# Awareness Growth

Daniel Rothschild, UCL, Term 3, 2021-22

## 1 Changing the Event Space

Recall Bayesian conditioning relies on the idea that in light of new evidence $E$ you update your probabilities $p$ by conditioning.

> Or something similar in the case of Jeffrey Conditioning and its ilk.

All this presupposes a single, unchanging event space $\mathcal{F}$ which you assign probabilities to. This can be thought of something like the agents *conceptualization* of the relevant possibilities. When this changes we have some form of *conceptual change*.

> Recall that $\mathcal{F}$ is a $\sigma$ alegbra over the set of worlds $W$

### 1.1 Examples

PROFOUND CONCEPTUAL CHANGE The case most discussed in the Bayesian literature is that of Kuhnian scientific revolutions.

> Earman [1992] is a classic discussion.

REFINEMENT Previously you thought Ernesto would either get strawberry or vanilla ice cream. Now you realize she can get strawberry, Madagascan vanilla, or Mexican vanilla.

A famous example:

"Is there any other point to which you would wish to draw my attention?"

"To the curious incident of the dog in the night-time."

"The dog did nothing in the night-time."

"That was the curious incident," remarked Sherlock Holmes.

> From "Silver Blaze" (1894), note Holmes is talking to Inspector Gregory, not Watson as commonly supposed in the secondary literature, starting with Geanakoplos [2021].

| 90 | Intruder |
|----|----------|
| 20 | No intruder |

vs.

| 10 | Intruder and no dog |
|----|----------|
| 80 | Intruder and dog barking |
| 20 | No intruder |

NEW POSSIBILITIES Previously you thought Sylvia would either get strawberry or vanilla ice cream. Now you realize he can also get chocolate!

> Note this case relates to learning you were wrong, since you previously thought the options were exclusive.

NOTE How we divide up these two kinds of cases in terms of redefining event space is complex. The problem is we don't usually have grip on what the underlying worlds are, so we will assume all that matters is the structure of the event space.

> We have two things to play with, the underlying set of worlds, and the event-space over it.

### 1.2 Ur-Prior Bayesianism

The ur-prior Bayesian says the only requirement is to have your credences at any one time determined by conditioning a total probability function on your

> See, Williamson [2000], Meacham [2016], and Das [2019]

total evidence. The probability function does not have to be your previous probabilities but can be some 'objective' priors. So no diachronic norm, and no particular problem about awareness growth.

## 1.3   Conditionalizing Generalized: reverse Bayesianism

Can we generalize conditioning to cover awareness growth? A standard proposal called *reverse Bayesianism* has been discussed and much criticized in the recent philosophical literature.

Karni and Vierø [2013], see Bradley [2017, ch. 12] for a more informal version.

The basic idea of reverse Bayesianism is that when we add or refine possibilities we ought to maintain the ratio of credences in old possibilities.

Actually implementing this is a bit tricky for the new possibility case. Take the case where you previously though it was 50/50, vanilla and strawberry. Then you consider the chocolate possibility. You might think you are forced to keep the relative possibility of vanilla and strawberry constant but add some chance.

i.e. 40/40/20, 33/33/33 are okay for V/S/C, but not 40/30/30 is not.

There is a technical problem here: if you want to keep the ratios the same in all proposition you have to keep the ratio of the probability of $A$ to the probability of $\neg A$ the same. But to do that you would need to keep the same probability in $A$. So we need to restrict which probabilities we keep the ratios constant in to some subset of probabilities (and not include any "negative propositions").

Thus even stating reverse Bayesianism requires a new concept, what is called "basic propositions"

**Reverse Bayesianism**  The ratio of credences in basic propositions does not change when you add more possibilities.

Note that it is not at all clear what reverse Bayesianism can say in cases in of *radical* theory change like scientific revolutions.

## 1.4   Counterexamples to Reverse Bayesianism

Here's a simple counterexample to Reverse Bayesianism. Considers a case in which you are in a flat and someone is singing. You assume it is either the landlord or Bob a tenant. You assign 50/50. Now you consider the possibility it is some other tenant. If the basic propositions are Bob, Tenant, and Landlord, you cannot keep their ratios the same and assign any probability to it being another tenant.

From Mahtani [2020]

Note the ratio of Landlord to Bob has to equal the ratio of Landlord to Tenant

Here's a case with refinement not adding possibilities:

Steele and Stefánsson [2021]

Suppose you are deciding whether to see a movie at your local cinema. You know that the movie's predominant language and genre will affect your viewing experience. The possible languages you consider are French and German, and the genres you consider are thriller and comedy. But then you realize that, because of your poor French and German skills, your enjoyment of the movie will also depend on the level of difficulty of the language. Since it occurs to you that the owner of the cinema is quite simple-minded, you are, after this realization, much more confident that the language of the movie will be simple than that it will be difficult. Moreover, since you associate simple language with thrillers,

this makes you more confident than you were before that the movie on offer is a thriller as opposed to a comedy.

Seems like this is a case of correcting sloppy reasoning before as in Steele and Stefánsson [2021, 1220]. Their response is that the reasoning required awareness growth so is a product of it.

> More distinctions among cases might be useful here.

In general, it does seem like merely drawing distinctions among cases should not alter your credences unless it makes you aware of some mistake in reasoning.

> For example, if you think it's likely you'll be given a red wine for dinner, distinguishing all the different countries from which it might come should not change your credences.

## 2   Types of Unawareness

There's been a large literature on modelling unawareness in economics which the current philosophical interventions are building on.

> Geanakoplos [2021] and Dekel et al. [1998] are classics in this lit.

From that literature we can find some different notions:

> Borrowing from Fritz and Lederman [2015].

1. Unawareness due to inattention.
   *Examples.* Holmes and the dog barking. The difficulty of language in a movie.

2. Unawareness due to conceptual limitations.
   *Example.* Not seeing possibility of public key encryption when trying to send a coded message in the early 20th century.

3. Unawareness due to processing limits.
   Not seeing what possibilities a sentence picks out due to lack of logical omniscience.

Seems like we are thinking primarily of types 1 and 2 here.

### 2.1   Basic formula of counterexamples

The design of the counterexamples is that the new possibilities itself have evidential impact on cells in the old partition. No way to in principle avoid this, unless you try to make a distinction between awareness and evidence that seem inextricably linked.

> This from Steele and Stefánsson [2021]

> It might be argued that our examples are not illustrative of a simple learning event (a simple growth in awareness); rather, our examples illustrate and should be expressed formally as complex learning experiences, where first there is a growth in awareness, and then there is a further learning event that may be represented as, say, a Jeffrey-style or Adams-style learning event. In this way, one could argue that the awareness-growth aspect of the learning event always satisfies Reverse Bayesianism (the new propositions are in the first instance evidentially irrelevant to the comparison of the old basic propositions). Subsequently, however, there may be a revision of probabilities over some partition of the possibility space, resulting in changes to the ratios of probabilities for the old basic propositions. The reason we reject this way of conceiving of the learning events described by our examples is that the two-part structure seems ultimately unmotivated. The

second learning stage is an odd, spontaneous learning event that would be hard to rationalize. Hence this would again seem to us to be an artificial and ad hoc way to save Reverse Bayesianism.

## 3 Modelling Awareness Growth

There are some decision points on how to model awareness growth. We could have one 'total' set of worlds $W$ and states space $\mathcal{F}$ and model all limited awareness as $\sigma$-algebras that are proper subsets of $\mathcal{F}$.

*A conservative approach*

Steele and Stefánsson [2021] instead assume what grows is the primitive (finite) set of basic propositions, **X**. We define the probability space out of **X** by taking worlds to be functions from **X** to $0, 1$ and propositions to be set of worlds. We can then identigy propositions by all the valuations that make them true and thus allow continuity of propositions across awareness growth.

*I don't really understand the motivation...*

### 3.1 Rescuing Conservativity

Reverse Bayesianism seems either *false* (due to counterexamples) or *vacuous* (by classifying all awareness growth that leads to violations of it as not pure awareness growth). So is there a better version?

Steele and Stefánsson [2021] consider an interesting principle that only applies in the expanded possibility case.

$$P(A|\top_o) = P_o(A)$$

*Where $\top_o$ corresponds to all the old possibilities in the new framework.*

Effectively new possibilities don't change probabilities conditional on those possibilities not arising. This principle is actually quite close to reverse Bayesianism, as they note for cases of refinement, but also for other cases.

They retreat back to claiming the only cases in which reverse Bayesianism holds are the ones in which new possibilities are evidentially irrelevant (so the vacuous version of RB).

*See Mahtani [2020] for a criticism of a formal definition of evidential relevance that is I think not in their published paper.*

In the end, perhaps there is no formal theory of how you should change your beliefs in the face of awareness growth?

*Note for refinement, at least, ceteris parabis, refining the partition should not change probabilities, so something right in RB.*

## 4 Pettigrew on Awareness Growth

Pettigrew [2022] takes a different approach and asks how dutch book and accuracy arguments for conditioning can be extended to cover the case of awareness growth.

He distinguishes, as we have, between *pragmatic* arguments, based on dutch books, and *accuracy* arguments based on epistemic value for conditioning. Another important distinction is between *planning* arguments which concern how you should plan to update you credences and *direct* arguments about how you actually should adjust your credences.

Some of the planning arguments go indirectly by way of what Pettigrew calls the **Weak Reflection Principle**.

**Weak Reflection Principle**  Suppose

1. $c$ defined on $\mathcal{F}$ is your prior credence function;

2. $r$ is your updating plan;

3. $r'_1, \ldots, r'_n$ defined on $\mathcal{F}'_1, \ldots, \mathcal{F}'_n$ are all the possible credence functions that your updating plan might require you to adopt, where $\mathcal{F} \subseteq \mathcal{F}1', \ldots, \mathcal{F}'_n$;

4. $r_1, \ldots, r_n$ are the restrictions of $r_1, \ldots, r_n$ to $F$.

Then it should be that $c$ is a mixture of $r_1, \ldots, r_n$. That is, there should be non-negative real numbers $\lambda_1, \ldots, \lambda_n$ that sum to 1 such that, for all propositions $X$ in $\mathcal{F}$,

$$c(X) = \lambda_1 r_1(X) + \ldots + \lambda_n r_n(X)$$

The justification for this principle is a bit complex. Standardly we justify such a position by saying that otherwise you must be planning to change your probabilities (in one direction) and there would be some other probability that you would know would better match your future probabilities (and hence be more accurate or result in better decisions). This doesn't extend that well to the awareness case as we can't be *aware* of how we will expand awareness.

As Pettigrew says, "As soon as I consider a particular way in which my awareness might grow, it immediately grows in exactly that way."

For example in the Holmes case if the only possible expansion of awareness is the dog barking, then we would, of course, violate this principle if it shifts one's beliefs in old propositions. Pettigrew tries to evade this general worry as follows:

> I think there are two responses to this concern. On the first, we think of updating plans not as commitments that we consciously entertain and adopt, but rather as dispositions to update in a particular way. You can easily have a disposition to do something in response to a particular stimulus without being at all aware of the possibility of that stimulus. So the worry from above does not arise if we conceive of updating plans like that. On the second response, we think of updating plans not as mental states that an individual is ever actually in, whether consciously or unconsciously, whether dispositions or intentions, but rather as devices we use to determine the rationality of an individual's posterior. So, instead of imagining that the individual actually has an updating plan and then assessing its rationality and demanding that they follow through on that plan when their evidence or their awareness grows, we instead look at how they actually updated and say that they did so rationally if there is any updating plan that it would have been rational for them to adopt had they been able to adopt it, and which would have given rise to their actual posterior had they followed it.

In the end, Pettigrew argues that the requirement imposed by Weak Reflection is incredibly weak: any form of bizarre change in belief is fine to have possible as long as you balance it out with another one.

I am not convinced a) that the Weak Reflection norm for planning is toothless, or b) that it is valid.

Pettigrew [2022] also considers direct arguments for conditioning, focusing on Gallow [2019]. As far as I can tell Gallow's argument works by assuming

For instance, in the Holmes case the expansion could satisfy weak reflection if something else could be thought of to balance the changes thinking about the dog barking would add.

More specifically I don't really see why necessarily unconscious dispositions can be rational/irrational if they are not sustainable if made conscious.

that the epistemic value function changes in result to learned proposition. (By no longer caring about what you believe in worlds learned not to be possible.) Then one can argue that the best function one can have by the lights of the previous one is gotten by conditioning.

Pettigrew suggests that, in the case of refinement, your epistemic utility function does not change with respect to the propositions in the old partition so should lead you to condition with respect to those.

He suggests though that in the case of expansion, your priors don't determine your beliefs in the old propositions, so the new utility function itself doesn't tell you what to believe.

Epistemic utility functions assign scores for your beliefs in each proposition: a weird feature of them.

# References

Richard Bradley. *Decision Theory with a Human Face*. Cambridge University Press, 2017.

Nilanjan Das. Accuracy and ur-prior conditionalization. *Review of Symbolic Logic*, 12(1):62–96, 2019.

Eddie Dekel, Barton L Lipman, and Aldo Rustichini. Standard state-space models preclude unawareness. *Econometrica*, 66(1):159–173, 1998.

John Earman. *Bayes or Bust*. MIT Press, 1992.

Peter Fritz and Harvey Lederman. Standard state space models of unawareness. *Theoretical Aspects of Rationality and Knowledge*, 15, 2015.

J. Dmitri Gallow. Learning and value change. *Philosophers' Imprint*, 19:1–22, 2019.

John Geanakoplos. Game theory without partitions, and applications to speculation and consensus. *The B.E. Journal of Theoretical Economics*, 21(2): 361–394, 2021.

Edi Karni and Marie-Louise Vierø. "reverse bayesianism": A choice-based theory of growing awareness. *American Economic Review*, 103(7):2790–2810, 2013.

Anna Mahtani. Awareness growth and dispositional attitudes. *Synthese*, 198 (9):8981–8997, 2020.

Christopher J. G. Meacham. Ur-priors, conditionalization, and ur-prior conditionalization. *Ergo: An Open Access Journal of Philosophy*, 3, 2016.

Richard Pettigrew. How should your beliefs change when your awareness grows? unpublished manuscript, 2022.

Katie Steele and H. Orri Stefánsson. Belief revision for growing awareness. *Mind*, 130(520):1207–1232, 2021.

Timothy Williamson. *Knowledge and its Limits*. Oxford University Press, 2000.