The Structure of Epistemic Probabilities

Abstract The epistemic probability of A given B is the degree to which B evidentially supports A, or makes A plausible. This paper is a first step in answering the question of what determines the values of epistemic probabilities. I break this question into two parts: the *structural* question and the *substantive* question. Just as an object's weight is determined by its mass and gravitational acceleration, some probabilities are determined by other, more basic ones. The structural question asks what probabilities are *not* determined in this way—these are the *basic* probabilities which determine values for all other probabilities. The substantive question asks how the values of these basic probabilities are determined.

I defend an answer to the structural question on which basic probabilities are the probabilities of atomic propositions conditional on potential direct explanations. I defend this against the view, implicit in orthodox mathematical treatments of probability, that basic probabilities are the unconditional probabilities of complete worlds. I then apply my answer to the structural question to clear up common confusions in expositions of Bayesianism and shed light on the "problem of the priors."

Keywords Bayesian Epistemology; Bayesian Networks; Explanation; Probability; Inference to the Best Explanation

1. Introduction

Were the dinosaurs killed by an asteroid? I don't know—and neither do you. How confident ought we to be that this proposition is true?

A plausible answer is that our confidence that the dinosaurs were killed by an asteroid ought to be equal to the *probability of that proposition given our evidence*. This raises two further questions: what is our evidence, and how is the probability of a proposition given some evidence determined? This paper is a first step in the (very large) project of answering the second of these questions.

The relevant sense of probability here is *epistemic* probability. The epistemic probability of A given B—notated P(A|B)—is a relation between the propositions B and A. It is the degree to which B supports A, or makes A plausible. Entailment is a limiting case of this relationship; if B entails A, then P(A|B) = 1. It constrains rational degrees of belief, in that, if P(A|B) = *n*, then someone with B as their evidence ought to be confident in A to degree *n*.[1]

Keynes (1921), Jeffreys (1939), Cox (1946), Carnap (1950), Williamson (2000: ch. 10), Swinburne (2001), Jaynes (2003), Hawthorne (2005), and Maher (2006) offer similar explications of probability.[2] There is a great deal more that could be said about the nature of epistemic probability. Most of the above authors claim that epistemic probability relations are necessary and knowable *a priori*. I am sympathetic to these claims, but the approach to the

[1] If some probabilistic support relations are imprecise or non-unique, then we would need to amend this bridge principle so that degrees of support constrain but do not always determine rational degrees of belief. Note also that some philosophers (e.g., Williamson 2010) take epistemic probabilities to be *identical* to rational degrees of belief, rather than a determinant of what degrees of belief are rational. I defend a degree-of-support interpretation of probability over a degree-of-rational-belief interpretation in my manuscript, "Epistemic Probabilities are Degrees of Support, not Degrees of (Rational) Belief." Philosophers skeptical of this degree-of-support interpretation can see this paper as provisionally working out the best way to develop it. (In addition, much of my project could be recast in terms of degrees of belief, for philosophers so inclined: see note 5 for further discussion.)

[2] These authors use various terms to describe their conceptions of probability, including 'logical probability,' 'inductive probability,' 'evidential probability,' and 'degree of support.' For the most part any differences between these conceptions of probability are not important for my purposes here.

structure of epistemic probabilities I go on to defend could also be accepted by philosophers who conceive of probabilistic support relations in an externalist or subjectivist manner.[3]

Epistemic probabilities conform to the laws of the probability calculus. However, these laws do not suffice to determine the values of epistemic probabilities. We can break down the project of explaining how these values are determined into two parts, which I will call the *structural* project and the *substantive* project. The structural project asks what probabilities' values are determined by the values of other probabilities, and what probabilities' values are not determined by the values of other probabilities. The substantive project then asks how the values of the latter probabilities are determined. (For example, one traditional answer would be that they are determined by the Principle of Indifference.) In this paper I undertake the structural project, leaving the substantive project for another time.

The premise of the structural project is that just as an object's weight is determined by its mass and its gravitational acceleration, the values of some probabilities are determined by the values of other probabilities.[4] I will call probabilities the values of which are determined by the values of other probabilities *non-basic*. *Basic* probabilities, by contrast, are the elementary quantities out of which other probabilities are built; they are the 'atoms' of probability theory. Given values for basic probabilities, we can compute values for all non-basic probabilities.[5]

---

[3] Some of my arguments rely on assumptions about the objective values of certain probabilities—e.g., that the probability that the ball drawn out of an urn will be black, given that the urn has 1 black ball and 2 white balls, is 1/3. That there are objective probabilities in easy cases like these is compatible, though, with the subjectivist intuition that in other, harder cases there is no one unique degree to which one proposition supports another.

[4] Plausibly, this determination relation is metaphysical grounding, but one need not assume this to pursue the structural question. I briefly discuss the possibility of other kinds of non-causal explanatory priority relations in section 2.3.

[5] Although some of the arguments that I go on to make regarding the structural project turn on a conception of epistemic probabilities as degrees of support, we could recast the structural project in terms of degrees of belief. In particular, when requiring that an agent's credences be coherent, we could ask which credences (if any) should be assigned *directly*, and which ones should be conformed to these basic credences by the laws of probability. The approach I defend, applied to subjective probability, would make a (to be specified) subset of conditional credences at a time basic, and require an agent to conform her other credences at that time to those. Suitably reformulated,

So the structural project asks: what probabilities are basic? And the substantive project asks: how are the values of these basic probabilities determined? Although these questions are both metaphysical, they are interesting mainly because of their epistemological upshot. We want to be able to *figure out* how probable our evidence makes the hypothesis that the dinosaurs were killed by an asteroid. The structural and substantive projects aid us in this to the extent that they help us figure out the values of basic probabilities, and then compute the values of non-basic probabilities (like, I will argue, this one) as a function of those.

In the past philosophers who have addressed the question of how we might figure out the values of epistemic probabilities have mainly focused on the substantive project, jumping straight to arguing for or against substantive methods like the Principle of Indifference. But what probabilities should we (for example) assign equal probabilities to? We must answer the structural question before we can know how to apply the Principle of Indifference (or some other substantive method).

Some philosophers have suggested that the values of some epistemic probabilities can be directly perceived (e.g., Keynes 1921: ch. II.8). If this is so, it again raises the question: which ones? In sections 3.3 and 3.4, I argue that the values of basic probabilities are more epistemically accessible than the values of non-basic probabilities. This means that determining which probabilities are basic can help us more reliably figure out the values of probabilities we care about even in the absence of an answer to the substantive question.

I call my answer to the structural question *Explanationism*. Informally, Explanationism says that the basic probabilities are the probabilities of atomic propositions conditional on

---

arguments 2-5 in section 3 could provide reason for subjective Bayesians who interpret probabilities as degrees of belief to endorse Explanationism about the structure of degrees of belief, and argument 6 could provide reason for objective Bayesians who interpret probabilities as rational degrees of belief to endorse Explanationism about the structure of rational degrees of belief.

potential direct explanations of those propositions. In section 2, I explain Explanationism in more depth, contrasting it with the Orthodox view about the structure of probabilities. In section 3, I argue for Explanationism against Orthodoxy. In section 4, I explore some philosophical implications of Explanationism. Section 5 concludes with some questions for further research.

2. Rival Views on the Structure of Probabilities

Before us is an urn. We know that it was selected by coin flip from two urns, U_1 and U_2. U_1 contains 1 black ball and 2 white balls, and U_2 contains 2 black balls and 1 white ball. We propose to learn about the contents of the urn by sampling from it at random. Let B and W stand for the propositions that the ball we draw is black or white, respectively.

In this problem there are two variables: the contents of the urn, and what color ball we draw. For each value that a variable can take on (e.g., the color of the ball drawn taking on the value black), there is an associated proposition (e.g., the proposition that the ball drawn out of the urn is black). Hence, each variable has an associated partition, that is, set of mutually exclusive and jointly exhaustive possibilities: $\{U_1, U_2\}$, $\{B, W\}$. (For ease of exposition, I will often informally speak of the members of these partitions as the values of their associated variables.)

In this problem we have four atomic propositions: U_1, U_2, B, and W.[6] We also have various complex disjunctions and conjunctions of these propositions which we can consider. Of particular interest are the following complex propositions:

U_1&B

U_1&W

U_2&B

[6] By 'atomic proposition,' I mean a proposition that is not truth-functionally decomposable into other propositions.

U_2&W

These propositions are *state-descriptions*—conjunctions in which one member of each partition appears once. State-descriptions are maximally complete descriptions of the world of our problem. They answer all our questions, assign a value to all our variables. In general, if we have *n* partitions with *m* members each, then we have m^n possible state-descriptions.[7]

For any pair of propositions in our problem X and Y, we can consider P(X|Y). We can also consider "unconditional" probabilities like P(X), which is the probability of X conditional only on the background knowledge given in the statement of the problem. (For ease of exposition, I suppress this background in this and the next section, e.g., writing $P(U_1\&B)$ rather than $P(U_1\&B|K)$.) Our question, applied to this problem, is which of these probabilities are basic, and which are non-basic.

2.1 Non-starters

One answer is that all these probabilities are basic. The lack of attention to the structural question suggests that many philosophers tacitly assume this. A second answer is that all the *unconditional* probabilities are basic. On this view, $P(U_1)$ and P(B) are basic, but $P(U_1|B)$ and $P(B|U_1)$ are not. This view is suggested by Hedden's (2015b: 470) claim that the "unique rational prior probability function … represents the a priori plausibility of each proposition," and Williamson's (2000: 211) remark that evidential probability "measures something like the intrinsic plausibility of hypotheses prior to investigation." This second view is also implicit in Ramsey (1926) and Jeffrey (1983)'s subjective Bayesian theories, which define unconditional

[7] The term 'state-description' comes from Carnap (1950), but my definition is slightly different from Carnap's. First, the conjuncts of state-descriptions in my sense are propositions rather than sentences. Second, and setting aside the first difference, state-descriptions in Carnap's sense are the special case of state-descriptions in my sense where the partitions are all of the form {A, ~A}. This latter difference is purely formal: each Carnapian state-description will be materially equivalent to a state-description constructed from more fine-grained partitions, and vice-versa.

degrees of belief first, and then define conditional degrees of belief in terms of these.[8]

Accepting either of these views makes it difficult to give an account of how the values of basic probabilities are determined. Standard answers to the substantive question would lead to probabilistic incoherence if applied to all probabilities, or applied to all unconditional probabilities. For example, the Principle of Indifference tells us to assign equal probabilities to a set of possibilities when our information does not support one over another. But it is impossible to assign equal values to all probabilities, or all unconditional probabilities; doing so will always be probabilistically incoherent. (For example, in the problem at hand, suppose that $P(U_1) = P(U_2) = P(U_1\&B) = P(U_1\&W)$. Since $P(U_1) = P(U_1\&B) + P(U_1\&W)$, it follows that $P(U_1) = 2P(U_1)$, and so $P(U_1) = P(U_2) = 0$. But this is impossible, since $\{U_1, U_2\}$ is a partition, and so $P(U_1) + P(U_2) = 1$.) So the Principle of Indifference can never directly determine the values of all (unconditional) probabilities; if it determines the values of these probabilities at all, it must determine some indirectly, by determining the values of others. Or consider a substantive view on which simpler propositions have higher probabilities than more complex propositions. Presumably $U_1 v B$ is a more complex proposition than U_1. If this criterion of simplicity is applied unrestrictedly, it then implies that $P(U_1) > P(U_1 v B)$, which is impossible.

I discuss further how answers to the structural question combine with substantive principles for determining the values of basic probabilities in section 3.6.1. For now the important thing to note is that principles like the above were designed to be applied to *partitions* of propositions, like $\{U_1, U_2\}$ and $\{B, W\}$. What went wrong in the above examples is that the different propositions being assigned probabilities are not mutually exclusive. I will now

[8] Other answers that may initially appear appealing would not actually fix values for all probabilities. For example, an assignment of values to all unconditional probabilities of *atomic* propositions would not determine values for either conditional probabilities or unconditional probabilities of state-descriptions. Knowing $P(U_1)$, $P(U_2)$, $P(B)$, and $P(W)$ would not enable us to determine the values of, e.g., $P(U_1|B)$ or $P(U_2\&W)$.

consider two structural views on which basic probabilities are assigned across partitions, in a way that makes it easier to combine these views with an answer to the substantive question.

2.2 Orthodoxy

The first of these views focuses on the partition of *state-descriptions*: in this case, $\{U_1\&B, U_1\&W, U_2\&B, U_2\&W\}$. On this view, the basic probabilities are the unconditional probabilities of state-descriptions: $P(U_1\&B)$, $P(U_1\&W)$, $P(U_2\&B)$, and $P(U_2\&W)$. This answer to the structural question takes its cue from orthodox mathematical probability theory, in which the probabilities of state-descriptions are assigned first, and other probabilities are determined as a function of these. Because of this, I call this view *Orthodoxy*.

In Kolmogorov's (1933) axiomatization of probability, the set of different state-descriptions is the "sample space." The sample space is one of the three basic notions in Kolmogorov's axiomatization. The second notion is an "algebra" on this sample space, that is, a set of subsets of the sample space. We can understand this as a set of state-descriptions and disjunctions of state-descriptions. The third notion is a "probability function" from members of the algebra to the unit interval [0,1].

While Kolmogorov's axioms for this probability function do not themselves require that any particular members of the algebra get assigned numbers first, the most standard way to construct a function that obeys these axioms is to begin by assigning probabilities to each member of the sample space (i.e., each state-description) such that these probabilities sum to 1.[9] (One can think of each state-description as taking up a certain proportion of the total space of possibilities, which has measure 1.) Kolmogorov's axioms, together with the ratio definition of

[9] This assumes a finite number of state-descriptions. It is controversial whether the elements of the sample space need to sum to 1 if the sample space is infinite. For simplicity, I only discuss finite sample spaces in this paper.

conditional probability, then determine P(X|Y) for any pair of propositions in our algebra X and Y, because any such proposition is logically equivalent to a disjunction of state-descriptions.[10]

Since orthodox probability theory assigns unconditional probabilities directly to each state-description, it treats the unconditional probabilities of state-descriptions as basic. Perhaps for this reason, most philosophers who have given precise quantitative (as opposed to merely qualitative) solutions to the substantive problem, including Carnap (1950), Solomonoff (1964), and Williamson (2010), have assumed Orthodoxy in their solutions.[11]

2.3 Explanationism

A final answer to our question, and the one I will defend, is *Explanationism*.[12] According to Explanationism, basic probabilities are the probabilities of atomic propositions conditional on propositions directly explanatorily prior to them. Because $\{U_1, U_2\}$ is directly prior to {B, W}, and nothing is prior to $\{U_1, U_2\}$, we have here six basic probabilities—$P(U_1)$, $P(U_2)$, $P(B|U_1)$, $P(W|U_1)$, $P(B|U_2)$, and $P(W|U_2)$.

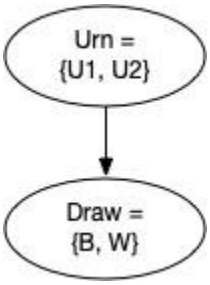


Figure 1

[10] Although I do not here address the question of which axiomatization of probability is best, in light of Explanationism's commitment to conditional probabilities as basic, it is plausible that the Explanationist should endorse non-standard axioms of probability, like those presented by Cox (1946), Jaynes (2003), and Maher (2004), which make conditional probability a primitive notion.

[11] For contemporary followers of Carnap and Solomonoff, see Tooley (2012) and Rathmanner and Hutter (2011), respectively. I discuss Williamson's system further in section 3.6.

[12] 'Explanationism' is sometimes used (e.g., in Lipton 2004) to describe the view that explanatory considerations are central to empirical inference. Explanationism in my sense implies, and gives specific content to, this more general claim. 'Explanationism' has also been recently used to describe non-Bayesian methods of updating credences (Douven 2013, Douven and Schupbach 2015) which I do not endorse (see Climenhaga 2017b).

Before offering a more formal statement of Explanationism, it will be helpful to go through this reasoning more slowly. According to Explanationism, the first step in determining the values of probabilities is to order the variables/partitions in our problem by their explanatory priority. In our current case, the Urn variable is explanatorily prior to the Draw variable—the contents of the urn influence what ball we draw out, but what we draw from the urn does not influence its (initial) composition. Fig. 1 formalizes these priority relations. It has two nodes, representing our two variables, with an arrow from the Urn node to the Draw node because the former is prior to the latter.

After ordering our variables, we take the basic probabilities to be those given to values of a variable by values of the variable(s) immediately prior to it. A basic probability, then, is the probability of a "downstream" proposition conditional on immediately "upstream" propositions. In the current case there are six such probabilities:

$P(U_1) = 1/2$

$P(U_2) = 1/2$

$P(B|U_1) = 1/3$

$P(W|U_1) = 2/3$

$P(B|U_2) = 2/3$

$P(W|U_2) = 1/3$

The Urn node is a root node; that is, there are no nodes pointing into it. As such, the (basic) probabilities of U_1 and U_2 are represented as unconditional. Really, they are conditional on the suppressed background knowledge given in the statement of the problem.

These six basic probabilities let us calculate any other probabilities we might be interested in. For example, Bayes' Theorem gives us:

$$P(U_1 \mid B) = \frac{P(U_1)P(B \mid U_1)}{P(U_1)P(B \mid U_1) + P(U_2)P(B \mid U_2)} = \frac{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)} = \frac{\left(\frac{1}{3}\right)}{1} = \frac{1}{3}$$

In this simple example, we had only two variables to order. Consider now two modifications of the above case. In the first modification, we make two draws with replacement from the urn. Then our diagram looks like Fig. 2. In the second modification, we make two draws from the urn but do *not* replace the ball after the first draw. Then our diagram looks like Fig. 3. Alternatively, we could represent the choice to replace or not replace the first draw as a separate variable, as in Fig. 4.

In these diagrams, we include an arrow from one variable to another if we think it possible that the value of the first variable somehow *influences* the value of the second. If we are sampling with replacement, the outcome of the first draw does not influence the outcome of the second. If we are sampling without replacement, it does; drawing black the first time lowers the probability that we draw it the second time. In Fig. 4, the lack of an arrow from the Draw 1 variable to the Replacement variable represents the assumption that the outcome of the first draw will not influence our choice of whether to put the ball back in the urn.

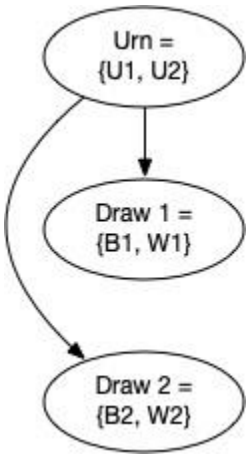


Figure 2

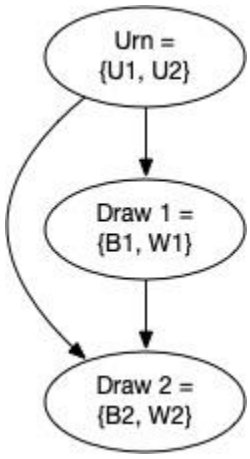


Figure 3

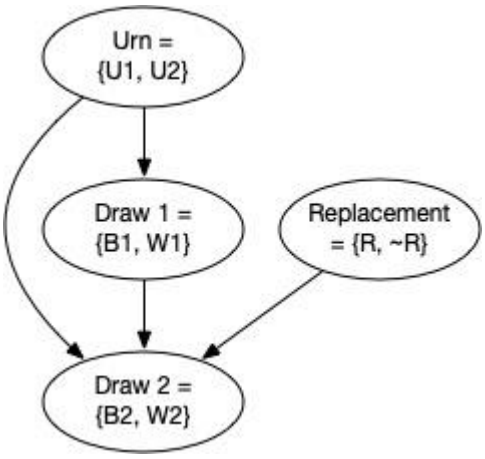


Figure 4

Figures 1-4 are *directed acyclic graphs* (DAGs). A DAG is a directed graph with no

loops. It consists of a finite number of nodes, with arrows drawn from some nodes to other nodes such that the arrows never form a loop. We can interpret a DAG as giving us the *ordering* of the variables in our algebra which allows us to determine which probabilities are basic. To do this we employ the language of ancestors and descendants. We say that X is a *parent* of Y iff there is an arrow from X to Y, and an *ancestor* of Y iff it is a parent, parent of a parent, etc. (that is, there is a directed path from X to Y). If X is a parent/ancestor of Y, Y is a child/descendant of X.

The variables represented by a DAG are said to obey the *Markov condition* just in case a variable's parents screen it off from all non-descendants. For example, in Fig. 2 the Urn variable screens off Draw 1 from Draw 2—if we know what urn we are sampling from, learning the outcome of the second draw provides us no information about the outcome of the first draw, and vice-versa. Formally:

> **Markov Condition**
> A DAG obeys the *Markov condition* iff for all atomic X, X is conditionally independent, given its parents, from any (conjunction of) non-descendants of X.

Where pa(X) represents the parents of X, the Markov condition says that P(X|pa(X)) = P(X|pa(X)&Z), unless (some conjunct of) Z is a descendant of X.

The Markov condition is intuitively plausible when we think of a DAG as representing causal structure (and there are no relevant causal variables omitted from the DAG). If Y already tells us everything relevant to predicting X in advance, then we can only get more information about whether or not X is true by learning about its effects. For example, if we know that the only thing that directly causally influences one's getting lung cancer is the amount of tar in one's lungs, then it is plausible that the amount of tar in one's lungs screens off getting cancer from one's smoking habits—that is, P(cancer | tar) = P(cancer | tar&smoking).[13]

[13] The Markov condition is not completely uncontroversial in either epistemic or causal contexts, but I lack the space

A directed network of partitions that obeys the Markov condition is called a *Bayesian network*. On the Explanationist answer to the structural problem, we start off by ordering the partitions we are interested in in a Bayesian network. The basic probabilities will be those given to an atomic proposition by assignments of values to all its parents. All other probabilities in the network can be determined as a function of those (Pearl 2000: 14-16). For example, in Fig. 4, $P(B_2 \mid R\&B_1\&U_1)$ is basic, but $P(B_2 \mid B_1\&U_1)$ is not, because the latter probability is not conditioned on all the parents of B_2. Rather, $P(B_2 \mid B_1\&U_1)$ must be calculated as a weighted average of $P(B_2 \mid R\&B_1\&U_1)$ and $P(B_2 \mid \sim R\&B_1\&U_1)$, weighted by $P(R \mid B_1\&U_1)$ and $P(\sim R \mid B_1\&U_1)$.[14]

Here, then, is a first statement of Explanationism:

Explanationism 1.0
P(X|Y) is basic iff X is atomic, and Y is a conjunction of values for all parents of X in a Bayesian network that includes all variables immediately explanatorily prior to X, and correctly relates all the variables it includes.

Later on, in section 3.6, I will relativize this statement to higher-order hypotheses about Bayesian networks, to allow for uncertainty about the correct Bayesian network. For the moment, I set these complications aside, as there is plenty to unpack here already.

First, in saying that a Bayesian network correctly relates the variables it includes, I mean that it includes an arrow from V_1 to V_2 iff V_1 is immediately explanatorily prior to V_2.[15] By 'immediately explanatorily prior' (or 'directly explanatorily prior'), I mean that V_1 is

---

to address objections to it here. For discussion, see Pearl 1988: ch. 3 and Hitchcock 2012.

[14] More generally, we can break down the probability of a conjunction A&B where A is a parent of B into basic probabilities using the Conjunction Rule P(A&B) = P(A)P(B|A), break down the probability of a proposition given its descendant into basic probabilities using Bayes' Theorem, and break down the probability of a proposition given non-descendants into basic probabilities using the Theorem of Total Probability (conditioning on all the ways the proposition's parents could be). We will see further examples of these operations in sections 3 and 4.

[15] Similarly, Pearl (1988: 123) and Bovens and Hartmann (2003: 68) both suggest that Bayesian networks should be constructed so that the parents of a variable are its "direct influences."

explanatorily prior to V_2, and there is no other variable that mediates the explanatory relation between them.

When is V_1 explanatorily prior to V_2? Causal priority, as in the above urn examples, is one kind of explanatory priority, and the most common kind to which Bayesian networks have been applied. Schaffer (2016) also uses Bayesian networks to formalize metaphysical grounding. Plausibly, causal and metaphysical priority are the only two kinds of *direct* explanatory priority, so that V_1 is directly explanatorily prior to V_2 iff it is either directly causally prior to V_2 or directly metaphysically prior to V_2. But if V_1 is metaphysically prior to V_2, which is causally prior to V_3, then even if V_1 is neither metaphysically or causally prior to V_3, it is still explanatorily prior to it: indirect relations of explanatory priority need not be solely metaphysical or solely causal, but can be combinations of both (cf. Lange 2018: 1345).

Whether causal and metaphysical priority are really the only two kinds of direct explanatory priority is disputable. Mathematical priority might be distinct from metaphysical grounding. Huemer (2009: 352-53) discusses temporal, part-whole, in-virtue-of, and supervenience priority. Henderson et al. (2010: 180) speak of more specific theories as being "constructed" out of more general theories, giving examples in which the probability of the specific theory conditional on the general theory is apparently treated as basic by scientists. I leave the question of whether these are really (distinct) kinds of explanatory priority, and whether there are other kinds, as an area for further research.

Although I am aware of no philosopher who has explicitly formulated Explanationism in the above manner, the view has several important predecessors. It sides with defenders of inference to the best explanation (e.g., Thagard 1978, Lipton 2004, Henderson 2014, Hedden 2015a: sec. 4, Climenhaga 2017a) in holding that explanatory relations are central to uncertain

inference. Mathematically, it is indebted especially to Pearl's (1988, 2000) groundbreaking work on Bayesian networks.[16] For the most part, philosophers who have applied Bayesian networks to epistemology (e.g., Bovens and Hartmann 2003) do not discuss the foundational issues explored in this essay; the same goes for statisticians such as Gelman et al. (2014: ch. 5) who employ hierarchical Bayesian models (a special case of Bayesian networks)[17] in statistics. Explanationism can justify these applications of Bayesian networks, as well as (I argue in section 3.4) many other ordinary applications of probability that do not appeal to graphical modeling. The philosophers who have come closest to endorsing Explanationism are Henderson et al. (2010), who defend hierarchical Bayesian modeling in the philosophy of science, and Huemer (2009) and Weisberg (2009: 140-41), who defend the application of the Principle of Indifference to explanatorily basic partitions. Both of these are special cases of Explanationism.[18]

In the next section I give six arguments for Explanationism.[19] The first is that it fits more naturally with the characteristics of epistemic probability than does Orthodoxy. The second is that in some cases, conditional probabilities may be well-defined while the associated state-description probabilities are not, making the latter unavailable as a ground for the former. The third and fourth are that the probabilities that Explanationism identifies as basic are precisely

[16] In the past two decades, Bayesian networks have been applied to two main areas: artificial intelligence/machine learning (e.g., Pearl 1988, Russell and Norvig 2009: ch. 14, Korb and Nicholson 2011) and causality (e.g., Pearl 2000, Spirtes et al. 2000, Hitchcock 2012). Explanationism is more in keeping with the epistemic interpretation of probability usually adopted in the artificial intelligence literature, but it agrees with contributors to the causality literature that causal relationships—and explanatory relationships more generally—cannot be reduced to probabilistic relationships.

[17] In particular, a hierarchical Bayesian model is a Bayesian network in which the variables are totally ordered from V_1 to V_n, and the only arrows are from V_1 into V_2, V_2 into V_3, and so on.

[18] Jon Williamson, who I mentioned earlier (section 2.2) as a proponent of Orthodoxy, is a notable recent advocate for the use of Bayesian networks in epistemology. However, unlike me, Williamson endorses the use of Bayesian networks for purely pragmatic, computational purposes. For Williamson, the probability distribution is determined first, and the Bayesian network is a way to represent it; whereas for me, the Bayesian network comes first, and helps determine the probability distribution. See the discussion in section 3.6.1.

[19] See Climenhaga 2017b for some related arguments against taking state-descriptions (there called "world-states") to be the primary objects of inference.

those which we find ourselves able to more easily judge the value of, both in urn-sampling thought experiments and in more realistic applications. The fifth is that Explanationism can be more easily extended to calculate probabilities conditioned on interventions rather than observations. The final, and most important, argument is that plausible substantive methods deliver incorrect results when combined with Orthodoxy, but not when combined with Explanationism.

3. Six Arguments for Explanationism

3.1 Explanationism fits better with the nature of epistemic probability.

Orthodoxy about a kind of probability may look initially appealing partly because it offers to *reduce* conditional probabilities to unconditional probabilities. Epistemic probability, though, is a *relation* between propositions: the degree to which one proposition makes another plausible. This means that all epistemic probabilities are conditional, because only conditional probabilities have two relata. The "unconditional" epistemic probability of a state description is really the state-description's probability conditional only on *a priori* truths (Hájek 2003: 315)—the degree to which *a priori* truths make that state-description plausible.

On the epistemic interpretation of probability, then, Orthodoxy becomes less motivated: it becomes unclear why we should think that the probabilities that Orthodoxy identifies as basic are basic. If these conditional probabilities can be basic, why must other conditional probabilities be defined in terms of them? What is special about the Orthodox basic probabilities?

By contrast, Explanationism can give a principled explanation of why, say, $P(B|U_1)$ is basic—it is basic because U_1 directly *gives* a probability to B in virtue of the Urn variable being the sole variable influencing B's truth. U_1 (which says that the urn contains 1 black and 2 white balls) directly makes B plausible to degree 1/3 because of the role it plays in explaining the truth

or falsity of B. This fits well with a conception of epistemic probability as measuring a quantity (namely, plausibility) that U_1 *confers* on B.

3.2 Conditional probabilities of atomic propositions may be well-defined when associated unconditional state-description probabilities are not.

It is controversial whether all probabilities are well-defined. Hájek (2003: 303-05, 309-10) suggests that there may not be well-defined physical or subjective probabilities for some of a person's future free actions. Similarly, one might think that the unconditional epistemic probabilities of some future free actions are undefined. Consider again the urn example represented in Fig. 4, in which we include a variable for whether we sample with replacement. If the choice whether or not to replace is a free choice, it might be that P(R) is undefined.

It is obvious that $P(B_2 \mid R\&B_1\&U_1) = 1/3$—for U_1 says that we are drawing from the urn with 1 black ball and 2 white balls, and R says that we replace our first draw, so that it does not impact the composition of the urn. However, Orthodoxy would have it that this value is determined by the equation

$$P\left(B_2 \,\middle|\, R\&B_1\&U_1\right) = \frac{P(B_2\&R\&B_1\&U_1)}{P(R\&B_1\&U_1)} = \frac{P(B_2\&R\&B_1\&U_1)}{P(B_2\&R\&B_1\&U_1) + P(W_2\&R\&B_1\&U_1)}$$

But if P(R) is undefined, then so presumably are these state-description probabilities. So according to Orthodoxy, $P(B_2 \mid B_1\&R\&U_1)$ should be undefined too. By contrast, Explanationism identifies $P(B_2 \mid B_1\&R\&U_1)$ as basic, and so can easily let it be well-defined.

It is not obvious that some epistemic probabilities are undefined. But it is also not obvious that all epistemic probabilities are well-defined. Orthodoxy would make obviously well-defined conditional probabilities undefined if the unconditional probabilities of some state-descriptions turn out to be undefined. By contrast, Explanationism can allow that these obviously well-defined conditional probabilities are well-defined, even if the unconditional probabilities of

the associated state-descriptions turn out to be undefined. Inasmuch as we should leave open the possibility that some epistemic probabilities are undefined, we should prefer a structural theory that does not let potentially undefined probabilities extend their influence too widely.

3.3 The probabilities Explanationism identifies as basic can be more directly perceived than those Orthodoxy identifies as basic.

In the above urn cases, you can immediately tell that $P(B|U_1) = 1/3$ as soon as you understand what B and U_1 say. On the Orthodox treatment of probability, however, $P(B|U_1)$ is not basic, but is instead defined as

$$P(B \mid U_1) = \frac{P(U_1\&B)}{P(U_1)} = \frac{P(U_1\&B)}{P(U_1\&B) + P(U_1\&W)}$$

However, whereas you can immediately see that $P(B|U_1) = 1/3$, the unconditional probabilities of the two state-descriptions are not immediately obvious. Were you called upon to determine $P(U_1\&B)$, the way to proceed would be to reduce it to $P(U_1)P(B|U_1) = (1/2)(1/3) = 1/6$. But this way of determining its value appeals to $P(B|U_1) = 1/3$, and so cannot be the means by which we gain knowledge of that equality (cf. Pearl 1988: 31 and 2000: 4).

It does not follow from the fact that we can more immediately *see* the value of $P(B|U_1)$ than $P(U_1\&B)$ that the former is more *metaphysically* basic than the other. In many contexts, less metaphysically basic properties are more epistemically accessible. For example, we can more easily determine the weight of an object than its mass, even though the weight depends on the mass. In the *a priori* case, many of us can readily tell that, if we have four cards with “Beer” or “not-Beer” on one side and “Over 21” or “Under 21” on the other, then in order to make sure that no card violates the rule “If you are drinking beer you are over 21,” we must turn over any card with Beer face up and any card with Under 21 face up. But we cannot as readily tell that, if we have four cards with “P” or “not-P” on one side and “Q” or “not-Q” on the other, then, in order

to make sure that no card violates the rule "If P then Q," we must turn over any card with P face up *and* any card with not-Q face up.

Nevertheless, the metaphysical basicality of $P(B|U_1)$ is the most plausible explanation of its epistemic directness in the present case. We are able to discover empirical properties without any knowledge of their metaphysical grounds because we can examine the way they affect the environment. For example, we can determine an object's weight by placing it on a scale. But this is not how we determine the value of $P(B|U_1)$: we do not measure the effects of this value on some external stimulus. Similarly, we can sometimes more readily perceive less basic *a priori* facts because of our implicit knowledge of the more basic facts which make them true. But our knowledge that $P(B|U_1) = 1/3$ does not appear to be based on any implicit grasp of the values of $P(U_1\&B)$ and $P(U_1\&W)$, in the way that our knowledge of how to react in the beer-rule example is based on implicit knowledge of how conditionals work.

Instead, in this case we appear to judge that $P(B|U_1) = 1/3$ simply because we understand what B says and we understand what U_1 says. If you were to ask a layperson, unfamiliar with Kolmogorov's axiomatization, *why* $P(B|U_1) = 1/3$, the most likely answer would appeal to the content of B and U_1, and their explanatory relation: "Well, U_1 says that 1 out of the 3 balls is black, and B says that we draw a black ball." (And perhaps: "And we've got no reason to think we're more likely to draw one ball than the other.") So in the present case, it is plausible that we perceive the value of $P(B|U_1)$ either directly or in virtue of grasping some substantive rule like the Principle of Indifference.

3.4 Explanationism better models actual probabilistic reasoning.

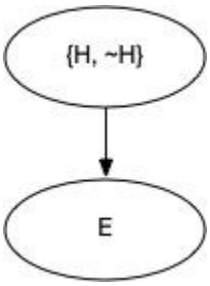
In the last sub-section I observed that the propositions Explanationism identifies as metaphysically basic in our urn example are exactly the ones that are most epistemically direct,

and argued that their being metaphysically basic is a plausible explanation of their being epistemically direct. You might worry that the urn example is cherry-picked, and that in other examples we would be more easily able to see the values of state-description probabilities. However, when we turn to real-life applications of Bayesian reasoning, we find that—despite orthodox mathematical probability theory's commitment to Orthodoxy—philosophers and scientists reason more in accord with Explanationism than Orthodoxy.

For example, consider Bayes' Theorem,

$$P(H \mid E) = \frac{P(H)P(E \mid H)}{P(H)P(E \mid H) + P(\sim H)P(E \mid \sim H)}$$

Expositions of Bayes' Theorem frequently advocate its use in cases where H is a "hypothesis" or "theory" and E is some "empirical data" "predicted" by H (see, e.g., Howson & Urbach 2006: 20-22, Joyce 2008: sec. 1, Weisberg 2015: sec. 1.2.2). These terms connote H's being explanatorily prior to E, as in Fig. 5. If Fig. 5 is our entire network, then according to Explanationism, the basic probabilities in the network are exactly the ones in Bayes' Theorem above.[20]



[20] The concurrence between Explanationism and applications of Bayes' Theorem is even clearer in older Bayesian terminology. Whereas today philosophers and statisticians follow R.A. Fisher in speaking of posterior probabilities and likelihoods, older writers (e.g., Venn 1876: sec. VI.9) referred to these as "inverse probabilities" and "direct probabilities," respectively. (These terms have occasionally survived, as in Joyce 2008: sec. 1.) The term "inverse probability" embodied the idea that in employing Bayes' Theorem we are moving "backwards" from effects to causes (Fienberg 2006: 5), and the term "direct probability" connotes a probability the value of which we are able to directly see or determine.

Figure 5

When we turn to examples writers use to illustrate Bayes' Theorem, they are invariably ones in which the hypothesis H is explanatorily prior to the evidence E. Salmon (1990: 178) illustrates Bayes' Theorem with an example in which H is the hypothesis that a particular can opener was produced by a machine with a given propensity for producing defective can openers, and E is the (explanatorily downstream) proposition that this can opener is defective. All four examples (drawing balls from an urn, finding a spider in a batch of bananas, hearing a witness report the color of a taxi, and getting a positive result on a medical test) in the "Bayes' Rule" chapter from Ian Hacking's introductory textbook (2001: ch. 7) likewise conform to this pattern.

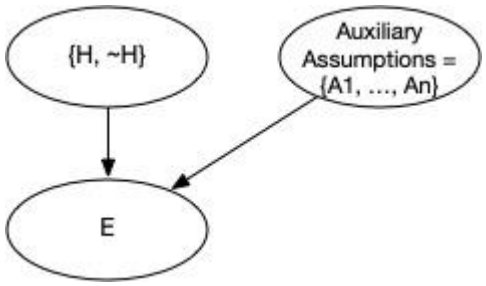


Figure 6

Again, consider the standard Bayesian treatment of Duhem's problem that most scientific hypotheses only make definitive predictions when combined with auxiliary assumptions. If H is our hypothesis and E is our empirical data, as before, this amounts to the problem of determining P(E|H) and P(E|~H) when applying Bayes' Theorem. The standard Bayesian resolution is to make explicit different possible auxiliary assumptions $\{A_1, \ldots, A_n\}$ and incorporate them into Bayes' Theorem as follows (Howson & Urbach 2006: 103-14):

$$P(H \mid E) = \frac{\sum_i P(H\&A_i)P(E \mid H\&A_i)}{\sum_i \left[P(H\&A_i)P(E \mid H\&A_i) + P(\sim H\&A_i)P(E \mid \sim H\&A_i)\right]}$$

If H and the A_i are independent (relative to any implicit background knowledge), then $P(H\&A_i)$

= P(H)P(A_i), and we have:

$$P(H \mid E) = \frac{\sum_i P(H)P(A_i)P(E \mid H\&A_i)}{\sum_i [P(H)P(A_i)P(E \mid H\&A_i) + P(\sim H)P(A_i)P(E \mid \sim H\&A_i)]}$$

In this context, the A_i are understood to be additional assumptions about, e.g., experimental set-up, the accuracy of our measurements, and any background theory relevant to making predictions about the outcome of our experiment. This suggests (if H and the A_i are independent) the network in Fig. 6. According to Explanationism, in this network, the terms on the right-hand side of the above equation are all basic.[21]

The above examples furnish us with another argument for Explanationism: many applications of Bayesian reasoning break down probabilities into precisely those quantities which Explanationism says are basic (or closer to being basic, inasmuch as the above networks approximate the actual evidential situation). As Pearl (1988: 78) notes,

> Human performance shows the opposite pattern of complexity [from Orthodoxy]: probabilistic judgments on a small number of propositions … are issued swiftly and reliably, while judging the likelihood of a conjunction of propositions entails much difficulty and hesitancy. This suggests that the elementary building blocks of human knowledge are not entries on a joint-distribution table. Rather, they are low-order marginal and conditional probabilities defined over small clusters of propositions.

Inasmuch as it is plausible that the more metaphysically basic probabilities will also be more epistemically direct, Explanationism explains the way people reason probabilistically in both philosophical and empirical contexts. By contrast, if Orthodoxy is true it is unclear why philosophers and scientists so often apply rules like Bayes' Theorem to break down complex probabilities into precisely those probabilities which Explanationism identifies as basic.

3.5 Explanationism combines more easily with a probabilistic calculus for causal interventions.

[21] See Bovens and Hartmann 2003: 107-11 for a more detailed application of Bayesian networks to Duhem's problem, including cases in which H and the A_i are not independent.

Another advantage of Explanationism presents itself when we consider adding the possibility of "direct causal interventions" to our problem. Explanationism, but not Orthodoxy, can easily tell us what probabilities to assign propositions conditional on such interventions.

In his influential 2000 book *Causality*, Pearl argues that we need to expand the syntax of the probability calculus to include probabilities of the form P(X|do(Y)), where do(Y) says that we directly make Y true, rather than observe that Y is true. Pearl (2000: 110) observes,

> By specifying a[n Orthodox] probability function $P(s)$ on the possible states of the world, we automatically specify how probabilities should change with every conceivable observation e, since $P(s)$ permits us to compute (by conditioning on e) the posterior probabilities $P(E|e)$ for every pair of events E and e. However, specifying $P(s)$ tells us nothing about how probabilities should change in response to an external action $do(A)$.

Constructing a Bayesian network relating X and Y allows us to determine P(X|do(Y)) by simply deleting any arrows going into Y, and calculating P(X|Y) in our mutilated network. Consider again the case of sampling twice from our urn with replacement in Fig. 2. Because we are sampling with replacement, the outcome of the first draw does not influence the outcome of the second—hence there is no arrow between them. However, learning that the first draw was black gives us information about the contents of the urn, and so is evidence that the second draw will also be black. By breaking down the value of $P(B_2|B_1)$ into basic probabilities, we can see that B_1 raises the probability of B_2 by raising the probability of U_2 from 1/2 to 2/3:

$$P(B_2 \mid B_1) = P(U_1 \mid B_1)P(B_2 \mid U_1) + P(U_2 \mid B_1)P(B_2 \mid U_2) =$$

$$= \frac{P(U_1)P(B_1 \mid U_1)}{P(U_1)P(B_1 \mid U_1) + P(U_2)P(B_1 \mid U_2)} P(B_2 \mid U_1)$$

$$+ \frac{P(U_2)P(B_1 \mid U_2)}{P(U_1)P(B_1 \mid U_1) + P(U_2)P(B_1 \mid U_2)} P(B_2 \mid U_2)$$

$$= \frac{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)}\left(\frac{1}{3}\right) + \frac{\left(\frac{1}{2}\right)\left(\frac{2}{3}\right)}{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)}\left(\frac{2}{3}\right) = \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{2}{3}\right)\left(\frac{2}{3}\right)$$

$$= \frac{1}{9} + \frac{4}{9} = \frac{5}{9}$$

The value of $P(B_2)$ can similarly be obtained by summing over U_1 and U_2 as above. In that calculation, the weights $P(U_1)$ and $P(U_2)$ are both equal to 1/2, and $P(B_2)$ is a simple average of $P(B_2|U_1) = 1/3$ and $P(B_2|U_2) = 2/3$, so that $P(B_2) = 1/2 < P(B_2|B_1) = 5/9$.

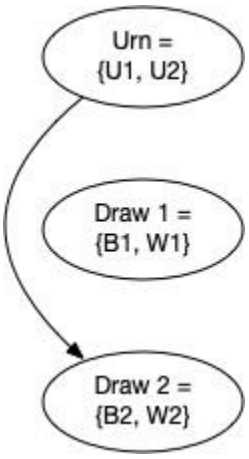


Figure 7

But now suppose that we directly "set" the value of the first draw to black, e.g., we hire someone to look inside the urn and intentionally pull out a black ball. If we then put the ball back in the urn, we learn nothing about the outcome of the second draw. Explanationism can deliver this result if we take $P(B_2|do(B_1))$ in our original network to be equal to $P(B_2|B_1)$ in the mutilated

network in Fig. 7. Here $P(B_2|B_1) = P(B_2)$, because B_1 neither raises the probability of B_2 directly nor *via* some intermediary, as in the original network.

By contrast, the Orthodox probability distribution over our four state-descriptions imposes no constraints on $P(B_2|do(B_1))$. The Orthodox probabilist could assign a new probability distribution over a new set of state-descriptions that includes actions like $do(B_1)$. But nothing in Orthodoxy requires that this distribution give probabilities like $P(B_2|do(B_1))$ the intuitively correct values. If it does give the correct values, this is simply a brute fact about those probability distributions. Inasmuch as Explanationism requires intuitively correct equalities that Orthodoxy must stipulate *ad hoc*, this gives us reason to prefer Explanationism.

3.6 Substantive methods for determining the values of basic probabilities get the wrong result if applied to Orthodox basic probabilities.

Recall that the task of determining the values of epistemic probabilities has two parts. We have been exploring the structural part, which asks which probabilities are basic and which are non-basic. The substantive part involves assigning values to the basic probabilities. Substantive methods will have different implications if applied to different (allegedly basic) probabilities. One of the most important reasons to settle the structural question is to guide the application of substantive methods in probabilistic reasoning. I will now argue that when we combine Orthodoxy and Explanationism with proposed substantive methods and they deliver different results, it is Orthodoxy that goes wrong. The proposed substantive methods I will consider are Maximum Entropy (a generalization of the Principle of Indifference) and assigning higher probabilities to simpler hypotheses.

I should stress that I am not committed to the correctness of these proposed substantive methods. My argument is conditional: *if* Maximum Entropy or simplicity are the correct criteria of basic probability, they get the right result only if combined with Explanationism. I argue,

moreover, that the basic problematic phenomenon I identify—the addition of explanatorily posterior variables affecting the probability of explanatorily prior variables—will take place with any method that assigns probabilities directly to state-descriptions.

My argument in this sub-section will be most effective with objectivists who think there are privileged probability assignments determined by some substantive method or other. However, I would note two points. First, applying Maximum Entropy to Explanationist basic probabilities rather than state-descriptions allows us to avoid many of the paradoxes the Principle of Indifference is often held to lead to (see Huemer 2009). As such, some objections to objectivism may be undermined by my argument in this sub-section. Second, many subjectivists about probability think of the impact of evidence as something individuals are free to determine based on how they weigh conflicting substantive criteria—such as symmetry and simplicity—against each other. So subjectivists who use these criteria to determine their own personal probabilities might still be moved by my arguments in this sub-section, provided that they share my intuitions about which applications of these criteria seem unsatisfying.

3.6.1 Example 1: Maximum Entropy

The Principle of Indifference says that we should assign equal probability to a space of alternatives if our knowledge does not favor any of these alternatives over any other. The Maximum Entropy principle (MaxEnt) generalizes this by telling us to assign probabilities that are as close to equal as is consistent with our knowledge (Williamson 2005: 80, 2010: 28-29).[22]

Orthodox probabilists like Williamson would have us apply MaxEnt to the set of all

[22] For the technical details of how to spell out "closeness to equality," see Williamson (2005: 79-84, 2010: 28-30 and 49-66) and Jaynes (2003: ch. 11). The equivocality, or uninformativeness, of a distribution is measured by its entropy, and we seek to maximize this entropy consistent with constraints provided by our knowledge—hence the name Maximum Entropy. In the text I rely on an intuitive understanding of closeness to equality; the results I give are those we would find by applying the mathematical methods in the above texts.

possible state-descriptions. On Williamson's version of objective Bayesianism, "the probabilities of the atomic states [i.e., state-descriptions] are basic: all other probabilities can be defined in terms of them" (2010: 27). According to Williamson, when one has no information favoring one state-description over another, one should assign equal probabilities to all of them. If one does have information favoring one state-description over another, one should assign probabilities as close to equal as is consistent with one's information.

I will now argue that applying MaxEnt to state-descriptions in this way leads to absurd results. Suppose I tell you that I have an urn in front of me that contains 1 black ball and 1 white ball. If I sample from the urn only once and we apply the Principle of Indifference to the partition $\{B_1, W_1\}$, we get the result that $P(B_1) = P(W_1) = 1/2$.

But now suppose that I tell you that I am going to sample from the urn twice, and that the outcome of the first draw will influence the outcome of the second one. In particular, if I draw the black ball the first time, I will set it aside, and so be ensured to draw the white ball the second time. If I draw the white ball the first time, I will set it aside, but also *add* a green ball to the urn.

Now we have two partitions: $\{B_1, W_1\}$, $\{B_2, W_2, G_2\}$. This gives us six state-descriptions: $\{B_1\&B_2, B_1\&W_2, B_1\&G_2, W_1\&B_2, W_1\&W_2, W_1\&G_2\}$. Your background knowledge that $B_1 \leftrightarrow W_2$ and $W_1 \leftrightarrow (B_2)v(G_2)$ allows you to eliminate the first, third, and fifth outcomes, leaving you with $\{B_1\&W_2, W_1\&B_2, W_1\&G_2\}$. If you apply the Principle of Indifference to those state-descriptions not excluded by your knowledge, they each get 1/3 probability. This implies that, before either draw has been made, $P(B_1) = 1/3$ and $P(W_1) = 2/3$. So without giving you any new knowledge about how I make the first draw and without telling you about any actual (as opposed to merely possible) effects of that draw, *I have made it more initially likely for you that the first draw is white.*

This is the intuitively wrong result. The outcome of the first draw is determined prior to the outcome of the second. B_1 and W_1 should both be assigned unconditional probability 1/2, and B_2 and G_2 should each be assigned equal probability conditional on B_1. This gives probabilities of 1/2, 1/4, and 1/4 to our state-descriptions.

Explanationism delivers the intuitively correct result in this case. First, we order our variables: Draw 1 is prior to Draw 2. Then we have the following basic probabilities:

$P(B_1) = P(W_1) = 1/2$,

$P(B_2|B_1) = P(G_2|B_1) = 0$,

$P(W_2|B_1) = 1$,

$P(B_2|W_1) = P(G_2|W_1) = 1/2$,

$P(W_2|W_1) = 0$.

The first and fourth lines follow from the application of MaxEnt to $\{B_1, W_1\}$ and to $\{B_2, W_2, G_2\}$ conditional on W_1. It follows that

$P(B_1\&W_2) = P(B_1)P(W_2|B_1) = (1/2)(1) = 1/2$,

$P(W_1\&B_2) = P(W_1)P(B_2|W_1) = (1/2)(1/2) = 1/4$,

$P(W_1\&G_2) = P(W_1)P(G_2|W_1) = (1/2)(1/2) = 1/4$.

Williamson (2005: 95-106; cf. 2010: 46-47) recognizes the above problem with applying MaxEnt to state-descriptions when we have causal information, discussing a similar example raised by Pearl (1988). His solution is to introduce causal constraints in addition to the quantitative constraints $P(W_2|B_1) = 1$ and $P([B_2 v G_2] \mid W_1]) = 1$ imposed by the above information. These causal constraints say that if our knowledge tells us that variables $\{V_1, \ldots V_i\}$ are causally ordered from 1 to *i*, then we begin by maximizing entropy over the propositions in V_1, giving us a probability distribution P_1. Next, we select the highest entropy probability

distribution over $V_1 \times V_2$ (i.e., the Cartesian product of V_1 and V_2) which is *consistent* with P_1, giving us P_2. Then, we select the highest entropy probability distribution over $V_1 \times V_2 \times V_3$ which is consistent with P_2, and so on.

In the above case, we begin by maximizing entropy over $\{B_1, W_1\}$, assigning probability 1/2 to each possibility, and then choose the probability distribution which maximizes entropy over $\{B_1\&W_2, W_1\&B_2, W_1\&G_2\}$ among those distributions consistent with $P(B_1) = P(W_1) = 1/2$. This gives us the same result as the Explanationist method.

There are two problems with using the above method to save Orthodoxy. First, it appears to be Orthodox in name only. From the perspective of Orthodoxy, Williamson's causal constraint looks *ad hoc* and unmotivated, wheeled in only to stave off counterexamples. If the probabilities of state-descriptions really are basic, then why does our background knowledge require us to conform them to probabilities first assigned to what are, from the perspective of Orthodoxy, disjunctions of state-descriptions (e.g., $[W_1\&B_2]v[W_1\&G_2]$)?

Indeed, this constraint gets the right results only because it parrots the Explanationist approach. For example, if $P(W_1) = 1/2$, then

$$P(W_1\&B_2) = P(W_1)P(B_2|W_1) = (1/2)P(B_2|W_1)$$

and

$$P(W_1\&G_2) = P(W_1)P(G_2|W_1) = (1/2)P(G_2|W_1).$$

We obtain the most equal distribution over $\{B_1\&W_2, W_1\&B_2, W_1\&G_2\}$ by setting $P(B_2|W_1) = P(G_2|W_1) = 1/2$, which gives us the {1/2, 1/4, 1/4} distribution over this partition. At each step we maximize entropy over the new set of variables consistent with the causal constraints *precisely by* maximizing entropy over the conditional probabilities $P(X|pa(X))$, i.e., the probabilities that Explanationism says are basic.

Second, and more seriously, Williamson's method only applies to the special case in which we *know* which variables causally influence which other variables. But consider a modification to the above case. As before, I tell you that an urn will be sampled from twice, and that in the one draw the possible outcomes are $\{B_1, W_1\}$ and in the other they are $\{B_2, W_2, G_2\}$. And as before, I tell you that $B_1 \leftrightarrow W_2$ and $W_1 \leftrightarrow (B_2)v(G_2)$. However, now I do not tell you which draw takes place first.

We can continue denoting the draw in which the only possibilities are black and white as Draw 1 and the other as Draw 2, but now these should be understood simply as labels, and not as denoting temporal information. So for all you know, the situation could be as represented in Fig. 8, or it could be as represented in Fig. 9. In this latter scenario, if I draw white in Draw 2 (which is now the first draw), I set the white and green balls aside, ensuring that I draw black on Draw 1; and if I draw black or green in Draw 2, set the black and green balls aside, ensuring that I draw white in Draw 1.

Williamson's method only applies when we know what the causal constraints are (2005: 99). As such, it will not preclude the ordinary application of MaxEnt to the state-descriptions $\{B_1\&W_2, W_1\&B_2, W_1\&G_2\}$. So we will again assign 1/3 probability to each of these, since that makes our distribution maximally equivocal. But inasmuch as you have no reason to think that either draw comes first, the Principle of Indifference should advise you to assign equal probability to both these possibilities, and then determine how likely each of these possibilities would make each of these state-descriptions. Letting N_1 stand for the hypothesis that the network in Fig. 8 is correct (i.e., Draw 1 comes first), and N_2 stand for the hypothesis that the network in Fig. 9 is correct (i.e., Draw 2 comes first), this gives us

$$P(B_1\&W_2) = P(N_1)P(B_1\&W_2 \mid N_1) + P(N_2)P(W_2\&B_1 \mid N_2)$$

$$= P(N_1)P(B_1 \mid N_1)P(W_2 \mid B_1\&N_1) + P(N_2)P(W_2 \mid N_2)P(B_1 \mid W_2\&N_2)$$

$$= (1/2)(1/2)(1) + (1/2)(1/3)(1) = 1/4 + 1/6 = 3/12 + 2/12 = 5/12$$

$$P(W_1\&B_2) = P(N_1)P(W_1\&B_2 \mid N_1) + P(N_2)P(B_2\&W_1 \mid N_2)$$

$$= P(N_1)P(W_1 \mid N_1)P(B_2 \mid W_1\&N_1) + P(N_2)P(B_2 \mid N_2)P(W_1 \mid B_2\&N_2)$$

$$= (1/2)(1/2)(1/2) + (1/2)(1/3)(1) = 1/8 + 1/6 = 3/24 + 4/24 = 7/24$$

$$P(W_1\&G_2) = P(N_1)P(W_1\&G_2 \mid N_1) + P(N_2)P(G_2\&W_1 \mid N_2)$$

$$= P(N_1)P(W_1 \mid N_1)P(G_2 \mid W_1\&N_1) + P(N_2)P(G_2 \mid N_2)P(W_1 \mid G_2\&N_2)$$

$$= (1/2)(1/2)(1/2) + (1/2)(1/3)(1) = 1/8 + 1/6 = 3/24 + 4/24 = 7/24$$

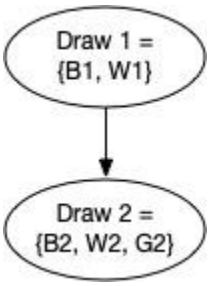


Figure 8

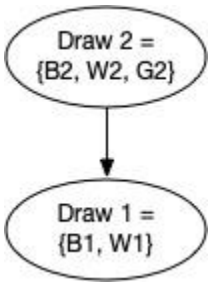


Figure 9

My initial statement of Explanationism in section 2.3, Explanationism 1.0, assumed that our background information entailed a unique explanatory network. We can amend Explanationism to recommend the above calculation by representing uncertainty about $\{N_1, N_2\}$ as higher-order uncertainty about what network is correct, and then taking basic probabilities to be relative to the network endorsed by N_i (cf. Huemer 2009: 363-65, Weisberg 2009: 141)—e.g., $P(B_1|N_1)$ and $P(W_2|B_1\&N_1)$ in the first equation above.

More formally:

Explanationism 2.0

> $P(X|Y\&N_i)$ is basic iff X is atomic, and Y is a conjunction of values for all parents of X in a Bayesian network that, according to N_i, includes all variables immediately explanatorily prior to X, and correctly relates all the variables it includes.[23]

Given values for the basic probabilities identified by Explanationism 2.0, we can determine $P(W|Z\&N_i)$ for any W and Z in N_i. From these we can then obtain $P(W|Z)$ by averaging over $P(W|Z\&N_j)$ for all possible networks N_j, weighted by the network-probabilities $P(N_j|Z)$, as above. These latter probabilities are a function of the prior probabilities of the networks, $P(N_j)$, and the degree to which these networks predict Z, $P(Z|N_j)$. Explanationism 2.0 can hold that the prior probability of a network $P(N_j)$ is basic by holding that this probability is implicitly relative to a higher-order network which contains a partition of the different possible first-order networks N_j. In the example above, this higher-order network would contain a single node, with the partition $\{N_1, N_2\}$.[24]

Although Williamson, like me, advocates the use of Bayesian networks in calculating probabilities, he cannot accommodate this higher-order uncertainty about networks into his framework. On my view Bayesian networks are logically prior to probability assignments, and basic probabilities are determined by means of them. But for Williamson, Bayesian networks play the purely pragmatic role of simplifying computations (2010: ch. 6), except in the special case in which a Bayesian network is uniquely determined by our causal information. If we do not know which of the Draw variables comes first, Williamson's method for constructing Bayesian networks (2005: 84-95) would lead to the network represented in Fig. 9, simply because Draw 2

[23] Explanationism 2.0 is my final statement of Explanationism for the purposes of this essay. In what follows I again mostly ignore network-relativity for simplicity's sake, but it is important to keep in mind for many applications, because network uncertainty is very common.

[24] Higher-order networks could get much more complicated—e.g., we could imagine a higher-order network that made which of Draw 1 and Draw 2 is first itself prior to the order of variables in some other problem. We could also complicate things by making the higher-order network uncertain rather than given in the statement of the problem. A question for further exploration is whether we need to eventually reach an *a priori* higher-order network relating all uncertain lower-order networks, in order to avoid an infinite regress in the determination of some probabilities.

has more variables than Draw 1 and so placing it prior to Draw 1 on the network, combined with successive applications of MaxEnt in the way that Explanationism recommends, gives us the same result as directly maximizing entropy over the state-descriptions $\{B_1\&W_2, W_1\&B_2, W_1\&G_2\}$. So except in the special case in which causal knowledge forces us to adopt a particular network, what Bayesian network to employ is determined by what will maximize entropy over the state-descriptions. By contrast, according to Explanationism what Bayesian network or networks to employ is determined by explanatory relations that are prior to the application of MaxEnt or any other substantive method for determining the values of basic probabilities.

3.6.2 Example 2: Simplicity

I have considered the application of MaxEnt to determining the values of basic probabilities, and argued that it gives us the wrong result if the basic probabilities are those posited by Orthodoxy, whereas it gives us the right result if the basic probabilities are those posited by Explanationism. The same phenomenon occurs if we employ other proposed criteria of basic probability, such as simplicity. For illustrative purposes, let us follow Hesse (1974: 234-36) and Swinburne (2001: 87) in taking one facet of simplicity to be quantitative parsimony, so that a theory is simpler to the extent that it posits fewer entities.[25]

Suppose that we know that either 1 male or 1 male and 1 female bird (of the same species) flew to an island off the coast of the Americas 2 generations ago. We further know that

[25] Hesse (1974: ch. 10) appears to endorse Orthodoxy, writing that "*ceteris paribus*, the universe is to be postulated to be as homogeneous as possible consistently with the data" (230). Bradley (forthcoming) defends a view similar to Hesse's. Swinburne (2001: ch. 4) sometimes speaks of the importance of simplicity in a way that suggests that he also thinks it attaches fundamentally to state-descriptions. For example, he writes: "We should postulate, on grounds of simplicity as most likely to be true, that theory of a narrow region *which makes our overall world theory the simplest for the total data*. That theory may not be the simplest theory of the narrow aspect of the world, considered on its own" (96, emphasis mine). Elsewhere, though, he clarifies his view in a way that makes clear that his view is closer to Explanationism than Orthodoxy: "the intrinsic probability of a world is a function of how simple are the highest-level hypotheses which it contains and how well they are able to explain all the other propositions which the world contains" (Swinburne 2011: 394).

each pair of male-female birds has 5 male and 5 female children in a generation. Then the total number of birds (in all generations) under the second hypothesis is 2 + 10 + 50 = 62. Since, on the first hypothesis, the bird has no mate with which to reproduce, the total number of birds given the first hypothesis is 1.

If we read quantitative parsimony as attaching to the total number of entities to which we are committed in our overall worldview, then the 2-bird hypothesis is much less simple than the 1-bird hypothesis—it posits 62 times as many birds! Intuitively, however, the 2-bird hypothesis is only slightly less simple than the 1-bird hypothesis, inasmuch as it posits only one more (comparatively) fundamental entities (birds). And as with the adding-a-green-ball-to-the-urn example above, comparing the simplicity of state-descriptions would lead to the implausible conclusion that learning that one more generation has gone by should lower the relative probability of the 2-bird hypothesis. It seems, then, that if we want to give preference to simpler hypotheses, we should compare the simplicities of atomic hypotheses on the same level of explanation, and not the simplicities of overall worldviews.

In this sub-section I have considered the application of substantive methods for determining the values of basic probabilities, and argued that they give us the wrong result if we adopt Orthodoxy, and the right result if we adopt Explanationism. The defender of Orthodoxy might object that the methods I have considered are not the correct ones, or have been misapplied. But the same general phenomenon of the addition of explanatorily posterior variables wrongly affecting the probability of explanatorily prior variables will take place with any method that assigns probabilities directly to state-descriptions, unless a safeguard is built into the method to avoid this, as in Williamson's version of MaxEnt. And such a safeguard will likely, as above, either fail to avoid all counterintuitive consequences, reveal an implicit

commitment to the order of explanation as prior to the assignment of probabilities, or both.

4. Why Explanationism Matters

In section 3 I gave six arguments for Explanationism over Orthodoxy. First, it is philosophically better motivated than Orthodoxy as a theory of basic *epistemic* probabilities. Second, it allows for conditional probabilities to be well-defined even when the state-description probabilities to which Orthodoxy would reduce them may not be well-defined. Third, we are more easily able to judge the values of the probabilities Explanationism identifies as basic than those Orthodoxy identifies as basic. Fourth, it better describes actual (good) scientific and empirical reasoning. Fifth, it can more easily be combined with Pearl's (2000) probabilistic do-calculus. Finally, it leads to more intuitive probability assignments when combined with substantive methods like the Principle of Indifference.

In light of my fourth argument, that applications of Bayesian reasoning tend to conform better to Explanationism than Orthodoxy, you may wonder what Explanationism can really teach us. Even if philosophers don't explicitly endorse the view, don't they already tacitly assume it in their reasoning? Unfortunately, while many applications of probability conform to Explanationism, the lack of explicit attention to the structure of probabilities leads to both incorrect expositions of basic concepts and bad reasoning about more complicated examples. This is especially so when it comes to the use of Bayes' Theorem in calculating probabilities.

Expositors of Bayes' Theorem,[26]

$$P(H \mid E\&K) = \frac{P(H \mid K)P(E \mid H\&K)}{P(H \mid K)P(E \mid H\&K) + P(\sim H \mid K)P(E \mid \sim H\&K)}$$

frequently speak as if the "empirical data" E that enters into it is always "new evidence we have

[26] For the rest of this essay I make our background knowledge K explicit in all probabilities. As we will see, this becomes philosophically important when we try to apply Explanationism more generally.

just acquired" (Salmon 1990: 177). Others describe Bayes' Theorem "as a normative rule for *updating* beliefs in response to evidence" (Pearl 1988: 32-33, emphasis mine). However, our having just learned a proposition E is neither necessary nor sufficient for Bayes' Theorem to break P(H|E&K) into more basic quantities. All that is necessary is that the evidence E is explanatorily downstream from the hypothesis H.

The terms in Bayes' Theorem are often divided into "priors" (P(H|K), P(~H|K)), "likelihoods" (P(E|H&K), P(E|~H&K)), and "posteriors" (P(H|K)). Many philosophers of probability attach undue metaphysical weight to these divisions, holding that there is a special problem with determining the values of prior probabilities. Talbott (2015: 4.2.F) is representative: "Are there constraints on prior probabilities other than the probability laws? This is the issue that divides the Subjective from the Objective Bayesians."[27]

Other philosophers have pointed out that the assumption that only prior probabilities are difficult to determine is dubious: for example, Earman (1992: 84) writes that "while much of the attention on the Bayesian version of the [Duhem] problem has focused on the assignments of prior probabilities, the assignments of likelihoods involves equally daunting difficulties." But the assumption that likelihoods are objective, while priors are not, is not only dubious—it is impossible. There can be no intrinsic difference between prior probabilities and likelihoods because these terms describe not the intrinsic nature of different probabilities, but their functional role in a particular application of Bayes' Theorem. In different instances of Bayes' Theorem, one and the same probability can be both a prior probability and a likelihood.

[27] Compare Gillies (1991: 530), who writes that likelihoods "can usually be calculated in a quite unproblematic manner," and Hawthorne (1994: 241), who contrasts "objective" likelihoods and "subjective" priors. So-called "swamping solutions" to the problem of the priors, according to which agents with different priors eventually converge on their posteriors given enough data, likewise presuppose intersubjective agreement on likelihoods. The older terminology of 'direct probability' instead of 'likelihood' also seems to take for granted that the values of likelihoods are obvious.

For example, consider the proposition C: a coin will be flipped to choose between urns U_1 and U_2. $P(U_1|C)$ will be a "likelihood" if we are calculating the posterior probability of C, $P(C|U_1)$, and it will be a "prior probability" if we know C and are calculating the posterior probability of U_1 given that we draw black, $P(U_1|C\&B)$. Either way, $P(U_1|C)$ is a basic probability, and we can see that its value is 1/2. What matters for determining the values of probabilities is not whether they are likelihoods or priors, but whether they are basic or non-basic, and if they are non-basic, what basic probabilities they can be reduced to.

The assumption that there is a special problem with the objectivity of prior probabilities has led most philosophers who discuss the problem of determining the values of probabilities to misconstrue it as the "problem of the priors." In turn, most existing solutions to the problem of the priors are based on a false presupposition—namely, that the unconditional, or "intrinsic," probabilities of hypotheses are basic.[28] On Explanationism, this amounts to the assumption that when we have no background knowledge, the partition of rival hypotheses being assigned (unconditional) prior probabilities in a problem is a root node in the Bayesian network representing our hypothesis space; that is, it has no parents. Substantive methods like the ones discussed in section 3.6 can then be applied to that partition: for example, a flat (indifferent) distribution can be assigned over the partition, or the hypotheses in the partition can be ranked in order of simplicity, with higher probabilities given to simpler hypotheses.

In idealized cases (including the urn scenarios above) it is often useful to assume that

[28] Jaynes (2003: ch. 11-12) seems to assume this in his defense of MaxEnt. And in his defense of simplicity as a criterion of prior probability, Swinburne (2001: ch. 4) suggests that when we have no background knowledge and two hypotheses have equal scope, the simpler hypothesis will always have higher prior probability. (For similar claims, see Plantinga 1993: 145-146 and Draper 2016. Draper's proposed criterion for which he thinks this is true is coherence, rather than simplicity.) Bayesian discussions of the history of science, such as Salmon 1990: 181-87, tend to do better than more abstract solutions to the problem of the priors at acknowledging the role background theories play in determining prior probabilities.

prior probabilities are basic. But in real-life Bayesian reasoning the prior probability of almost any hypothesis is non-basic. This is because there are almost always other theories explanatorily prior to the hypothesis which make a difference to how likely it is to be true.

For example, consider the formulation of Darwin's theory of evolution by natural selection. The prior probability of Darwinism (i.e., its probability apart from the data explanatorily downstream from it) was not basic. Rather, it was influenced by such considerations as empirical data suggesting that the earth was comparatively young, so that there had not been sufficient time for the speciation required by Darwin's theory to take place (McGrew et al. 2009: 242). The age of the Earth is explanatorily prior to the origins of Earth's species, and so in evaluating the prior probability of a theory about the latter we need to sum over different hypotheses about the former and about other relevant higher-level possibilities. For example, the network in Fig. 10 lets us calculate the probability of Darwinism as follows:

$$P(Darwinism|K) = \sum_i \sum_j P(A_i|K)P(T_j|K)P\big(Darwinism|A_i\&T_j\&K\big)$$

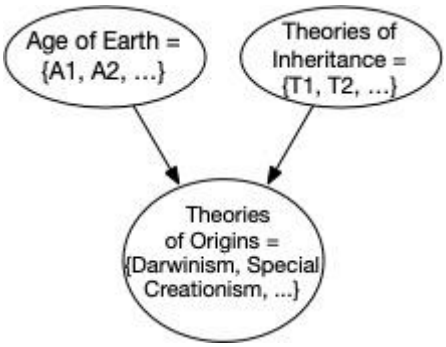


Figure 10

Historically we had empirical data relevant to the higher-level hypotheses in this network. But the structure of the network is not dependent on the existence of these data. Evidence that the earth is young is not necessary for us to see that the possibility of the degree of speciation necessary to produce the variety of life on earth today depends on how old the earth is.

So even in the absence of such background knowledge, the prior probability of Darwinism would still be a function of its probability on different combinations of higher-order theories like those in Fig. 10, weighted by the prior probability of those combinations. (These priors will be influenced by even more explanatorily basic hypotheses, suggesting that we need to expand the above Bayesian network. How far back we need to expand it—at what point we reach explanatorily fundamental theories, or ultimate explanations—is a large question which I do not have space to address here.)

It follows that how well Darwinism and Special Creationism satisfy proposed criteria of theory choice, such as simplicity, is not directly relevant to their relative prior probabilities, *when those simplicities are measured in the absence of potential background explanations*. Their prior probabilities are a function of their probabilities conditional on conjunctions of higher-order theories. These conditional probabilities may partially be a function of the simplicity of Darwinism and Special Creationism *relative to these conjunctions*; but in this case what matters is not how simple the two theories are unconditionally, but how simple they are when we assume the truth of particular higher-order theories.[29]

5. Conclusion

How are the values of epistemic probabilities determined? In this paper I have taken a first step towards answering this persistently difficult question. In particular, I have addressed the

[29] Perhaps because the role of background considerations is so obvious here, I know of no Bayesian attempts to directly apply substantive methods to the prior probability of Darwinism and rival theories of the origins of Earth's species. But some Bayesians have tried to apply substantive methods to the prior probability of rival physical theories in a similarly misguided way. For example, Swinburne (2001) suggests that Newton's law of gravity has a higher intrinsic probability than its rivals because it is simpler than them. But the argument here shows that if this law is more intrinsically probable than its rivals, this is because it is more likely given particular theories about the origins of the universe and its physical laws, not because it is simpler (except insomuch as this makes a difference to how likely it is on these theories of cosmic origins). (It might be legitimate to apply substantive methods to rival physical laws directly if one thought it *a priori* true that there are no deeper explanations of these laws. But Swinburne is a theist, and thinks that God's actions explain why our universe has the physical laws it does.)

structural problem of how to “break down” a non-basic probability into basic probabilities of which it is a function. I have defended a view on which the explanatory structure of probabilities is determined by the explanatory structure of the propositions these probabilities relate. We obtain basic probabilities by explanatorily ordering different partitions of propositions, and determining which propositions potentially explain the truth of other propositions. Consideration of both simple thought experiments and actual applications of probabilistic reasoning reveals that we do conceive of basic probabilities in this way.

On the Orthodox approach, the probabilities of complete state-descriptions are basic, and other probabilities are determined as a function of those. Because Orthodoxy ignores the explanatory relations between the conjuncts of state-descriptions, it conflicts with our intuitive judgments about what probabilities are basic. Moreover, when combined with substantive methods for determining probabilities, it delivers the wrong results. By ignoring the asymmetry of explanation, it wrongly allows the addition of future, explanatorily downstream, variables to alter the probability distribution over past, explanatorily upstream, variables.

Explanationism has important implications for many debates in epistemology and philosophy of science. In particular, it sheds light on informal debates about the substantive problem, such as the literature on the so-called “problem of the priors.” According to the Explanationist, these debates are largely misconceived, treating prior probabilities of empirical hypotheses as *sui generis*, rather than imposed on them by explanatorily prior theories.

There remain significant open questions about how to flesh out the Explanationist picture:

- Besides causal and metaphysical priority, are there other kinds of explanatory priority relevant to constructing a Bayesian network?

- Is an infinite explanatory regress possible? Or is the Explanationist committed to there

being a first cause/ultimate explanation?

- In cases of network uncertainty, is an infinite regress of higher-order networks possible? Or is the Explanationist committed to there being some *a priori* higher-order network that relates all lower-order networks?

All these issues deserve further investigation. In addition, the substantive question of what determines the values of basic probabilities continues to loom large.

Daunting questions remain, then. Nevertheless, the Explanationist picture seriously advances the project of determining the values of epistemic probabilities, laying a foundation for further work and dispelling much of the dust and confusion surrounding this thorny project.

Acknowledgments

I am grateful to Daniel Immerman, Andrew Brenner, Jeff Tolly, Lane DesAutels, Fritz Warfield, Branden Fitelson, Christopher Hitchcock, Alex Pruss, Steve Finlay, Al Hájek, John Hawthorne, Simon Goldstein, Nicholas DiBella, and an anonymous reviewer for *Philosophical Studies* for comments on earlier drafts; and to audiences at Houston Baptist University, Messiah College, LMU Munich, Australian National University, and the Eastern APA for feedback on presentations of this project.

References

Bovens, Luc and Stephan Hartmann (2003). *Bayesian Epistemology* (Oxford University Press).

Bradley, Darren (forthcoming). “Naturalness as a Constraint on Priors.” *Mind*.

Carnap, Rudolf (1950). *Logical Foundations of Probability* (University of Chicago Press).

Climenhaga, Nevin (2017a). “How Explanation Guides Confirmation.” *Philosophy of Science* 84: 359–68.

Climenhaga, Nevin (2017b). “Inference to the Best Explanation Made Incoherent.” *Journal of*

Philosophy 114: 251–73.

Cox, R.T. (1946). “Probability, frequency and reasonable expectation,” *American Journal of Physics* 17: 1–13.

Douven, Igor (2013). “Inference to the best explanation, Dutch books, and inaccuracy minimisation,” *The Philosophical Quarterly* 63*:* 428-444.

Douven, Igor, & Jonah N. Schupbach (2015). “Probabilistic alternatives to Bayesianism: the case of explanationism.” *Frontiers in psychology* 6: 1-9.

Draper, Paul (2016). “Simplicity and Natural Theology,” in *Reason and Faith: Themes from Richard Swinburne*, ed. Michael Bergmann and Jeffrey E. Brower (Oxford University Press).

Earman, John (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory* (MIT Press/Bradford Books).

Fienberg, Stephen E. (2006). “When Did Bayesian Inference Become “Bayesian”?” *Bayesian Analysis* 1: 1-40.

Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Bunson, Aki Vehtari, and Donald B. Rubin (2014). *Bayesian Data Analysis*, Third Edition (Chapman & Hall).

Gillies, Donald (1991). “Intersubjective Probability and Confirmation Theory,” *British Journal for the Philosophy of Science* 42: 513-33.

Hacking, Ian (2001). *An Introduction to Probability and Inductive Logic* (Cambridge University Press).

Hájek, Alan (2003). “What Conditional Probability could not be,” *Synthese* 137: 273–323.

Hawthorne, James (1994). “On the Nature of Bayesian Convergence” *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1: 241-249.

Hawthorne, James (2005). “Degree-of-Belief and Degree-of-Support: Why Bayesians Need Both Notions,” *Mind* 114: 277-320.

Hedden, Brian (2015a). “A Defense of Objectivism about Evidential Support.” *Canadian Journal of Philosophy* 45: 716-743.

Hedden, Brian (2015b). “Time-Slice Rationality.” *Mind* 124: 449-491.

Henderson, Leah (2014). “Bayesianism and Inference to the Best Explanation,” *British Journal for the Philosophy of Science* 65: 687-715.

Henderson, Leah, Noah D. Goodman, Joshua B. Tenenbaum, and James F. Woodward (2010).

"The structure and dynamics of scientific theories: a hierarchical Bayesian perspective." *Philosophy of Science* 77: 172-200.

Hesse, Mary (1974). *The Structure of Scientific Inference* (University of California Press).

Hitchcock, Christopher (2012). "Probabilistic Causation," *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2012/entries/causation-probabilistic/>.

Howson, Colin and Peter Urbach (2006). *Scientific Reasoning: The Bayesian Approach*, Third Edition (Open Court).

Huemer, Michael (2009). "Explanationist aid for the theory of inductive logic," *British Journal for the Philosophy of Science* 60: 345–375.

Jaynes, E.T. (2003). *Probability Theory: the Logic of Science* (edited by G.L. Bretthorst) (Cambridge University Press).

Jeffrey, Richard (1983). *The Logic of Decision, 2nd edition* (University of Chicago Press).

Jeffreys, Harold (1939/1998). *Theory of Probability*, reprinted in Oxford Classics in the Physical Sciences series (Oxford University Press).

Joyce, James (2008). "Bayes' Theorem," in *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/bayes-theorem/>.

Keynes, John Maynard (1921). *A Treatise on Probability* (Macmillan and Co.).

Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitrechnung* (Ergebnisse Der Mathematik); translated as *Foundations of Probability* (Chelsea Publishing Company, 1950).

Korb, Kevin B. and Ann E. Nicholson (2011). *Bayesian Artificial Intelligence*, 2nd edition (Chapman & Hall).

Lange, Marc (2018). "Transitivity, Self-Explanation, and the Explanatory Circularity Argument against Humean Accounts of Natural Law," *Synthese* 195: 1337-1353.

Lipton, Peter (2004). *Inference to the Best Explanation*, 2nd edition (Routledge).

Maher, Patrick (2004). "Probability captures the logic of scientific confirmation," in *Contemporary Debates in Philosophy of Science*, ed. Christopher R. Hitchcock (Blackwell), 69–93.

Maher, Patrick (2006). "The Concept of Inductive Probability," *Erkenntnis* 65: 185-206.

McGrew, Timothy, Marc Alspector-Kelly, and Fritz Allhoff, ed. (2009). *Philosophy of Science: An Historical Anthology* (Wiley).

Pearl, Judea (1988). *Probabilistic Reasoning in Intelligent Systems* (Morgan Kaufmann).

Pearl, Judea (2000). *Causality: Models, reasoning, and inference* (Cambridge University Press).

Plantinga, Alvin (1993). *Warrant and Proper Function* (Oxford University Press).

Ramsey, Frank (1926/1990), "Truth and Probability," in his *Philosophical Papers* (Cambridge University Press).

Rathmanner, Samuel and Marcus Hutter (2011). "A Philosophical Treatise of Universal Induction," *Entropy* 13: 1076-1136.

Russell, Stuart and Peter Norvig (2009). *Artificial Intelligence: A Modern Approach*, 3rd edition (Pearson).

Salmon, Wesley (1990). "Rationality and Objectivity in Science or Tom Kuhn Meets Tom Bayes," in *Scientific Theories*, C.W. Savage (ed.), *Minnesota Studies in the Philosophy of Science* (University of Minnesota Press) 14: 175–204.

Schaffer, Jonathan (2016). "Grounding in the Image of Causation," *Philosophical Studies* 173: 49-100.

Solomonoff, R.J. (1964). "A Formal Theory of Inductive Inference, Part I," in *Information and Control* 7: 1-22.

Spirtes, Peter, Clark Glymour, and Richard Scheines (2000). *Causation, Prediction and Search*, second edition (M.I.T. Press).

Swinburne, Richard (2001). *Epistemic Justification* (Oxford University Press).

Swinburne, Richard (2011). "Gwiazda on the Bayesian Argument for God," *Philosophia* 39: 393-396.

Talbott, William (2015). "Bayesian Epistemology," in *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2015/entries/epistemology-bayesian/>.

Thagard, Paul (1978). "The Best Explanation: Criteria for Theory Choice," *Journal of Philosophy* 75: 76-92.

Tooley, Michael (2012). "Inductive Logic and the Probability that God Exists: Farewell to Skeptical Theism," in *Probability in the Philosophy of Religion*, ed. Jake Chandler and Victoria S. Harrison (Oxford University Press).

Venn, John (1866). *The Logic of Chance* (London and Cambridge: Macmillan and Co.).

Weisberg, Jonathan (2009). “Locating IBE in the Bayesian Framework,” *Synthese* 167: 125–144.

Weisberg, Jonathan (2015). “Formal Epistemology,” in *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), edited by Edward N. Zalta, URL = <http://plato.stanford.edu/archives/sum2015/entries/formal-epistemology/>.

Williamson, Jon (2005). *Bayesian Nets and Causality: Philosophical and Computational Foundations* (Oxford University Press).

Williamson, Jon (2010). *In Defence of Objective Bayesianism* (Oxford University Press).

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