}$. Thus, because credences in ema and $\mathsf{20c}''$ are not constrained by Updated Reverse Bayesianism, it is possible to assign them credences such as to get the intuitively correct result while not violating Updated Reverse Bayesianism. To illustrate, suppose that Jill's new credences are $\frac{Cr(\mathsf{MA})}{Cr(\mathsf{MP})} = \frac{Cr^+(\mathsf{hma} \vee \mathsf{lma})}{Cr^+(\mathsf{mp})}$ and $\frac{Cr(\mathsf{20C})}{Cr(\mathsf{19C})} = \frac{Cr^+(\mathsf{20c}')}{Cr^+(\mathsf{19c})}$, as per Updated Reverse Bayesianism and the earlier embedding assumption. We are then free to have $Cr^+(\mathsf{ema})$ and $Cr^+(\mathsf{20c}'')$ such that $\frac{Cr(\mathsf{MA})}{Cr(\mathsf{MP})} < \frac{Cr^+(\mathsf{ema} \vee \mathsf{hma} \vee \mathsf{lma})}{Cr^+(\mathsf{mp})}$ and $\frac{Cr(\mathsf{20C})}{Cr(\mathsf{19C})} < \frac{Cr^+(\mathsf{20c}' \vee \mathsf{20c}'')}{Cr^+(\mathsf{19c})}$.

However, these embeddings make $\mathsf{GamesShow1}$ and $\mathsf{GamesShow2}$ cases of expansion, when intuitively they are not. In these cases, Jill further partitions one of the initial options. Content-wise, ema is clearly related to MA and analogously for $\mathsf{20c}''$ and $\mathsf{20C}$. Whereas in cases of awareness growth by expansion, the expansion is driven by becoming aware of a proposition whose content is unrelated (other than having the same subject matter) to the propositions pre awareness growth. Both cases are straightforward cases of awareness growth by refinement and I see no reason to think otherwise, other than that it endangers Updated Reverse Bayesianism.

Another defensive line to take in GameShow1 is to say that awareness growth corrected faulty reasoning. Jill shouldn't have taken the periods to stretch for the length she initially did. Thus, this response dismisses GameShow1 as a counterexample on the grounds that it features an irrational agent whose behaviour a theory of rationality ought not to take as a datum. However, Jill wasn't in a position to recognise the fact she underestimated the length of the Middle Ages prior to becoming aware of the distinction amongst the Middle Ages and so it is not the case that this is somehow a defective case of credal change. Surely awareness growth is interesting in cases where it pushes our conceptual resources to conclusions we were previously unable to draw. Thus, the case is an example of a core phenomenon of interest and is in no way defective. This response is even less plausible in case of GameShow2 because it is straightforwardly a case of awareness growth which enables Jill to structure credal assignment from a vantage point of being more informed.[58]

Another line of reply is that really, we do not have a single process of a change in credences but at least two distinct processes.[59] Initially, Updated Reverse Bayesianism takes over, and then there is a separate stage that lands us with the final credal assignment where the ratios mandated by Updated Reverse Bayesianism are violated (Bradley, 2022).

What we will need from a defender of Updated Reverse Bayesianism is a proposal as to how, formally, the second stage behaves. While we wait for this proposal we can nonetheless say a few things in reply even at this stage. Let us grant that we can distinguish, at a conceptual level, between the two stages of changing credences.[60] First, as we saw, there are at least two distinct ways for awareness growth to be evidentially relevant to pairs of incompatible propositions, so this proposal would need to provide a unified treatment for both of those ways. Seeing as though the two ways are quite distinct, I am sceptical that a future proposal for the second stage would

[58]As it happens, in 1920 neither the Russian Empire nor Germany were part of the League of Nations and they were the signatories of the Treaty of Rapallo in 1922. So Jill seizes on information that is orthogonal to reality, but rationality is not a matter of always being correct and so Jill's behaviour is at least permissible.

[59]Insofar as it is hard to rationalise two stages instead of one, it is even harder to rationalise more than two stages. So herein I will assume that if one takes this line of response to counterexamples to Updated Reverse Bayesianism, it posits only one more additional stage.

[60]Steele and Stefánsson say of the two stage proposal that "The second learning stage is an odd, spontaneous learning event that would be hard to rationalise" (Steele and Stefánsson, p. 1221). To some extent I agree; the awareness growth and the associated change in credences really seem to be part of one process. However, especially when thinking about awareness growth, one should be mindful of the possibility of a proposal coming along to rationalise the breakdown of the process into two stages.

adequately account for the difference, but I am prepared to be proven wrong on this count by future work. After all, when one is writing about awareness growth it is only reasonable to leave open the possibility that something one has not yet encountered will change one's opinion.

Second, one might protest that there is no motivation for positing two (or more) stages of credal update, other than to save Updated Reverse Bayesianism. I think there is some merit to this view, but imagine that an actual proposal for how to revise credences in the second stage comes around. Would it be so unmotivated then? If it saves Updated Reverse Bayesianism, we have a way to proceed in cases of awareness growth. Without Updated Reverse Bayesianism, it seems we have nothing. Indeed, many have come to the conclusion that not only is Updated Reverse Bayesianism not a general norm of rationality in the case of awareness growth, but there is reason to think that no such norm is forthcoming. In the face of such pessimism, holding onto Updated Reverse Bayesianism by positing a second stage seems like an eminently good idea; something is better than nothing. However, I will argue that we can profess scepticism about a general requirement of rationality in the case of awareness growth while carving out a role for Updated Reverse Bayesianism, or a closely related principle, without positing a second stage. The remainder of this chapter is concerned with putting forward this proposal and rationalising it.

I will first examine the rationale behind Updated Reverse Bayesianism. Once we bring this rationale to the fore, it will allow us to better see what role it, or a closely related norm, can play even when they fail.

5.2 Conservative upgrade

What kind of arguments have been offered in favour of (Updated) Reverse Bayesianism beyond intuitive appeal? The philosophical arguments are offered not in support of (Updated) Reverse Bayesianism directly, but a related principle called Awareness Rigidity (Wenmackers and Romeijn, 2016; Bradley, 2017). Recall Rigidity, the common core of Conditionalisation and Jeffrey Conditionalisation.

Rigidity. For all $a, b \in \mathcal{X}^{cl}$ such that $Cr(b) > 0$, $Cr^+(a|b) = Cr(a|b)$, where b and only b was learnt between $Cr(\cdot)$ and $Cr^+(\cdot)$.

Both Bradley and Wenmackers and Romeijn reflect on Rigidity and are thereby guided to a principle that can be stated as follows.

Awareness Rigidity.[61] Given an embedding $h : \mathbb{B}_{\mathcal{X}} \to \mathbb{B}_{\mathcal{X}^+}$, for all $a \in \mathcal{X}^{cl}$, $Cr^+(\cdot)$ over $\mathbb{B}_{\mathcal{X}^+}$ ought to be such that:

$$Cr^+(h(a)|h(\top)) = Cr(a).$$

What is the guiding thought from Rigidity as part of Conditionalisation to Awareness Rigidity? Bradley states that

> "[w]ithin the Bayesian framework, conservation of the agent's relational beliefs is ensured by the rigidity of her conditional probabilities. So we can conclude that conservative belief change requires the agent's new conditional probabilities, given the old domain, for any members of the old domain should equal her old unconditional probabilities for these members" (2017, p. 258)

Similarly, Wenmackers and Romeijn say that:

> "[i]n analogy with Bayes' rule, one natural conservativity constraint is that the new probability distribution must respect the old distribution on the preexisting parts of the algebra" (2016, p. 1235)

The main thought is as follows. Rigidity guarantees that Conditionalisation results in a minimal change of credence in light of learning a proposition. It is a minimal change because it does not alter credences conditional on that proposition. That is, when we are updating credence in a having learnt b for certain, the relative credence of a given b doesn't change upon learning b because learning b does not give evidence or reason to change one's relative credence of *a given b*. Of course, in the case of awareness growth we do not have conditional credences given the proposition(s) we become aware of, and therefore we need something else to stay constant across awareness growth. Awareness Rigidity says that what one should keep constant, is one's credences given the old domain. If one takes the ratios of credences in old propositions to be conceptually fundamental, then one gets Updated Reverse Bayesianism. Thus, the guiding idea behind both Awareness Rigidity and Updated Reverse Bayesianism is that they capture the notion of conservative credence update in the context of awareness growth.[62] We

[61]This formulation is an adapted version of the principle found in Bradley (2017). I adapted it by putting things in terms of an embedding, since the need to identify propositions across awareness growth does not go away with Updated Reverse Bayesianism.

[62]Such conservatism is not without analogues in the wider formal epistemology literature. In Dynamic Epistemic Logic, Conservative Upgrade is a particular rule for transforming so-called plausibility models van Benthem (2007). A conceptually related notion of minimal revision is also present in AGM belief revision theory, as can be seen in: Gärdenfors (1984), Boutilier (1996) and Arló-Costa and Levi (2006), among others.

change credences just enough to take into account awareness growth, but no more.

However, the two principles are not only related conceptually, but also formally. In fact, the two are equivalent in the framework where pre and post awareness growth algebras are linked by an embedding.[63] The immediate consequence is that GamesShow1 and GamesShow2 are counterexamples to Awareness Rigidity too.[64] Of course, even if two principles are formally equivalent, it still leaves open the question of which is conceptually the more fundamental one. I will not delve deeper into this and on the count of equivalence I will speak just of Updated Reverse Bayesianism.

We saw that Updated Reverse Bayesianism is not a general principle of rationality since it suffers from counterexamples. Now, with the underlying reasons behind Updated Reverse Bayesianism fully in view, let us turn to the question of where that leaves us with respect to the role that Updated Reverse Bayesianism can play even when it fails.

5.3 Localised investigation and failure of conservative upgrade

I have followed others in arguing that (Updated) Reverse Bayesianism is not a general requirement of rationality in the case of awareness growth (Mahtani, 2020; Steele and Stefánsson, 2021a; Steele and Stefánsson, 2021b; Pettigrew, n.d.). Now, what kind of conclusion do these authors thereby draw about the prospects of Updated Reverse Bayesianism and the prospect of a general norm governing credal change in cases of awareness growth?

On the first point, Steele and Stefánsson state "we can conclude that we should not impose Reverse Bayesianism as a general constraint on how a rational agent can revise her credences when her awareness grows" (Steele and Stefánsson 2021a, p. 1219). Pettigrew, after having discussed a counterexample to Reverse Bayesianism due to Mahtani (2020) concludes: "I think Reverse Bayesianism must be wrong" (Pettigrew n.d., p. 7).

On the second point, Steele and Stefánsson say "In other words, does rationality (or

[63] Proof under Claim 2, Appendix B.

[64] As noted earlier, both GamesShow1 and GamesShow2 are cases of awareness growth by refinement. The reader might wonder whether analogous counterexamples are available for awareness growth by expansion, and indeed there are. The "possible affair" case (Steele and Stefánsson 2021a, p. 1218) is one such counterexample. However, there is a reason why my counterexamples to Updated Reverse Bayesianism are cases of refinement. This is because Awareness Rigidity does not imply Reverse Bayesianism in a possible worlds-based framework of Steele and Stefánsson (2021a), but it does in the case of refinement. Thus, someone who prefers the framework that Steele and Stefánsson (2021a) use to the one I use can nonetheless take GamesShow1 and GamesShow2, and use them as counterexamples to Awareness Rigidity too. Thus, focusing on cases of awareness growth by refinement makes my argument more general and applicable beyond the framework I use.

stability) impose any general constraints on the relationship between one's credences prior to and post some growth in awareness? [...] we argued that there are no such general constraints" (Steele and Stefánsson 2021b, p. 122). Pettigrew (n.d.) is similarly sceptical. Speaking in the context of norms governing credal change in case of awareness growth that are supported by the most prominent arguments for Conditionalisation, he says "any such norms that follow in the case of awareness growth will apply much less often than those that follow in the case of new evidence" (Pettigrew n.d., p. 27). So the consensus view is that (Updated) Reverse Bayesianism is not a general norm of rationality in the case of awareness growth and, more broadly, there is reason to think that no such norm is forthcoming. Steele and Stefánsson (2021a) is a book-length argument for this scepticism about a general norm and, whilst I cannot hope to even state all of the reasons they give, the core idea is simple. Awareness growth is capable of being so unpredictable that the changes it can bring are limitless.

Now suppose that there is not an exceptionless norm for revising credences in light of awareness growth. For one, this makes Updated Reverse Bayesianism a partner in crime with literally any other extant principle. In this case waiting for a second stage to save Updated Reverse Bayesianism from extant counterexamples looks like the best way to proceed. I disagree. There is a way to proceed while acknowledging scepticism about a general rule for revising and carving out a distinct role for Updated Reverse Bayesianism without positing the fact that awareness growth sometimes involves more than one stage of credence update. Moreover, it allows us to ward off undue pessimism about the prospects of a well-worked out proposal for dealing with awareness growth in a Bayesian setting.

I myself argued that Updated Reverse Bayesianism suffers from counterexamples. My suggestion is that even if further evidence mounts against it, it should not be abandoned outright; we just need to re-purpose it. Updated Reverse Bayesianism plans for awareness growth in the sense that it allows for new possibilities to come to light, but maintains that something remains constant as awareness grows. The kind of awareness growth it plans for is awareness growth that won't threaten certain features of the way the agent takes things to be pre awareness growth. When awareness growth is just of this kind, we should stick with Updated Reverse Bayesianism. I cannot emphasise this point enough - in cases when Updated Reverse Bayesianism is not broken, we should not fix it.

But what about when it fails? The idea is that when Updated Reverse Bayesianism fails, it is to be used as a way to diagnose that the agent originally conceived of the problem in a defective way, be it at the level of the priors or even at the level of propositions, to which credences are attached. To that end, recall a point due to Teller

(1975) that while Bayesians like Savage were keen to emphasise that the theory is to be applied to suitably circumscribed situations (Savage's "small worlds") and sketched ways in which this circumscription should proceed, they have not identified when we should "start from scratch, restructuring his problem and making new subjective evaluations of a prior probability function" (Teller 1975, p. 170). We now have identified when the agent's initial way of conceiving of the situation was, in the face of awareness growth, ill-judged. It is just those cases when our norm of conservative upgrade is violated. Thus, cases when the conservative norm fails provide a natural stopping procedure that prompts us to start from scratch. Moreover, the manner of failure of the norm can be used as a guide on how to restructure the problem: what propositions to use, what priors to assign and so on. To that end, recall that in GamesShow1. What goes wrong is that the thing that the priors reflected, namely the perceived length of historic periods, changes. In GamesShow2 the mechanism of credal assignment itself changes.

Now let us turn to how the restructuring of the problem should proceed. Ought it be the case that we assign the priors to match the credences that violate our conservatvity norm? In GamesShow1, plausibly yes. For concreteness, suppose that instead of taking the Middle Ages to stretch from 750 AD to 1400 AD, Jill now takes them to start from 500 AD. Thus, her credences post awareness growth are: $Cr^{+}(\mathsf{AN}) = \frac{500}{2000}$, $Cr^{+}(\mathsf{MA}) = \frac{900}{2000}$, $Cr^{+}(\mathsf{EM}) = \frac{400}{20000}$ and $Cr^{+}(\mathsf{MP}) = \frac{200}{2000}$. Since what awareness growth changes is the relation between the propositions and the mechanism of credal assignment and nothing more, it seems fair that in our restructuring we reassign priors to reflect this change and do no more. Thus, when we restructure the problem, we have just the same propositions as before, but with new priors.

However, in GamesShow2, I do not think we should adopt as priors those credences that violate the conservativity constraint. This is because awareness growth changed the very mechanism of credal assignment. Jill previously didn't consider international bodies as a salient factor in assessing when a given treaty was signed. After awareness growth, she does. There is cause to inquire into the matter further; awareness growth destabilised things and further information can help stabilise things before setting up the problem again. For instance, Jill could look into salient points in international relations that are to do with co-operation: Peace of Westphalia (1648), Congress of Vienna (1815), establishment of the League of Nations (1920), founding of the UN (1945) and so on. If she doesn't know where to start, plausibly she would consult an expert for guidance. In sum, awareness growth sometimes shakes things up enough to prompt the agent to inquire into the matter further, before proceeding with new priors or even a new space of propositions. Of course, the example is somewhat artificial since in a games show Jill cannot do any of these things, and if she could, she could also find out when the Treaty of Rapallo was signed! But the general point stands, awareness growth shook up things enough to warrant new priors and if it were possible, Jill should investigate the matter further before re-assigning priors.

I would like to note that everyone in the recent literature, with the notable exception of de Canson (2020), at least implicitly takes Bayesianism to only be applicable in well circumscribed cases.[65] Why? Because these authors themselves go through the effort to set up the problem in terms of a set of scientific hypotheses of some domain or a set of propositions representing an agent in more mundane cases of uncertainty. Since all the aforementioned authors start from somewhere, vis-a-vis the problem at hand, I think they should find my proposal amenable. The difference between myself and these authors, with the notable exception of Strevens (2010), is that my proposal involves not just intra-problem changes but sometimes inter-problem changes. So, these authors structure the problem and then all changes occur within said problem - updating by Conditionalisation, awareness growth, credal update in light of awareness growth and so on. On my proposal, however, sometimes the way the problem was originally set-up runs its course and we then restructure it and proceed from there.

My proposal differs from Strevens' in that there is a lot more continuity. For one, even if it fails, the norm not only provides further guidance as to how to restructure the problem further but also constrains how it is to be done. Strevens, however, leaves this restructuring as a completely open-ended endeavour.

Moreover, in light of what has been said, I think there is no need to wait for a potential proposal of how the second stage of Updated Reverse Bayesianism proceeds. Instead, we should use those principles to diagnose just when we need to restructure the agent's space of possibilities. Furthermore, we are not left without further work. For one, it is still an open question as to whether Awareness Rigidity or Updated Reverse Bayesianism is *the* norm of conservative upgrade. In this chapter I presented two different ways in which awareness growth by refinement can be evidentially relevant to pairs of propositions. Future work should expand on that or show why those two are the only ways. This further work is consequential on my proposal because the manner of failure of the conservative norm matters for how we will restructure the space of possibilities, how we will re-assign the priors and so on.

[65] The recent literature includes (Maher, 1995; Strevens, 2010; Wenmackers and Romeijn, 2016; Bradley, 2017; Mahtani, 2020; Roussos, 2020; Steele and Stefnsoon, 2021a; Steele and Stefnsson, 2021b).

6 Conclusion

6.1 Thesis Summary

The chief contribution of this thesis is threefold, corresponding to Chapters 3, 4 and 5 respectively. In Chapter 3, I clarified the Bayesian Problem of Awareness Growth. I distinguished two different versions of the problem, argued against one of these versions on the grounds that it is at odds with the normative ambitions of Bayesianism, and finally, delineated for whom and when my preferred version of the Bayesian Problem of Awareness Growth is a *problem*. In particular, the Problem of Awareness Growth isn't a problem for those who boil down credal assignment to assigning credences at discrete points in time with no link between those times, but that results in giving up Conditionalisation and Bayesianism with it. There is also a way out of the problem that just re-works the space of propositions at every instance of awareness growth with new priors being re-assigned each time awareness growth takes place.

In Chapter 4, I turned to one of the key debates around the Problem of Awareness Growth: whether a catch-all proposition has to be included in the agent's algebra. The dialectic of Chapter 4 came in three stages. First, I argued that if a catch-all is to be included, it should not be assigned a definite prior or likelihood. The intractability of the prior and the likelihood in the catch-all is at once a negative result for the catch-all and a point in the right direction. It is a negative result because it was shown that whatever the difficulties are with assigning priors and likelihoods to ordinary propositions, doing so for the catch-all is categorically harder. However, the intractability also points to a way out since it automatically privileges all those catch-all proposals that do not assign a definite prior and likelihood. If we are to have a successful implementation of a catch-all proposition, it should not be assigned a prior and likelihood. This completes the first step of the dialectic.

In the second step, I further privileged a catch-all proposal due to Wenmackers and Romeijn (2016). Their proposal not only does not assign a definite prior and likelihood to the catch-all but solves two additional problems for the catch-all, as pressed by Earman (1992). The second step of the dialectic was completed by responding on behalf of Wenmackers and Romeijn to possible objections that could further be raised against their proposal. In the third step, I argued that extant arguments purportedly showing that the catch-all should always be included in the agent's algebra fail. Thus, if one wants to include a catch-all proposition, there is a way to proceed but there's no general requirement to always include it.

Chapter 5 was concerned with a key putative norm for revising credences in light of awareness growth - Reverse Bayesianism. The chapter first proceeded to update

Reverse Bayesianism with a criterion for specifying the relation between the pre and post awareness growth algebras. This was done via the notion of an embedding first used in this context by Roussos (2020). The resultant principle, Updated Reverse Bayesianism, turned out to be equivalent to a related principle, Awareness Rigidity, in the embedding framework.

On top of this, Chapter 5 did three things. First, I presented two counterexamples to Updated Reverse Bayesianism (and hence Awareness Rigidity) that built on a counterexample due to Steele and Stefánsson (2021a). This apparent failure led me to take a step back and consider why Updated Reverse Bayesianism is a plausible norm in the first place. In short, it is because it provides a norm for changing credences in light of awareness growth that is most closely aligned with the main Bayesian norm - Conditionalisation. Updated Reverse Bayesianism does this because it offers a conservative norm of credal update in light of awareness growth - the minimal change in credence brought about by awareness growth. Conditionalisation does the same in the case of evidential learning. I then used this insight to point to the role that Updated Reverse Bayesianism can play even when it fails.

When Updated Reverse Bayesianism fails, the agent's initial way of conceiving of the situation was, in the face of awareness growth, ill-judged. It is just those cases when our norm of conservative upgrade is violated. Thus, cases when the conservative norm fails provide a natural stopping procedure that prompts us to start from scratch. Moreover, the manner of failure of the norm can be used as a guide on how to restructure the problem: what propositions to use, what priors to assign and so on.

This last point is noteworthy because it is able to acknowledge that Updated Reverse Bayesianism is not an exceptionless norm and even that in the case of awareness growth no single exceptionless norm is forthcoming, yet this still points to a systematic way to proceed in the case of awareness growth. This goes against the grain of the literature where large-scale scepticism about an exceptionless norm leads people to conclude that there is no systematic way to proceed when it comes to changing credences in light of awareness growth. My proposal acknowledges the scepticism while denying the pessimistic conclusion.

6.2 Future Work

All this being said, there is still plenty of work left to do. In closing, I would like to note some strands of future research that seem fruitful to me. These strands concern the representation of awareness in an algebraic framework and its relation to other framework for representing awareness, future work in analysing different and interesting cases of awareness growth and finally, the study of the limits of Bayesianism. I take these in turn.

One assumption that we have been working on throughout this thesis, because it is shared by the wider literature, is that there are two types of awareness growth: expansion and refinement. Plausibly, there are other ways for awareness to grow that we can reflect in an algebraic framework. For one, sometimes we become aware that two options we thought are distinct are actually one and the same. Instead of splitting, two propositions merge into one. This would mean that a notion of an embedding is not fit to be the sole operation we use to identify pre and post awareness growth propositions. This is because an embedding is injective and a merging of propositions would require an operation that is not injective.

Moreover, while cases of forgetting are not instances of awareness growth, it would be welcome if the framework for representing awareness could also account for a loss in awareness. Again, this means we wouldn't be able to rely on the notion of an embedding because it is a total function, hence when the algebra loses elements, injectivity would have to be violated. Thus, there is more work to be done in analysing awareness in an algebraic framework. While this line of work is of independent interest, it is also not idle for Bayesian purposes.

Notably, if embedding is the suitable operation linking pre and post awareness growth algebras, then Awareness Rigidity and Updated Reverse Bayesianism are equivalent. This is not so in a framework that represents awareness and changes in awareness with possible worlds (Steele & Stefánsson, 2021a). In particular, in that framework Awareness Rigidity does not imply Reverse Bayesianism in the case of awareness growth by expansion. This is an interesting divergence, and more work developing the two frameworks for representing awareness more generally, would offer a better way to compare the two approaches and the associated formal results.

A line of work that naturally arises from my suggestion that we should re-work the space of propositions we had pre awareness growth in cases when Updated Reverse Bayesianism fails, is a more careful cataloguing of the different cases in which it fails. As I noted, there are at least two distinct mechanisms by which Updated Reverse Bayesianism fails, and there are plausibly more. Since I suggest that the way we re-work either the priors or the propositions (or both) depends on the mechanism of failure, I think it is natural that awareness growth is treated not as a single unified phenomenon, but as a family of disparate phenomena. For instance, the phenomenon of transformative experience has attracted interest in the philosophical literature since Paul (2014). A transformative experience is an experience that is so radically unlike what one has previously encountered that one can't assess its effect prior to the experience. A paradigm example is that of becoming a parent. The kind of work on awareness growth I would advocate is a careful consideration of a sub-species of awareness growth, like transformative experience, and then an analysis of how to revise credences in that particular case. Perhaps further reflection even makes distinctions amongst different cases of what is considered a transformative experience. Thus, the literature on awareness growth becomes more fragmented, but in light of my proposal the fragmentation would be welcome.

This thesis, tacitly, also raises a foundational question for Bayesianism. What are the proper limits of Bayesianism? What kind of phenomena is it concerned with? For instance, should it aim to provide a complete theory of epistemic rationality, or simply play a role in a patchwork of principles? The proposal in Chapter 5 favours the second option. After all, on my proposal, credences in propositions at a given time are only a part of the full picture. For one, the agent takes the space of propositions to be provisional; it might change in light of the kind of awareness growth that violated Updated Reverse Bayesianism. Moreover, a credence of 1 in the disjunction of all basic propositions, at a given point in time can hardly be interpreted as unconditional commitment to the truth of one of the basic propositions and so the picture presented by an agent's space of propositions and their credence is only a part of the full picture.

However, this patchwork view comes at a cost. To my knowledge, Bayesianism is the only theory of epistemic rationality that links up with a decision theory, thus forming a unified theory of rational decision-making, action and belief. However, recall that the inclusion of a catch-all meant that we can't have numerical posteriors even in the explicit propositions. This now means that we can't straightforwardly mix the credences in propositions with the utility attached to them. Thus, the unifying power of Bayesianism is threatened by a more patchwork approach. Settling the limits of Bayesianism would go a long way to delineate what kind of moves are forbidden when it comes to thinking about the Problem of Awareness Growth and other issues as well.

To conclude, I think progress has been made when it comes to the Problem of Awareness Growth. The problem itself has been clarified and we have moved forward in two key debates. The fact new questions have arisen along the way, goes to show that there is a long way for our awareness to go.

A Probability Proofs

A1

Appendix A1 proves some elementary facts about Boolean algebras and rational credences, which are just (subjective) probabilities. The reader who is somewhat familiar with Boolean algebras and probabilities should feel free to skip proofs in Appendix A1.

Recall that the domain of the credence function is a Boolean algebra, $\mathbb{B}_{\mathcal{X}}$, spanned by a set of basic propositions $\mathcal{X}$. Correspondingly, $\mathcal{X}^{cl}$ is the closure of $\mathcal{X}$ under $\neg, \vee$ and $\wedge$. The relation $\models$ is defined as follows: $a \models b$ iff $b = a \vee b$. The Boolean laws are as follows.

Commutativity	$a \vee b = b \vee a$	$a \wedge b = b \wedge a$
Associativity	$a \vee (b \vee c) = (a \vee b) \vee c$	$a \wedge (b \wedge c) = (a \wedge b) \wedge c$
Idempotence	$a \vee a = a$	$a \wedge a = a$
Absorption	$a = a \vee (a \wedge b)$	$a = a \wedge (a \vee b)$
Distributivity	$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$	$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$
Join Bound	$a \vee \neg a = \top$	$a \vee \bot = a$
Meet Bound	$a \wedge \neg a = \bot$	$a \wedge \top = a$

In the forthcoming probability proofs some facts about Boolean algebras will be useful and I prove them before moving on with the main claims.

Fact 1: $b = a \vee b$ iff $a = a \wedge b$.

Proof.

($\Longrightarrow$) Suppose $b = a \vee b$. By Absorption $a = a \wedge (a \vee b)$. Since by assumption $b = a \vee b$, this simplifies to $a = a \wedge b$.

($\Longleftarrow$) Suppose $a = a \wedge b$. By Absorption $b = b \vee (b \wedge a)$. By Commutativity from $a = a \wedge b$ we obtain $a = b \wedge a$. Thus, we obtain $b = b \wedge a$. Which by Commutativity yields $b = a \vee b$.

□

Fact 2: If $a \models b$ and $b \models a$ then $a = b$.

Proof.

Suppose $a \models b$ and $b \models a$. By definition of $\models$ this means that $b = a \vee b$ and $a = b \vee a$. By Commutativity, $b \vee a = a \vee b$. Thus, $a = b$.

$\square$

Fact 3: $a = (a \wedge b) \vee (a \wedge \neg b)$.

Proof.

By Meet Bound $a = a \wedge \top$. By Join Bound $\top = b \vee \neg b$. Thus, $a = a \wedge (b \vee \neg b)$. By Distributivity we obtain $a = (a \wedge b) \vee (a \wedge \neg b)$.

$\square$

Fact 4: If $a \models b$ then $a = a \wedge b$.

Proof.

Suppose $a \models b$. By definition of $\models$, $b = a \vee b$. By Fact 1, $a = a \wedge b$.

$\square$

Fact 5: $\neg a \wedge b \models \neg a$**.**

Proof.

By Idempotence $\neg a = \neg a \wedge \neg a$. Thus $\neg a \wedge b = (\neg a \wedge \neg a) \wedge b$. By Associativity and then Commutativity, this yields $\neg a \wedge b = (\neg a \wedge b) \wedge \neg a$. Thus, by definition of $\models$ and Fact 1, $\neg a \wedge b \models \neg a$.

$\square$

Fact 6: $\neg \top = \bot$.

Proof.

By Meet Bound $\bot = \top \wedge \neg \top$. Commutativity and Meet Bound $\top \wedge \neg \top = \neg \top$. $\square$

Fact 7: $(a \wedge b) \wedge (a \wedge \neg b) = \bot$.

Proof.

By Associativity we obtain $(a \wedge b) \wedge (a \wedge \neg b) = (a \wedge a) \wedge (b \wedge \neg b)$. Which by Idempotence and Meet Bound yields $(a \wedge b) \wedge (a \wedge \neg b) = a \wedge \bot$. Which by another application of Meet Bound yields $(a \wedge b) \wedge (a \wedge \neg b) = a \wedge (a \wedge \neg a)$. By Associativity and Idempotence $a \wedge (a \wedge \neg a) = a \wedge \neg a$. The right-hand side, by Meet Bound, equals $\bot$. □

With theses facts in hand we can go onto deducing consequences of the Axioms of Probability. The Axioms being as follows.

(1) $Cr(a) \geq 0$.

(2) $Cr(\top) = 1$.

(3) If $a \wedge b = \bot$, then $Cr(a \vee b) = Cr(a) + Cr(b)$.

Claim 1: For all $a \in \mathcal{X}^{cl}$, $Cr(a) + Cr(\neg a) = 1$

Proof.

By Axiom (3), $Cr(a \vee \neg a) = Cr(a) + Cr(\neg a)$ since by Meet Bound we have $a \wedge \neg a = \bot$. But then since $a \vee \neg a = \top$ by Join Bound, functionality of $Cr(\cdot)$ yields $Cr(\top) = Cr(a \vee \neg a)$. Thus, by Axiom (2) we obtain $Cr(a \vee \neg a) = 1$ and thus, $Cr(a) + Cr(\neg a) = 1$.

□

Claim 2: For all $a \in \mathcal{X}^{cl}$, $Cr(a) \leq 1$.

Proof.

By Axiom (1), $Cr(\neg a) \geq 0$. But then by Claim 1 it follows that $Cr(a) \leq 1$.

□

Claim 3: For all $a, b \in \mathcal{X}^{cl}$ such that $a \models b$ and $b \models a$, $Cr(a) = Cr(b)$.

Proof.

Suppose $a \models b$ and $b \models a$. By Fact 2 this implies that $a = b$. So, by functionality of $Cr(\cdot)$ it follows that $Cr(a) = Cr(b)$. □

Claim 4: For all $a, b \in \mathcal{X}^{cl}$, $Cr(a) = Cr(a \wedge b) + Cr(a \wedge \neg b)$.

Proof.

By Fact 3, $a = (a \wedge b) \vee (a \wedge \neg b)$. By functionality of $Cr(\cdot)$, $Cr(a) = Cr((a \wedge b) \vee (a \wedge \neg b))$. By Fact 7, $(a \wedge b) \wedge (a \wedge \neg b) = \bot$. Thus, by Axiom (3) we obtain $Cr((a \wedge b) \vee (a \wedge \neg b)) = Cr(a \wedge b) + Cr(a \wedge \neg b)$.

□

Claim 5: For all $a, b \in \mathcal{X}^{cl}$, if $a \models b$, then $Cr(a) \leq Cr(b)$.

Proof.

Suppose $a \models b$. By Claim (4), $Cr(b) = Cr(a \wedge b) + Cr(\neg a \wedge b)$. Since $a \models b$ it follows that $a = a \wedge b$, by Fact 4. Thus, by Claim 3 we obtain $Cr(b) = Cr(a) + Cr(\neg a \wedge b)$. By Axiom (1), $Cr(\neg a \wedge b) \geq 0$ and thus it follows that $Cr(a) \leq Cr(b)$.

□

Claim 6: For all $a, b \in \mathcal{X}^{cl}$, if $Cr(a) = 1$ then $Cr(a \wedge b) = Cr(b)$.

Proof.

Suppose $Cr(a) = 1$. By Claim 4, $Cr(b) = Cr(a \wedge b) + Cr(\neg a \wedge b)$. By the assumption that $Cr(a) = 1$ and Claim 1 we obtain that $Cr(\neg a) = 0$. But note that $\neg a \wedge b \models \neg a$ and thus, by Claim 5 we obtain $Cr(\neg a \wedge b) \leq Cr(a)$. Which by axiom (1) implies that $Cr(\neg a \wedge b) = 0$ and thus $Cr(b) = Cr(a \wedge b) + Cr(\neg a \wedge b)$ simplifies to $Cr(b) = Cr(a \wedge b)$. □

Claim 7: $Cr(\bot) = 0$.

Proof.

By Fact 6, $\neg\top = \bot$. By Claim 1 we obtain $Cr(\bot) = 1 - Cr(\top)$. By Axiom (2), $Cr(\top) = 1$ and so $Cr(\bot) = 0$.

□

With all these proofs in hand it is still to prove the Total Probability Theorem. Let $\{b_i\}_{i \in I}$ be a partition of $\mathcal{X}$. Which is to say that for two distinct elements of $\{b_i\}_{i \in I}$, b_i and b_j, $b_i \wedge b_j = \perp$ and $\bigvee_{i \in I} b_i = \top$.

Claim 8: Given a partition $\{b_i\}_{i \in I}$, for all $a \in \mathcal{X}^{cl}$, $Cr(a) = \Sigma_{i \in I} Cr(a|b_i)Cr(b_i)$.

Proof.

By assumption that $\{b_i\}_{i \in I}$ is a partition, $Cr(\bigvee_{i \in I} b_i) = 1$. But then, by Claim 6, $Cr(a \wedge \bigvee_{i \in I} b_i) = Cr(a)$. By continuous application of Distributivity, from $a \wedge \bigvee_{i \in I} b_i$ we can obtain $\bigvee_{i \in I} a \wedge b_i$. So, $Cr(a) = Cr(\bigvee_{i \in I} a \wedge b_i)$. The right-hand side can be turned into $\Sigma_{i \in I} Cr(a \wedge b_i)$ by continuous application of Axiom (3) and the assumption that elements of the partition are mutually exclusive. Finally, by the definition of conditional probability $Cr(a \wedge b_i) = Cr(a|b_i)Cr(b_i)$ for each b_i and so we obtain $Cr(a) = \Sigma_{i \in I} Cr(a|b_i)Cr(b_i)$.

□

A2

Recall Conditionalisation, Rigidity and Certainty.

Conditionalisation. An agent with a credence function $Cr(\cdot)$, upon learning proposition $b \in \mathcal{X}^{cl}$ (such that $Cr(b) > 0$) with certainty and nothing else, ought to have a new credence function $Cr^+(\cdot)$ such that for all $a \in \mathcal{X}^{cl}$:

$$Cr^+(a) = Cr(a|b) := \frac{Cr(a \wedge b)}{Cr(b)} = \frac{Cr(b|a)}{Cr(b)} \times Cr(a).$$

Rigidity. For all $a, b \in \mathcal{X}^{cl}$ such that $Cr(b) > 0$, $Cr^+(a|b) = Cr(a|b)$, where b and only b was learnt between $Cr(\cdot)$ and $Cr^+(\cdot)$.

Certainty. For all $b \in \mathcal{X}^{cl}$ such that $Cr(b) > 0$, $Cr^+(b) = 1$, where b and only b was learnt between $Cr(\cdot)$ and $Cr^+(\cdot)$.

Now consider the following claim.

Claim 1: Conditionalisation is equivalent to the conjunction of Rigidity and Certainty.

Proof.

Conditionalisation ⇒ Certainty: It is to show that $Cr^+(b) = 1$ given the original credence function $Cr(\cdot)$ and the fact b and only b was learnt in the time between $Cr(\cdot)$ and $Cr^+(\cdot)$. By Conditionalisation $Cr^+(b) = Cr(b|b)$. By definition of conditional probability $Cr(b|b) = \frac{Cr(b \wedge b)}{Cr(b)}$. By Idempotence $b \wedge b = b$. So the right-hand side of the equation simplifies to $\frac{Cr(b)}{Cr(b)}$. This in turn equals 1 and thus $Cr^+(b) = 1$. That was to show.

Conditionalisation ⇒ Rigidity: It is to show that the conditional credence in a given b before having learnt b is equal to the conditional credence in a given b after having learnt b and only b. That is, we have to show that $Cr^+(a|b) = Cr(a|b)$.

By definition of conditional probability $Cr^+(a|b) = \frac{Cr^+(a \wedge b)}{Cr^+(b)}$. Since Conditionalisation implies certainty as shown above and the assumption that b and only b was learnt between the times of $Cr^+(\cdot)$ and $Cr(\cdot)$, $\frac{Cr^+(a \wedge b)}{Cr^+(b)}$ simplifies to $Cr^+(a \wedge b)$ since the denominator is equal to 1. $Cr^+(a \wedge b)$, by Conditionalisation equals $Cr(a \wedge b|b)$ which in turn, by the definition of conditional probability, equals $\frac{Cr(a \wedge b \wedge b)}{Cr(b)}$. By Idempotence $a \wedge b \wedge b = a \wedge b$. Thus, the right-hand side simplifies to $\frac{Cr(a \wedge b)}{Cr(b)}$. Which in turn, by definition of conditional probability equals $Cr(a|b)$. That was to show.

Rigidity + Certainty ⇒ Conditionalisation: It is to show that $Cr(a|b) = Cr^+(a)$ given that b and only b was learnt between the times of $Cr(\cdot)$ and $Cr^+(\cdot)$.

$Cr(a|b) = Cr^+(a|b)$ by assumption of Rigidity. The right-hand side equals $\frac{Cr^+(a \wedge b)}{Cr^+(b)}$, by the definition of conditional probability. The numerator is equal to 1, by the assumption of Certainty and we thus get $Cr^+(a \wedge b)$. Furthermore, by Claim 6 of A1, $Cr^+(a \wedge b) = Cr^+(a)$ since $Cr^+(b) = 1$ by Certainty. That was to show.

□

Recall Jeffrey Conditionalisation.

Jeffrey Conditionalsiation. An agent with a credence function $Cr(\cdot)$, who undergoes a (possibly) uncertain learning experience that has the effect of shifting, for each elements of the partition $\{b_i\}_{i \in I}$, credences from $Cr(b_i)$ to $Cr^+(b_i)$, ought to have credence $Cr^+(\cdot)$ such that for all $a \in \mathcal{X}^{cl}$:

$$Cr^+(a) = \sum_{i \in I} Cr(a|b_i) \cdot Cr^+(b_i).$$

Claim 2: Conditionalisation is a special case of Jeffrey Conditionalisation.

Proof.

Let the partition be just $\{b, \neg b\}$ and let $Cr^{+}(b) = 1$. The latter implies that $Cr^{+}(\neg b) = 0$. The Jeffrey Conditionalisation formula then gives us $Cr^{+}(a) = Cr(a|b) \cdot Cr^{+}(b) + Cr(a|\neg b) \cdot Cr^{+}(\neg b) = Cr(a|b) \cdot 1 + Cr(a|b) \cdot 0 = Cr(a|b)$. $\square$

B Awareness Rigidity and (Updated) Reverse Bayesianism

First recall Reverse Bayesianism.

Reverse Bayesianism. For all $a, b \in \mathcal{X}$ such that $Cr(a) > 0$ and $Cr(b) > 0$, $Cr^+(\cdot)$ over $\mathbb{B}_{\mathcal{X}^+}$ ought to be such that:

$$\frac{Cr(a)}{Cr(b)} = \frac{Cr^+(a)}{Cr^+(b)}.$$

Let $\mathcal{X} = \{\mathsf{Hyp}_1, \mathsf{Hyp}_2, \mathsf{CA}\}$ and $\mathcal{X}^+ = \{\mathsf{Hyp}_1, \mathsf{Hyp}_2, \mathsf{Hyp}_3, \mathsf{CA}'\}$. Our goal is to show that Reverse Bayesianism captures the shaving-off dynamics broached by Earman. Thus, we want to show that $Cr^+(\mathsf{CA}') = Cr(\mathsf{CA}) - Cr^+(\mathsf{Hyp_3})$. Assume that $\mathsf{CA}' \vee \mathsf{Hyp}_3$ just is the same proposition as CA, thus $\mathsf{CA}' \vee \mathsf{Hyp}_3 \in \mathcal{X}$ and similarly, $\mathsf{CA} \in \mathcal{X}^+$.

Claim 1: $Cr(\mathsf{Hyp_i}) = Cr^+(\mathsf{Hyp_i})$ for $i = 1, 2$ and $Cr^+(\mathsf{CA}) = Cr(\mathsf{CA})$.

Proof.

Without loss of generality suppose for a contradiction that $Cr^+(\mathsf{Hyp_1}) > Cr(\mathsf{Hyp_1})$. Since by assumption $\mathsf{Hyp_1} \vee \mathsf{Hyp_2} \vee \mathsf{CA} = \top$ and elements of both $\mathcal{X}$ and $\mathcal{X}^+$ are mutually exclusive, by Claim 8 of A1, $1 = Cr(\mathsf{Hyp_1}) + Cr(\mathsf{Hyp_2}) + Cr(\mathsf{CA})$ and $1 = Cr^+(\mathsf{Hyp_1}) + Cr^+(\mathsf{Hyp_2}) + Cr^+(\mathsf{CA})$. So, $Cr^+(\mathsf{Hyp_1}) > Cr(\mathsf{Hyp_1})$ implies that either $Cr^+(\mathsf{Hyp_2}) < Cr(\mathsf{Hyp_2})$ or $Cr^+(\mathsf{CA}) < Cr(\mathsf{CA})$ or both. In any case Reverse Bayesianism is violated since, for instance, $\frac{Cr(\mathsf{Hyp_1})}{Cr(\mathsf{Hyp_2})} < \frac{Cr^+(\mathsf{Hyp_1})}{Cr^+(\mathsf{Hyp_2})}$.

□

Claim 2: $Cr^+(\mathsf{CA}') = Cr(\mathsf{CA}) - Cr^+(\mathsf{Hyp_3})$.

Proof.

Reverse Bayesianism then gives us that $\frac{Cr^+(\mathsf{CA}' \vee \mathsf{Hyp_3})}{Cr^+(\mathsf{CA})} = \frac{Cr(\mathsf{CA})}{Cr(\mathsf{CA})}$. The right-hand side just equals to 1 and thus after re-arranging we get $Cr^+(\mathsf{CA}' \vee \mathsf{Hyp}_3) = Cr^+(\mathsf{CA})$. Since $\mathsf{CA}', \mathsf{Hyp_3} \in \mathcal{X}^+$, we have $\mathsf{CA}' \wedge \mathsf{Hyp}_3 = \perp$ and so by (3), $Cr^+(\mathsf{CA}' \vee \mathsf{Hyp}_3) = Cr^+(\mathsf{CA}') + Cr^+(\mathsf{Hyp_3})$. Thus, $Cr^+(\mathsf{CA}') + Cr^+(\mathsf{Hyp}_3) = Cr^+(\mathsf{CA})$. By Claim 1 of Appendix B, $Cr(\mathsf{CA}) = Cr^+(\mathsf{CA})$, so we obtain $Cr^+(\mathsf{CA}') = Cr(\mathsf{CA}) - Cr^+(\mathsf{Hyp_3})$.

$\square$

Let $\mathbb{B}_{\mathcal{X}} = \langle \mathcal{X}^{cl}, \vee, \wedge, \neg, \top, \bot \rangle$ and $\mathbb{B}_{\overline{\mathcal{X}}} = \langle \overline{\mathcal{X}}^{cl}, \overline{\vee}, \overline{\wedge}, \overline{\neg}, \overline{\top}, \overline{\bot} \rangle$ be two Boolean Algebras. Recall Updated Reverse Bayesianism and Awareness Rigidity.

Updated Reverse Bayesianism. Given an embedding $h : \mathbb{B}_{\mathcal{X}} \to \mathbb{B}_{\mathcal{X}^+}$, for all $a, b \in \mathcal{X}^{cl}$, $Cr^+(\cdot)$ over $\mathbb{B}_{\mathcal{X}^+}$ ought to be such that:

$$\frac{Cr(a)}{Cr(b)} = \frac{Cr^+(h(a))}{Cr^+(h(b))}.$$

Awareness Rigidity. Given an embedding $h : \mathbb{B}_{\mathcal{X}} \to \mathbb{B}_{\mathcal{X}^+}$, for all $a \in \mathcal{X}^{cl}$, $Cr^+(\cdot)$ over $\mathbb{B}_{\mathcal{X}^+}$ ought to be such that:

$$Cr^+(h(a)|h(\top)) = Cr(a).$$

Claim 3: Awareness Rigidity is equivalent to Updated Reverse Bayesianism.

Proof.

Awareness Rigidity $\Rightarrow$ Updated Reverse Bayesianism: We are in either of two cases. Either (i) we have a case of awareness growth by refinement or (ii) a case of awareness growth by expansion.

(i) Let $h : \mathbb{B}_{\mathcal{X}} \to \mathbb{B}_{\mathcal{X}^+}$ be a refinement embedding. In particular, this implies that $h(\top) = \overline{\top}$. By the assumption of Awareness Rigidity we have $\frac{Cr(a)}{Cr(b)} = \frac{Cr^+(h(a)|h(\top))}{Cr^+(h(b)|h(\top))}$ for arbitrary $a, b \in \mathcal{X}^{cl}$. By assumption that $h(\top) = \overline{\top}$, we then obtain $\frac{Cr(a)}{Cr(b)} = \frac{Cr^+(h(a)|\overline{\top})}{Cr^+(h(b)|\overline{\top})}$. But $\frac{Cr^+(h(a)|\overline{\top})}{Cr^+(h(b)|\overline{\top})}$ just simplifies to $\frac{Cr^+(h(a))}{Cr^+(h(b))}$.

To see that $\frac{Cr^+(h(a)|\overline{\top})}{Cr^+(h(b)|\overline{\top})}$ simplifies to $\frac{Cr^+(h(a))}{Cr^+(h(b))}$ note that by the definition of conditional probability $Cr^+(h(a)|\overline{\top})$ equals $\frac{Cr^+(h(a)\wedge\overline{\top})}{Cr^+(\overline{\top})}$. Similarly for b. In turn $Cr^+(\overline{\top}) = 1$ by Axiom (2) and so by Claim 6 of A1, $Cr^+(h(a) \wedge \overline{\top}) = Cr^+(h(a))$. Again, analogous reasoning applies to b.

(ii) Let $h : \mathbb{B}_{\mathcal{X}} \to \mathbb{B}_{\mathcal{X}^+}$ be an expansion embedding. In particular, the expansion embedding implies that $h(\top) = \bigvee_{i \in I} h(a_i)$ where $\{a_i\}_{i \in I} = \mathcal{X}$. By assumption of Awareness Rigidity $\frac{Cr(a)}{Cr(b)} = \frac{Cr^+(h(a)|h(\top))}{Cr^+(h(b)|h(\top))}$ for arbitrary $a, b \in \mathcal{X}^{cl}$. As discussed,

$Cr^+(h(a)|h(\top)) = Cr^+(h(a)|\bigvee_{i\in I} h(a_i))$, same goes for b. By definition of conditional probability $Cr^+(h(a)|\bigvee_{i\in I} h(a_i)) = \frac{Cr^+(h(a)\wedge\bigvee_{i\in I} h(a_i))}{Cr^+(\bigvee_{i\in I} h(a_i))}$.

Since $h(a) \models h(a) \wedge \bigvee_{i\in I} h(a_i)$, by definition of $\models$ and Fact 4 A1, $h(a) = h(a) \wedge \bigvee_{i\in I} h(a_i)$. So $Cr^+(h(a) \wedge \bigvee_{i\in I} h(a_i)) = Cr^+(h(a))$ and we get $Cr^+(h(a)|h(\top)) = \frac{Cr^+(h(a))}{Cr^+(\bigvee_{i\in I} h(a_i))}$. Similarly for b. But then $\frac{Cr^+(h(a)|h(\top))}{Cr^+(h(b)|h(\top))} = \frac{Cr^+(h(a))}{Cr^+(\bigvee_{i\in I} h(a_i))} \div \frac{Cr^+(h(b))}{Cr^+(\bigvee_{i\in I} h(a_i))} = \frac{Cr^+(h(a))}{Cr^+(h(b))}$. That was to show.

Updated Reverse Bayesianism $\Rightarrow$ Awareness Rigidity: By definition of conditional probability $Cr^+(h(a)|h(\top)) = \frac{Cr^+(h(a)\wedge h(\top))}{Cr^+(h(\top))}$ for arbitrary $a \in \mathcal{X}^{cl}$. Since $h(a) \models h(\top)$, by definition of $\models$ and Fact 4 A1, $h(a) = h(a) \wedge h(\top)$. Thus, $\frac{Cr^+(h(a)\wedge h(\top))}{Cr^+(h(\top))} = \frac{Cr^+(h(a))}{Cr^+(h(\top))}$. By assumption of Updated Reverse Bayesianism, we obtain $\frac{Cr^+(h(a))}{Cr^+(h(\top))} = \frac{Cr(a)}{Cr(\top)}$. By axiom (2), $\frac{Cr(a)}{Cr(\top)}$ simplifies to $Cr(a)$ and that was to show. $\square$

References

Alexandrova, A. (2018). Can the science of well-being be objective? *British Journal for the Philosophy of Science*, *69*(2), 421–445. doi: 10.1093/bjps/axw027

Arló-Costa, H., & Levi, I. (2006). Contraction: On the decision-theoretical origins of minimal change and entrenchment. *Synthese*, *152*(1), 129–154. doi: 10.1007/s11229-005-0351-4

Boutilier, C. (1996). Iterated revision and minimal change of conditional beliefs. *Journal of Philosophical Logic*, *25*(3), 263–305. doi: 10.1007/bf00248151

Bradley, R. (2017). *Decision theory with a human face*. Cambridge University Press.

Bradley, R. (2022). Attitudes, deliberation and decisions. *Synthese*, *200*(1), 1–18. doi: 10.1007/s11229-022-03468-6

Briggs, R. A., & Pettigrew, R. (2020). An accuracy-dominance argument for conditionalization. *Noûs*, *54*(1), 162–181. doi: 10.1111/nous.12258

Builes, D. (2020). Time-slice rationality and self-locating belief. *Philosophical Studies*, *177*(10), 3033–3049. doi: 10.1007/s11098-019-01358-1

Canson, C. d. (2020). *Beliefs about the unobserved* (Unpublished doctoral dissertation). The London School of Economics and Political Science (LSE).

Chihara, C. S. (1987). Some problems for bayesian confirmation theory. *The British Journal for the Philosophy of Science*, *38*(4), 551–560.

Davey, B. A., & Priestley, H. A. (2002). *Introduction to lattices and order*. Cambridge university press.

de Finetti, B. (1990). Theory of probability, vol. i (1974). *John Wiley & Sons*, *5*(8), 17.

Dekel, E., Lipman, B. L., & Rustichini, A. (1998a). Recent developments in modeling unforeseen contingencies. *European Economic Review*, *42*(3-5), 523–542.

Dekel, E., Lipman, B. L., & Rustichini, A. (1998b). Standard state-space models preclude unawareness. *Econometrica*, *66*(1), 159–173.

Diaconis, P., & Zabell, S. L. (1982). Updating subjective probability. *Journal of the American Statistical Association*, *77*(380), 822–830.

Earman, J. (1992). *Bayes or bust? a critical examination of bayesian confirmation theory*. Cambridge, ma: MIT Press.

Elliott, E. (n.d.). *Unawareness and implicit belief.*

Fagin, R., & Halpern, J. Y. (1987). Belief, awareness, and limited reasoning. *Artificial intelligence*, *34*(1), 39–76.

Finetti, B. d. (1992). Foresight: Its logical laws, its subjective sources. In *Breakthroughs in statistics* (pp. 134–174). Springer.

Gärdenfors, P. (1984). Epistemic importance and minimal changes of belief. *Australasian Journal of Philosophy*, *62*(2), 136–157. doi: 10.1080/00048408412341331

Gillies, D. (2001). Bayesianism and the fixity of the theoretical framework. In D. Corfield & J. Williamson (Eds.), *Foundations of bayesianism* (pp. 363–379). Kluwer Academic Publishers.

Glymour, C. (1980). *Theory and evidence*. Princeton University Press.

Greaves, H., & Wallace, D. (2006). Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind*, *115*(459), 607–632. doi: 10.1093/mind/fzl607

Hedden, B. (2015a). *Reasons without persons: Rationality, identity, and time*. Oxford University Press UK.

Hedden, B. (2015b). Time-slice rationality. *Mind*, *124*(494), 449–491. doi: 10.1093/mind/fzu181

Heifetz, A., Meier, M., & Schipper, B. C. (2008). A canonical model for interactive unawareness. *Games and Economic Behavior*, *62*(1), 304–324.

Hempel, C. G. (1945). Studies in the logic of confirmation (i.). *Mind*, *54*(213), 1–26.

Henderson, L., Goodman, N. D., Tenenbaum, J. B., & Woodward, J. F. (2010). The structure and dynamics of scientific theories: A hierarchical bayesian perspective. *Philosophy of Science*, *77*(2), 172–200. doi: 10.1086/651319

Hosiasson-Lindenbaum, J. (1940). On confirmation. *Journal of Symbolic Logic*, *5*(4), 133–148. doi: 10.2307/2268173

Howson, C., & Urbach, P. (2006). *Scientific reasoning: the bayesian approach*. Open Court Publishing.

Jeffrey, R. C. (1990). *The logic of decision*. University of Chicago press.

Joyce, J. M. (1998). A nonpragmatic vindication of probabilism. *Philosophy of science*, *65*(4), 575–603.

Kuhn, T. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Kuhn, T. (1977). Objectivity, value judgment, and theory choice. In *The essential tension: Selected studies in scientific tradition and change* (pp. 320–39). University of Chicago Press.

Lakatos, I. (1978). *Philosophical papers*. Cambridge University Press.

Laudan, L. (1977). *Progress and its problems: Toward a theory of scientific growth*. University of California Press.

Laudan, L. (1990). Demystifying underdetermination. In C. W. Savage (Ed.), *Scientific theories* (pp. 267–97). University of Minnesota Press.

Leitgeb, H., & Pettigrew, R. (2010a). An objective justification of bayesianism ii: The

consequences of minimizing inaccuracy. *Philosophy of Science*, *77*(2), 236–272.
Leitgeb, H., & Pettigrew, R. (2010b). An objective justification of bayesianism i: Measuring inaccuracy. *Philosophy of Science*, *77*(2), 201–235.
Lewis, D., David, L. M., et al. (1999). *Papers in metaphysics and epistemology: Volume 2*. Cambridge University Press.
Mackie, J. L. (1963). The paradox of confirmation. *The British Journal for the Philosophy of Science*, *13*(52), 265–276.
Maher, P. (1993). *Betting on theories*. Cambridge University Press.
Maher, P. (1995). Probabilities for new theories. *Philosophical studies*, *77*(1), 103–115.
Mahtani, A. (2020). Awareness growth and dispositional attitudes. *Synthese*, *198*(9), 8981–8997. doi: 10.1007/s11229-020-02611-5
McMullin, E. (1993). Rationality and paradigm change in science. In P. Horwich (Ed.), *World changes. thomas kuhn and the nature of science* (pp. 55–78). MIT Press.
Miller, D. (1966). A paradox of information. *British Journal for the Philosophy of Science*, *17*(1), 59–61.
Modica, S., & Rustichini, A. (1999). Unawareness and partitional information structures. *Games and economic Behavior*, *27*(2), 265–298.
Nguyen, J., & Marcoci, A. (2019). Objectivity, ambiguity, and theory choice. *Erkenntnis*, *84*(2), 343–357. doi: 10.1007/s10670-017-9961-8
Otte, R. (1994). A solution to a problem for bayesian confirmation theory. *The British Journal for the Philosophy of Science*, *45*(2), 764–769.
Paul, L. A. (2014). *Transformative experience*. Oxford University Press.
Pettigrew, R. (n.d.). *How should your beliefs change when your awareness grows?*
Ramsey, F. (1926). Truth and probability. In D. H. Mellor (Ed.), *F.p. ramsey: Philosophical papers* (p. 52-109). Cambridge University Press.
Rinard, S. (2014). A new bayesian solution to the paradox of the ravens. *Philosophy of Science*, *81*(1), 81–100.
Roush, S. (2005). *Tracking truth: Knowledge, evidence, and science*. Oxford University Press.
Roussos, J. (2020). *Policymaking under scientific uncertainty* (Unpublished doctoral dissertation). London School of Economics.
Salmon, W. (1990). Rationality and objectivity in science or tom kuhn meets tom bayes. In C. W. Savage (Ed.), *Scientific theories* (pp. 14–175). University of Minnesota Press.
Savage, L. J. (1954). *The foundations of statistics*. Wiley Publications in Statistics.
Savage, L. J. (1967). Difficulties in the theory of personal probability. *Philosophy of Science*, *34*(4), 305–310. doi: 10.1086/288168

Sen, A., Sen, M. A., Foster, J. E., Amartya, S., Foster, J. E., et al. (1997). *On economic inequality.* Oxford university press.

Shimony, A. (1970). Scientific inference. In R. Colodny (Ed.), *The nature and function of scientific theories* (p. 79-172). University of Pittsburgh Press.

Skipper, M., & Bjerring, J. C. (2020). Bayesianism for non-ideal agents. *Erkenntnis*, *87*(1), 93–115. doi: 10.1007/s10670-019-00186-3

Stanford, P. K. (2006). *Exceeding our grasp:science, history, and the problem of unconceived alternatives: Science, history, and the problem of unconceived alternatives.* Oxford University Press.

Steele, K., & Stefánsson, H. O. (2021a). Belief revision for growing awareness. *Mind*, *130*(520), 1207–1232. doi: 10.1093/mind/fzaa056

Steele, K., & Stefánsson, H. O. (2021b). *Beyond uncertainty: Reasoning with unknown possibilities.* Cambridge University Press.

Strevens, M. (n.d.). *Notes on bayesian confirmation theory.*

Strevens, M. (2010). Reconsidering authority: Scientific expertise, bounded rationality, and epistemic backtracking. *Oxford studies in epistemology*, *3*, 294–330.

Teller, P. (1975). Shimony's a priori arguments for tempered personalism. In *Induction, probability, and confirmation. minnesota studies in the philosophy of science, volume 6* (pp. 166–203). University of Minnesota Press.

van Benthem, J. (2007). Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics*, *17*(2), 129–155. doi: 10.3166/jancl.17.129-155

Walker, O. (2014). Unawareness with "possible" possible worlds. *Mathematical Social Sciences*, *70*, 23–33.

Wenmackers, S., & Romeijn, J. (2016). New theory about old evidence. a framework for open-minded bayesianism. *Synthese*, *193*(4). doi: 10.1007/s11229-014-0632-x

Williamson, T. (2000). *Knowledge and its limits.* Oxford University Press.