Testing the Maximum Entropy Approach to Awareness Growth in Bayesian Epistemology and Decision Theory

Rafael Fuchs (Rafael.Fuchs@campus.lmu.de)
Munich Center for Mathematical Philosophy and Graduate School of Systemic Neurosciences, LMU Munich
80539 Munich, Germany

Marko Tešić (mt961@cam.ac.uk)
Leverhulme Centre for the Future of Intelligence, University of Cambridge
Cambridge, CB2 1SB, U.K.

Ulrike Hahn (u.hahn@bbk.ac.uk)
Centre for Cognition, Computation, and Modelling, Birkbeck, University of London
London, WC1E 7HX, U.K.

Abstract

In this paper, we explore the objective-Bayesian principle of minimum information and Maximum Entropy as a solution to the problem of awareness growth: how should rational agents adjust their beliefs upon becoming aware of new possibilities? We introduce the Maximum Entropy principle as a theoretical solution to the problem of awareness growth and present the results of two experiments conducted to compare human reasoners' responses with the theoretical prescriptions of the Maximum Entropy approach. We discover that, although the MaxEnt method may appear computationally demanding, participants' responses are largely consistent with the theoretical prescription.

Keywords: Maximum Entropy estimation; sequential causal structure learning; sequential causal reasoning; causal Bayesian networks

Introduction

Uncertainty is a pervasive feature of decision making and scientific investigation. Bayesian models provide a strong normative standard for rational belief updating under uncertainty. Relative to a given probability distribution over a fixed set of alternatives, updating on evidence leads to an optimal posterior degree of belief (Rosenkrantz, 1992). However, often uncertainty runs even deeper than what is captured by standard Bayesian models: in order to entertain a probability of an event, we must at least be aware that said event is possible. However, there are many things we are not aware of, and so we have to adjust our beliefs once we learn about a new relevant alternative. Therefore, it is an important question how agents should change their rational beliefs if they learn about new alternative explanations of an observed event, or if their set of options under consideration increases. Is it possible to find a more general normative account of optimal belief updating after learning about new possibilities?

In this paper, we present and test the Maximum Entropy (MaxEnt) approach, which lies at the heart of objective Bayesianism (Jaynes, 1968; Williams, 1980; Williamson, 2002, 2003, 2007). In the first section, we explain the problem of awareness growth within the Bayesian framework, providing an illustrative running example. In the subsequent section, we introduce the principle of MaxEnt, and derive a theoretical solution for the running example. Finally, we present the results of two experiments conducted to test whether human reasoners' behavior aligns with the MaxEnt principle. Our findings suggest that, while the MaxEnt approach may seem computationally intensive, participants' responses are largely in line with the theoretical prescriptions.

Background: The Problem of Awareness Growth in Bayesian Epistemology and Decision Theory

In this section, we present the problem of awareness extension within a Bayesian setting. To make the problem intuitively accessible, we will use the following background story as a running example:

> Anne gets up in the morning and remembers that according to yesterday's weather forecast rain during the night was unlikely. Furthermore, she knows that if it rained, the lawn is almost certain to be wet while otherwise, the lawn is very likely to remain dry.

We can represent Anne's initial beliefs by a Bayesian Network; the directed acyclic graph (DAG) in figure 1 represents the causal structure, where R (rain) and W (wet lawn) are two-valued random variables:

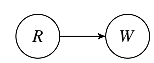

Figure 1: Causal DAG of the initial background story: rain (R) contributes to the wetness of the lawn (W)

R takes either value r (meaning that it rained during the night) or $\neg r$ (it didn't rain), and W takes either value w (the lawn is wet in the morning) or $\neg w$ (the lawn is not wet).

If we specify the qualitative assessments from the story to precise values (for illustrative purposes), we obtain the following probability distribution over R and S:

$$P(r) = 0.15, \quad P(w|r) = 0.99, \quad P(w|\neg r) = 0.05$$

From this, we can calculate Anne's initial expectation that the lawn will be wet, when she opens the window, via the law of total probability:

$$\begin{aligned} P(w) &= P(w|r)P(r) + P(w|\neg r)P(\neg r) \\ &= 0.99 \cdot 0.15 + 0.05 \cdot 0.85 = 0.191 \end{aligned}$$

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

Suppose that Anne now finds that the lawn is indeed wet. Consequently, she updates her degree of belief that it rained by standard Bayesian conditionalization:

$$P(r|w) = \frac{P(w|r)P(r)}{P(w)} = \frac{0.99 \cdot 0.1}{0.191} = 0.7857$$

This situation is the Bayesian standard case, which results in an optimal posterior degree of belief, given the initial probability values. However, what if later on Anne learns about an alternative cause of the lawn being wet? For example, Anne could learn that her neighbor bought a sprinkler, and if he turned it on during the night, the lawn is likely to be wet. How does this affect the posterior probability of rain? More generally, how should Anne change her beliefs, in response to learning about the new possibility?

These questions are related to a formal issue in Bayesian epistemology and decision theory, known under the labels of "domain extension", "algebra extension" or "awareness growth" (Karni & Vierø, 2015; Bradley, 2017; Steele & Stefánsson, 2021). Conceptually, the question concerns the relevant rationality constraints for learning new possibilities (causes, explanations), relative to a given set of observations. How much probability should be assigned to the new alternatives, and how does this affect the old alternatives? In a more technical sense, there are two kinds of awareness extension, discussed in (Steele & Stefánsson, 2021): refinement and expansion. We focus on the former case (refinement): this corresponds to adding a new random variable X to a propositional algebra $\mathcal{A}$, whereas expansions correspond to adding a new *value* to a random variable X (in this case, some of the old probabilities over X must decrease, so the new value fits in). Thus, the question is – as in the conceptual case – how this should be achieved, and what the relevant rationality constraints are. Importantly, even in the case of refinements, where changes of the old probabilities may not be *technically* necessitated (so that probabilities add up to 1), becoming aware of new possibilities may still be evidentially relevant for the assessment of the old prospects.

In the literature, general rationality conditions are discussed (Steele & Stefánsson, 2021), but there is no agreement on a single general updating procedure like Bayesian updating in the case of a fixed algebra. Most normative accounts of rational awareness extensions discussed in the literature don't propose unique prescriptions like Bayes' rule for updating (Bradley, 2017).

However, there is *one* proposal that makes more concrete recommendations that seem promising for moving towards a robust and general solution. This is the maximum entropy (MaxEnt) approach that is at the core of objective Bayesianism (Williams, 1980; Williamson, 2003). In the next section we will introduce MaxEnt, and use it to formally analyse two variations of our initial example, where the agent learns about the sprinkler before or after observing the wet lawn. We will show that the principle of minimum information (Williams, 1980)—i.e. MaxEnt for prior assignment and a corresponding updating procedure for learning evidence—can be used so that the final result of an awareness extension in combination with evidential learning (before or after the extension) will coincide with pure evidential updating of a fully aware agent. Finally, we present our experimental results where we investigated how people reason with each to each of these variations, whether their responses coincide with the MaxEnt solution, as well as whether there are effects that result from the order of learning the events.

The Maximum Entropy Solution

Maximum Entropy is the objective Bayesian approach to rational prior assignment, given any set of prior information. It consists in maximising the quantity

$$H(P) = -\sum_{\omega \in \Omega} P(\omega) \log P(\omega), \qquad (1)$$

which is called *information entropy* in Claude Shannon's (1948) mathematical theory of communication. There, it serves as a measure of expected information content. The global maximum of $H(P)$ is given when all events have the same probability, i.e. the distribution is uniform (conversely, $H(P)$ is minimal if there is exactly one ω' with $P(\omega') = 1$, i.e. the distribution is maximally concentrated). Prior information (e.g. about correlations between variables) can constrain the space of admissible distributions, and thus, entropy has to be maximised relative to these constraints (Jaynes, 1968). More generally, the *principle of minimum information* (Williams, 1980) states that rational agents should assume priors that are maximally equivocal, i.e. as uniform as possible relative to their prior information, and updates on new information should only deviate from the prior only as much as necessitated by the new constraints. This can be achieved by minimising the Kullback-Leibler divergence between posterior and prior distribution, which is a dynamic counterpart of MaxEnt, and generalises Bayesian conditoinalisation (Williams, 1980).

Objective Bayesians defend the principle of minimum information based on considerations of caution and minimising worst-case expected inaccuracy (Williamson, 2007; Landes & Williamson, 2013).

Importantly, the principle of minimum information allows an agent that only becomes sequentially aware of new possibilities to recompute their priors and all subsequent updates, so that their final belief eventually coincides with that of an ideal agent who was aware of all possibilities from the start. We briefly illustrate this with the sprinkler example, before testing how human reasoners change their beliefs upon becoming aware of new possibilities.

We now present two variations of our initial background story:

Variation 1: First thing after leaving the house, Anne observes that the lawn is wet ($P(w) = 1$), so she updates her belief to $Q(r) = P(r|w)$, as above. After that, she learns that her neighbor has bought a new sprinkler, but it is very unlikely that the sprinkler was actually turned on overnight – but if it was turned on, it would be likely that the lawn is wet.

Variation 2: Anne has observed that the lawn is wet, and learned about the existence of the new sprinkler. Before she leaves, her neighbor tells Anne that the sprinkler went off accidentally.

Let us now consider the MaxEnt solutions for each case. In Variation 1, we have a prior for S (the sprinkler was very unlikely to be turned on, say $P(s) = 0.01$), and a likelihood (*if* the sprinkler was turned on, it would be likely that the lawn was wet, say $P(w|s) = 0.8$). Furthermore, we know that S and R are independent, i.e. $P(s|r) = P(s)$. Thus, in order to obtain a prior distribution, we apply the following constraints:

$$P(r) = 0.15, \quad P(r|s) = P(s) = 0.01, \quad P(w|r) = 0.99, \\ P(w|\neg r) = 0.05, \quad P(w|s) = 0.8$$

With these constraints, we obtain the following prior distribution via MaxEnt:

$$P(r) = 0.15, \quad P(s) = 0.1, \quad P(w|r,s) = 0.99, \\ P(w|\neg r,s) \approx 0.765, \quad P(w|r,\neg s) \approx 0.99, \\ P(w|\neg r,\neg s) \approx 0.043$$

with the corresponding DAG representation in figure 2.

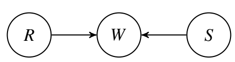

Figure 2: Causal DAG of the extended system: rain (R) and (S) are independent causes of a wet lawn (W)

This solution corresponds to the prior of an ideal agent who is fully aware of all relevant variables, as depicted in the Bayesian Network in Figure 2. Relative to the given constraints, the ideal agent obtains the specified distribution by MaxEnt. Hence, the bounded agent who becomes gradually aware of S as an alternative cause of W can recompute the prior distribution over all three variables, and then recompute their updated posterior distribution on the given observation by minimising the relative entropy (KL-divergence) between posterior and prior. Note that minimising relative entropy yields Bayesian conditionalisation as a special case, if the only new constraint is $P(w) = 1$ (Williams, 1980).

In Variation 1, where S is added to the algebra, and the prior is computed as above, the only new observation is $P(w) = 1$. Hence, the posterior $P(r|w)$ is computed by standard conditionalisation, which is the same as in the beginning, since $P(w)$ is already fixed by the constraints on $P(r), P(w|r), P(w|\neg r)$, and hence, $P(r|w) = 0.7857$.

This result can change slightly, if we remove the constraint $P(w|\neg r)$, i.e. there is no prior information on the probability of a wet lawn if it didn't rain (analogously, the likelihood $P(w|\neg s)$ was left unspecified). In this case, MaxEnt would largely give the same result, with the only difference being that $P(w|\neg s,\neg r) = 0.5$. Furthermore, for the smaller algebra R,W MaxEnt would yield $P(w|\neg r) = 0.5$, if left unconstrained. This would entail only slightly different results regarding the posterior $P(w|r)$ for distribution P_1 (over the small algebra R,W) and P_2 (over the full algebra R,S,W). In particular, $P_1(r|w) \approx 0.2579$ and $P_s(r|w) \approx 0.2589$.

In Variation 2, Anne also learns that $P(s) = 1$, together with $P(w) = 1$. Hence, she updates on both observations, and obtains $P(r|w,s) \approx 0.186$.

Hence, Anne's belief in r has dropped significantly in comparison to $P(r|w)$, due to the alternative explanation s. Hence, in qualitative terms, the MaxEnt solution yields the following results: upon learning a new variable S, with a prior $P(s)$ and an associated (partially specified) likelihood $P(w|s)$ (which is positively relevant for w), the result for $P(r|w)$ are the approximately identical in both variations (with a minimal differences if Anne has no prior information about $P(w|\neg r)$). Furthermore, upon learning $P(s) = 1$, we always obtain $P(r|w) > P(r|w,s) > P(r)$.

Experiments overview

Next present two experiments where we explore (i) how people revise their beliefs when they become aware of the elements of the algebra sequentially, as opposed to knowing all the elements from the start, and (ii) whether and how people change their beliefs when the old probability constraints are revised.

Experiment 1

In Experiment 1 we have three conditions. In the first condition ('Full') participants are told all the elements of the algebra (the two causes and one effect) before making any judgements. In the other two conditions ('Seq' and 'Seq aware') the algebra was introduced sequentially. The difference between the two sequential conditions was that in Seq aware one additional question was asked compared to the Seq condition: namely, after learning about the existence of the second cause (sprinkler) but before learning whether or not the second cause occurred, participants were asked to judge the probability of the first cause (rain) given that they knew the effect occurred (the lawn was wet). The purpose of this questions was to provide insight into how people update their beliefs (if at all) after becoming aware of the mere presence of another potential factor without knowing whether that factor occurred.

Participants & Design A total of 237 participants ($N_{\text{FEMALE}} = 147$, $M_{\text{AGE}} = 41.5$ years) were recruited from Prolific Academic (`www.prolific.co`). All participants were native English speakers residing in the UK, the US, or Canada with approval ratings of 95% or higher. They all gave informed consent and were paid £6.66 an hour rate for participating in the present study, which took on average 7.5 min to complete.

Participants were randomly assigned to one of the three conditions: *Full* condition ($N = 81$), *Seq* condition ($N = 76$), and *Seq aware* condition ($N = 80$).

Materials & Procedure To explore participants' probabilistic judgments, we created four scenarios. Each scenario constructed a background story for a common-effect model with two causes and one effect. One scenario featured rain and a sprinkler as causes with a wet lawn as the effect; another scenario depicted sleep deprivation and diuretics as causes of low magnesium levels; the third scenario described a weak car battery and starter motor issues as causes of a slow car start; and the final cover story identified hunger and thirst in dogs as causes of agitation in dogs.

After giving informed consent and basic demographic information, participants were presented with the four scenarios and questions related to these scenarios. The order in which the scenarios were presented was randomized for each participant.

For example, the *Full* condition for the rain, sprinkler, and lawn scenario looked as follows:

> One morning, you decide to do some gardening at a community garden just a mile away from your home. However, you will have to postpone gardening for another day if it had rained last night. You slept tightly and do not remember hearing any rain, but you recall the weather forecast saying that it was **unlikely** to rain overnight.
>
> You head outside to check the lawn outside your home, which is **almost always** wet in the morning if it rained overnight.
>
> On your way outside you realize you have to be careful in judging whether it rained last night on the basis of whether the lawn is wet. Your neighbor has a lawn sprinkler that can *accidentally* turn on overnight and make your lawn wet. This happens rarely and you believe that it is **very unlikely** that the sprinkler turned on overnight.
>
> If, however, the sprinkler turns on overnight, then the lawn is **often** wet in the morning.

The above introduced both causes (rain r, and sprinkler s) and the effect (wet lawn w). It also communicated to the participants the prior probabilities of each cause ($P(r)$ and $P(s)$) and the likelihoods ($P(w|r)$ and $P(w|s)$). The priors and the likelihoods were communicated to the participants in a verbal manner. This is to reduce the amount of the quantitative information which could anchor the participants responses (see e.g. Tešić & Hahn, 2019; Tešić, Liefgreen, & Lagnado, 2020). Nonetheless, reasoning with verbal information, particularly in the context of sequential reasoning, has been shown to be similar to reasoning with numeric information (Meder & Mayrhofer, 2017).

After the scenario and the relevant probabilities were communicated, participants were asked two questions regarding the prior probabilities of the two causes, i.e. about $P(r)$ and $P(s)$:

> **Q1**. How likely is it that it **rained** overnight?
>
> **Q2**. How likely is it that **the sprinkler turned on** overnight?

To provide answers to the two questions (and all other quantitative questions in this study), participants were asked to move a slider that ranged from 0% to 100%. The participants were then told that the lawn was wet ($P(w) = 1$), but it was still unknown whether the sprinkler had been on overnight. Participants were then reminded of the priors and the likelihoods for both causes and asked a questions eliciting their judgements regarding the probability of rain after learning the lawn was wet, i.e. $P(r|w)$:

> **Q3**. How likely is it that it **rained** overnight now that you know the lawn is wet?

This question was followed by a free format type text box where participants could explain their reasoning for selecting certain confidence/reliability estimates.

Participants were then told that the neighbour has informed them that the sprinkler was on overnight and they had to turn it off early in the morning. This new information was followed by a question eliciting the probability of rain after learning that the sprinkler was on and the lawn was wet, i.e. $P(r|w,s)$:

> **Q4**. How likely is it that it rained overnight now that you know the lawn is wet and the sprinkler turned on overnight?

Finally, participants were told that they just remembered that the sprinkler has water pressure issues, and if that occurs, the sprinkler doesn't reach the lawn. This new information leads them to revise their estimate of the frequency with which the sprinkler causes the lawn to be wet. They are now told to believe that the sprinkler only **sometimes** makes the lawn wet if it accidentally turns on overnight. They are then asked the questions about the probability of rain given this new constraint, i.e. $P(r|w,s,l\_diff)$, where l_diff indicates that the likelihood for the sprinkler is now different:

> **Q5**. How likely is it that it rained overnight now that you know the sprinkler only **sometimes** causes the lawn to be wet (you also know that the lawn is wet and the sprinkler turned on overnight)?

The format for the other three scenarios was the same. After the participants answered questions for all four scenarios they received debriefing information.

The participants in the sequential condition (*Seq*) were initially informed only about the rain and the lawn, including the prior for rain ($P(r)$) and the likelihood $P(w|r)$. They were not informed about the sprinkler. They first assessed the prior probability of rain (question **Q1** from above). After learning about the wet lawn, they were asked to judge the conditional probability $P(r|w)$ (question **Q3** from above). Participants were then made aware of the sprinkler, informed about the prior associated with the sprinkler ($P(s)$), and the likelihood ($P(w|s)$). Subsequently, they assessed the prior probability for the sprinkler (question **Q2** from above). As in the *Full* condition, they were informed that the sprinkler was on overnight and asked to judge $P(r|w,s)$ (question **Q4** from above). They were then told about the new likelihood for the sprinkler and asked to judge $P(r|w,s,l\_diff)$ (question **Q5** from above). The third condition (*Seq aware*) was exactly similar to the *Seq* condition except that after being told about the existence of the sprinkler, the associated prior, and the likelihood, but crucially before learning whether the sprinkler was on or off overnight, they were asked to judge the probability of rain, i.e. $P(r|w, \mathrm{A\_s}$, where A_s indicates that participant were made aware of the sprinkler but were not told whether it was on or not:

> **Q4**. How likely is it that it rained overnight now that you know the lawn is wet ***and*** you are aware of the lawn sprinkler?

After answering this question, participants were informed that the sprinkler was on, and the rest of the condition proceeded exactly the same as in the *Seq* condition.

Results & Discussion Participants' responses to all the questions in all three conditions in Experiment 1 are shown in Figure 3. We first note that participants' numeric estimates for the prior probabilities of the two causes were low across the three condition suggesting that the participants accepted the priors. In addition, these estimates were consistent across the conditions, indicating that even though these probabilities were communicated verbally, they still led to consistent estimates. Participants' estimates also exhibited a typical 'explaining away' pattern (Wellman & Henrion, 1993; Morris & Larrick, 1995), where the probability of one cause after learning about the presence of the effect ($P(C1|E)$) is higher than the prior probability of that cause ($P(C1)$). Then, after learning that the second cause has occurred, the probability of the first cause ($P(C1|E,C2)$) is again lower than the probability of that cause before learning about the presence of the second cause, i.e., $P(C1|E)$.

To analyze the data, we built a linear mixed-effects model (LMM) using the "lme4" package in R (Bates, Mächler, Bolker, & Walker, 2014). Our model included two fixed effects: Condition (*Full*, *Seq*, and *Seq aware*) and Question (5 levels), along with their interaction. The Question fixed effect accounted for all the questions regarding the probability estimates except for the question eliciting $P(r|w, \mathrm{A\_s}$, as this question was only asked in the *Seq aware* condition. The model's only random effect was the intercept for participants, with no random slope from the participant since the design was fully between subjects.

The LMM indicated a significant effect of Question ($t(948) = 6.3$, $p < .001$), a non-significant effect of Condition ($t(237) = 1.15$, $p = .25$), and no interaction between the two fixed effects ($t(948) = -0.81$, $p = .42$). This suggests that participants did change their estimates depending on the question they were asked. However, the non-significant effect of Condition suggests that participants did not adjust their probability estimates differently when learning the algebra sequentially compared to when the full algebra was presented at the start. This coincides with the prescription of the principle of minimum information, as explained in the previous section. From Figure 3, we also observe that in the *Seq aware* condition, participants' estimates for $P(C1|E, \mathrm{A\_}C2)$ were lower than their estimates for $P(C1|E)$. This suggests that merely learning about the existence of another variable (another cause) can affect people's beliefs, even if they do not know whether that variable is instantiated or not (whether the cause occurred). By contrast, in the theoretical MaxEnt solution we have observed that there can be a slight difference between the small- and the large algebra, if the agent does not have any prior knowledge about $P(w|\neg r)$. Since no prior information about $P(w|\neg r)$ was given to the participants, only slight differences might still be consistent with MaxEnt. However, we also note that a Kruskal-Wallis rank sum test showed a non-significant difference between the participants' estimates for these two probabilities ($\chi^2(1) = 1.9$, $p = 0.16$).

In summary, the results indicate that people are capable of reasoning with verbal probabilistic information and exhibit common causal reasoning patterns such as 'explaining away.' However, we found that people update their beliefs similarly whether all elements of the algebra were known from the start or when some elements were introduced later. We also observed that simply learning about the existence of another cause can potentially influence the probability of the first cause; however, this effect was not statistically significant.

Experiment 2

In Experiment 1, participants were directly informed about both the priors of the causes and the likelihoods related to those causes. However, only the priors were elicited from the participants. In Experiment 2, in addition to eliciting the priors, we also elicited the likelihoods from the participants to ensure that they have accepted the likelihoods. In all other aspects, Experiment 2 is exactly the same as Experiment 1 and, as such, it serves as a replication of Experiment 1.

Participants & Design A total of 288 participants ($N_{\mathrm{FEMALE}} = 171$, $M_{\mathrm{AGE}} = 43.8$ years) were recruited from Prolific Academic (`www.prolific.co`). All participants were native English speakers residing in the UK, the US, or Canada with approval ratings of 95% or higher. They all gave informed consent and were paid £6.66 an hour rate for participating in the present study, which took on average 9.2 min to complete.

Participants were randomly assigned to one of the three conditions: *Full* condition ($N = 94$), *Seq* condition ($N = 100$), and *Seq aware* condition ($N = 94$).

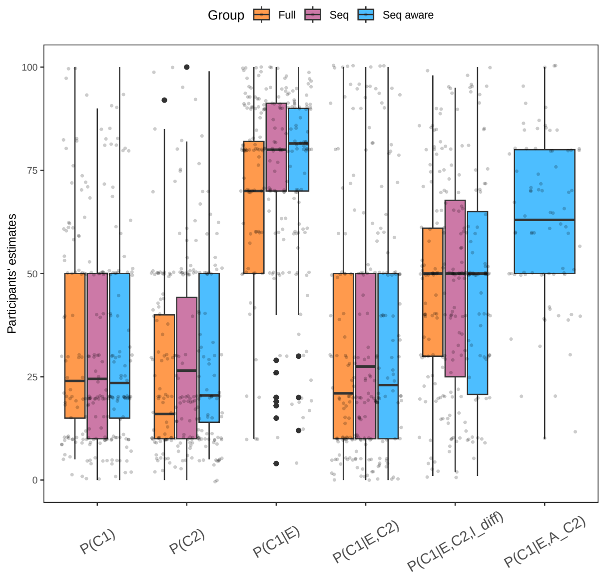

Figure 3: Participants' responses to questions in Experiment 1.

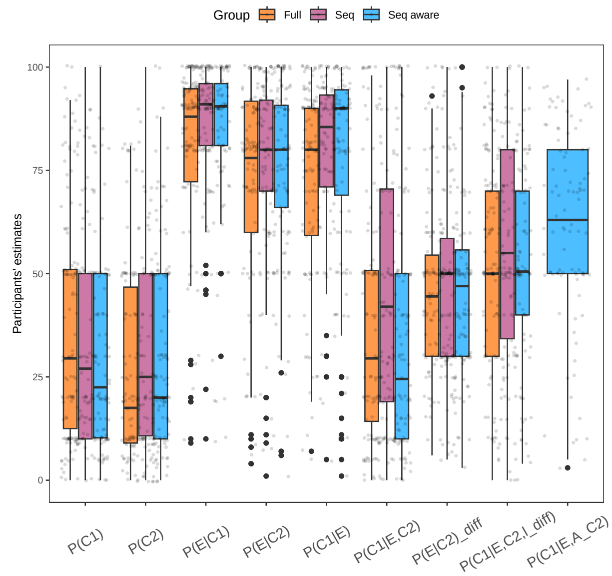

Figure 4: Participants' responses to questions in Experiment 2.

Materials & Procedure The materials and procedure were exactly the same as in Experiment 1, with the exception that in Experiment 2, participants were also asked the following two questions related to the likelihoods ($P(E|C1)$ and $P(E|C2)$):

Q3. How likely is it that the lawn is wet if it **rains** overnight?

Q4. How likely is it that the lawn is wet if **the sprinkler turns on** overnight?

Also, after learning about the new likelihood for the sprinkler, participants were again asked the same **Q4** from above. The estimates for this question were labeled as $P(E|C2)$_diff in Figure 4.

Results & Discussion From Figure 4, we find that, as in Experiment 1, participants have accepted the low prior and their estimates exhibit an 'explaining away' pattern. However, we also note that participants provided high likelihood estimates in line with the verbal information communicated to them ($P(E|C1)$ and $P(E|C2)$) and have revised their likelihoods for $P(E|C2)$_diff when they were informed that these had changed. This suggests that participants have accepted the likelihoods, including $P(E|C2)$_diff, indicating that they were attentive throughout the experiment.

We built an LMM with the same fixed- and random-effect structure as in Experiment 1. The LMM indicated a significant effect of Question ($t(948) = 6.28$, $p < .001$), a non-significant effect of Condition ($t(237) = 1.15$, $p = .25$), and no interaction between the two fixed effects ($t(948) = -0.81$, $p = .42$). These results resemble those from Experiment 1 and suggest that people reason similarly whether all elements of the algebra were known from the start or when some elements were introduced later. Like in Experiment 1, we also observe that in the *Seq aware* condition, participants' estimates for $P(C1|E, \text{A\_}C2)$ were lower than their estimates for $P(C1|E)$. This suggests that merely learning about the existence of another variable (another cause) can affect people's beliefs, even if they do not know whether that variable is instantiated (whether the cause occurred). However, unlike in Experiment 1, in this experiment a Kruskal-Wallis rank sum test showed a significant difference between the participants' estimates for these two probabilities ($\chi^2(1) = 9.1$, $p = 0.003$).

In summary, Experiment 2 largely replicated the findings from Experiment 1. Furthermore, we found that participants accepted the likelihoods and remained attentive throughout the experiment. We also observed a significant effect of merely learning about the existence of another cause (C2) on the probability of the first cause (C1), even when it is unknown whether C2 occurred.

Conclusion

In this paper, we explored the Maximum Entropy (MaxEnt) solution to the problem of awareness growth in Bayesian epistemology and decision theory. While MaxEnt is theoretically well-grounded in objective Bayesian principles (specifically, the principle of minimum information and arguments from epistemic cautiousness), the approach seems computationally intensive, making it seemingly unlikely that real human reasoners would adhere to it. However, our experimental results suggest that, under reasonable conditions, participants' responses were largely consistent with the MaxEnt solution. Notably, we observed the 'explaining away' effect and found that assessments of posterior probability that were independent of whether all possible causes were known from the beginning.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

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

Jaynes, E. T. (1968). Prior probabilities. *IEEE Transactions on Systems Science and Cybernetics*, *4*(3), 227–241.

Karni, E., & Vierø, M.-L. (2015). Probabilistic sophistication and reverse bayesianism. *Journal of Risk and Uncertainty*, *50*(3), 189–208.

Landes, J., & Williamson, J. (2013). Objective Bayesianism and the maximum entropy principle. *Entropy*, *15*(9), 3528–3591.

Meder, B., & Mayrhofer, R. (2017). Diagnostic causal reasoning with verbal information. *Cognitive Psychology*, *96*, 54–84.

Morris, M. W., & Larrick, R. P. (1995). When one cause casts doubt on another: A normative analysis of discounting in causal attribution. *Psychological Review*, *102*(2), 331.

Rosenkrantz, R. D. (1992). The justification of induction. *Philosophy of Science*, *59*(4), 527–539.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379–423.

Steele, K., & Stefánsson, H. O. (2021). Belief revision for growing awareness. *Mind*, *130*(520), 1207–1232.

Tešić, M., & Hahn, U. (2019). Sequential diagnostic reasoning with independent causes. In *Proceedings of the 41th annual conference of the cognitive science society* (pp. 2947–2953).

Tešić, M., Liefgreen, A., & Lagnado, D. (2020). The propensity interpretation of probability and diagnostic split in explaining away. *Cognitive psychology*, *121*, 101293.

Wellman, M. P., & Henrion, M. (1993). Explaining'explaining away'. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*(3), 287–292.

Williams, P. M. (1980). Bayesian conditionalisation and the principle of minimum information. *The British Journal for the Philosophy of Science*, *31*(2), 131–144.

Williamson, J. (2002). Maximising entropy efficiently. *Electronic Transactions in Artificial Intelligence Journal*, *6*.

Williamson, J. (2003). Bayesianism and language change. *Journal of Logic, Language and Information*, *12*(1), 53–97.

Williamson, J. (2007). Motivating objective Bayesianism: From empirical constraints to objective probabilities. In W. Harper & G. Wheeler (Eds.), *Probability and Inference: Essays in Honour of Henry E. Kyburg Jr* (p. 151-179). College Publications.