Not Enough There There

Evidence, Reasons, and Language Independence

Michael G. Titelbaum[1]

Draft of March 19, 2009.
Please do not circulate or cite without permission.

It is natural to think that given any body of evidence and set of hypotheses, that evidence favors some of the hypotheses over others. One strong way to maintain this position is to hold that any body of evidence confers a particular *evidential probability* on any particular hypothesis, which we can understand roughly as how likely the evidence makes it that the hypothesis is true.[2] One hypothesis is then favored over another hypothesis on the evidence just in case that evidence confers a higher probability on the former than the latter. It is sometimes suggested that an agent should assign a *degree of belief* to a hypothesis equal to the probability conferred on that hypothesis by his total evidence, but defenders of evidential probabilities need not accept this claim.

Maintaining the existence of evidential probabilities is a particularly strong way of holding that there are objective facts about when a body of evidence favors one hypothesis over another. Plenty of views suggest that such favoring relations can obtain without demanding that evidence confer anything so precise as a numerical probability (or even a range of probabilities) on a given hypothesis. For example, in epistemology we sometimes say that an agent's total evidence *supports* one hypothesis over another, or that it *provides more justification* for believing one hypothesis than the other. In metaethics and the theory of normativity, we say that an agent's total evidence gives him *more reason* to believe one hypothesis than another. Evidential favoring need not even be based on a gradable notion: we might say that an agent's evidence permits (or requires) him to believe one hypothesis but not another, and that in that sense it favors the former over the latter.

What all of these views have in common is belief in the existence of what I will call an evidential favoring relation: an objective three-place relation between two hypotheses (understood as propositions) and a body of evidence (also understood as a proposition). While subjective facts about an agent may play a role in determining what counts as his total evidence, defenders of an evidential favoring relation maintain that what a particular body of evidence favors is independent of any subjective considerations. My goal in this essay is to challenge the claim that there is such an evidential favoring relation.[3]

I will start with a problem that has plagued theorists of evidential (or "logical" or "epistemic") probabilities. Over the course of the twentieth century a number

[1]Besides the people mentioned at specific points in this text, I am grateful to audiences at the Australian National University's Philosophy Society and the University of Melbourne's Logic Seminar. I am also grateful to Fabrizio Cariani for teaching me about language independence, and to Branden Fitelson for conversations that helped me develop many of the views expressed here.

[2]Williamson (2000, Chapter 10) writes, "Given a scientific hypothesis h, we can intelligibly ask: how probable is h on present evidence? We are asking how much the evidence tells for or against the hypothesis."

[3]I assume that the evidential favoring in question is purely *epistemic*; if evidence can give an agent *pragmatic* reason to believe one hypothesis over another, such pragmatic favoring will lie outside our purview.

of formal theories proposed to reveal the probability of any hypothesis relative to any evidential proposition. Each of these theories had its own problems, but there was one problem they all shared: language dependence. The theory would assign one probability to a hypothesis relative to an evidential proposition, then when the hypothesis and evidence were re-expressed in a different language the probability assigned by the theory would change. Favoring relations were inverted as well: one hypothesis would receive a higher probability than another when the hypotheses and evidence were expressed in one language, then the order would be reversed relative to a different language. The most famous example of language dependence was Goodman's "grue" problem for Hempel's and Carnap's theories of evidential support, but more recent theories (such as Jaynes's "maximum entropy" approach) have suffered at the hands of language dependence as well.[4]

The first section of this paper shows that language dependence problems are an unavoidable obstacle for formal theories of three-place evidential favoring. The trouble isn't that no one has been clever enough in cooking up a theory to capture evidential favoring, and the trouble isn't even the project of understanding favoring relations in terms of probabilities. In Section 1 I articulate some basic, intuitive conditions on any evidential favoring relation, then present a general result showing that no relation expressible in a formal theory can meet all these conditions. (The result is proven in Appendix A.)[5]

It's tempting to read this as a lesson in the folly of formal theorizing. But in Section 2 I suggest that our formal failures are revealing something deeper about the evidential favoring relation itself, as it exists out there among the propositions (so to speak). If there is an evidential favoring relation, it must play favorites among properties; it must accord some properties more special status than others. This presents a challenge to defenders of an evidential favoring relation: Presumably reasoning agents can determine when evidential favoring relations obtain, which means they can determine the list of special properties (at least in part). But (I argue in Section 3) they cannot determine that list *from* their evidence, since they need to know the list in order to determine what their evidence supports. The only plausible option for the evidential favoring theorist is to adopt a strong theory of the *a priori* on which the special properties list can be discerned antecedent to our evidence. In Section 4 I criticize this approach and suggest what seems to me a better alternative: dropping the notion of an objective three-place evidential favoring relation and replacing it with a theory on which what an agent's evidence favors may be relative to other subjective facts about the agent.

1. The General Result

We begin with our general result about formal theories of evidential favoring. Since the technical details get a bit hairy and I want to keep our attention focused on the result's philosophical implications, the proof is reserved for Appendix A. Here I will sketch out the definitions and conditions needed for the result, then

[4] For grue, see (Goodman 1979, Chapter III). For Hempel's theory see (Hempel 1945); for Carnap's, (Carnap 1950) and (Carnap 1952). The maximum entropy approach was introduced in (Jaynes 1957a) and (Jaynes 1957b); for language dependence criticisms see (Seidenfeld 1986).

[5] While I have tended in this introductory discussion to emphasize what Carnap called "firmness" notions of favoring, the proof also applies to formal theories of "increase in firmness" notions such as confirmation. (For further discussion of the firmness/increase in firmness distinction and references to the Carnap, see (Fitelson 2006).)

state the result itself.[6] To help the reader keep track, I will **bold** a technical term when it is introduced. While I hope to say enough here to make the requirements for the proof intuitively appealing, responses to some specific objections and concerns will be put off until Appendix B.

Suppose there exists a three-place relation between hypotheses and bodies of evidence such that there are facts of the form "this evidence favors this hypothesis over this other hypothesis." We will take each of the three relata to be a proposition—if you prefer to think of a body of evidence as a *set* of propositions, we can use the conjunction of that set's members as our third relatum. A formal theory of evidential favoring will analyze the favoring relation by first representing the hypotheses and evidence as sentences in a formal language, then looking for structural patterns among those sentences that indicate a favoring relation among the propositions they represent.

So we'll begin by constructing formal **languages** to represent hypotheses and evidence. We will be working with finite first-order languages with no variables and no quantifiers. Each language will come with a syntactic consequence relation (symbolized "$\vdash$") and an interpretation. In a language $\mathcal{L}$ the predicates $(F, G, H, \ldots)$ will represent particular properties, the names $(a, b, c, \ldots)$ will represent particular objects, and sentences (well-formed formulas) will represent propositions. Our evidence and two hypotheses will be represented by the sentences $e, h_1, h_2 \in \mathcal{L}$ (respectively).[7]

Though each language we work with will have a finite number of predicates and names, we will impose no (finite) upper limit on the size of the languages available. This is in part because we require languages to be plentiful in the following sense:

> **Availability of Independent Properties:** Given any language $\mathcal{L}$, there exists another language whose names refer to all the same objects and whose predicates express all the properties expressed by the predicates of $\mathcal{L}$—as well as one additional property logically independent of those.

Once we have propositions represented in a language, the relation we are interested in is whether the evidence represented by e favors the hypothesis represented by h_1 over the hypothesis represented by h_2. When this relation holds, I will write $h_1 \succ_e h_2$. Please note that $h_1 \succ_e h_2$ is a metalinguistic abbreviation for something we say *about* elements in a language $\mathcal{L}$ (namely, that the proposition represented by h_1 is favored over the proposition represented by h_2 relative to the proposition represented by e); $\succ_e$ is not a symbol *in* the language $\mathcal{L}$ and $h_1 \succ_e h_2$ is not a sentence *of* $\mathcal{L}$.

Here's an example: You've been observing a number of sea creatures, some of whom have fins on their sides. We might represent some of your observations in a language $\mathcal{L}^{LR}$ with six names (a through f) and two predicates: L for the property of having a left fin and R for the property of having a right fin. Suppose your evidence to this point is represented in language $\mathcal{L}^{LR}$ as

[6] Some of the definitions and conditions are stated more precisely in Appendix A than they are here.

[7] I have now begun what will become a habit of leaving use/mention disambiguation to context. I will also leave it to the reader to discern when a symbol is being used as a name and when it is being used as a variable. For example, "h_1", "h_2", and "e" are metavariables, not actual sentences of $\mathcal{L}$.

$$\begin{array}{ccccccccccccc} La & \& & \sim Lb & \& & \sim Lc & \& & Ld & \& & \sim Le & \& & Lf & \& \\ Ra & \& & \sim Rb & \& & \sim Rc & \& & Rd & \& & \sim Re & & & & \end{array}$$

Looking at this evidence, we see a pattern: every creature with a left fin also has a right fin. Since you've already observed that creature f has a left fin, you should expect creature f to have a right fin as well. In other words, this evidence seems to favor the proposition that f has a right fin over the proposition that f lacks one. If we refer to the evidence sentence above as e, we have $Rf \succ_e \sim Rf$.

Setting up the $\succ_e$ relation in this fashion has the advantage that it automatically gives us the following condition:

> **Language Invariance:** Given any language $\mathcal{L}$ containing sentences h_1, h_2, and e, and any language $\mathcal{L}'$ containing h'_1, h'_2, and e' that express the same propositions as h_1, h_2, and e (respectively), $h_1 \succ_e h_2$ just in case $h'_1 \succ_{e'} h'_2$.

This condition follows automatically from our definitions because $h_1 \succ_e h_2$ just in case the *propositions* expressed by h_1, h_2, and e stand in a particular favoring relation; if we re-express the same propositions using different sentences in a different language, it will still be the case that the propositions represented by those sentences stand in the relevant relation. $\succ_e$ is language invariant because it tracks relations among the propositions that *underlie* a language's sentences; since evidential favoring is a relation among propositions, this is the kind of relation we want a formal theory of evidence to track.

The goal of such a theory is to tell us whether $h_1 \succ_e h_2$ obtains for any h_1, h_2, and e. In our sea creatures example, we would want our formal theory of evidence to be able to look at the evidence sentence e, detect the pattern in the evidence, and then predict that $Rf \succ_e \sim Rf$. But it would be unfair to expect a formal theory to detect favoring patterns when evidence is represented in *any* language. For example, suppose we are working with the language $\mathcal{L}^{LNT}$ that has three predicates: L for the property of having a left fin, N for the property of having no fins, and T for the property of having two fins. The evidence we expressed as e above could be re-expressed in this language as

$$Ta \quad \& \quad Nb \quad \& \quad Nc \quad \& \quad Td \quad \& \quad Ne \quad \& \quad Lf$$

This sentence (call it e') expresses exactly the same information as was expressed in our original evidence sentence. Moreover, the proposition expressed by e' favors the proposition expressed by Tf over the proposition expressed by $\sim Tf$: every creature for which both fins have been observed has either had no fins or two fins; since the evidence indicates that creature f does not fall into the former camp, it should be expected to fall into the latter. So $Tf \succ_{e'} \sim Tf$. But it would be unreasonable to expect a formal theory of evidence to reach that conclusion just by looking at sentence e'. e' has no pattern that such a formal theory could pick up on to generate this result; we were able to discern the relevant evidential relation only by thinking about information not available in the sentence e' (information about what properties the predicates involved represent).[8]

[8]You might think that e' displays a pattern of only affirming predicates of names, and that if we extend this pattern the proposition represented by e' will favor the proposition represented by Tf over the proposition represented by $\sim Tf$. But by that reasoning the proposition represented by e' would favor the proposition represented by Nf over the proposition represented over $\sim Nf$, which it most certainly does not—the proposition represented by e' actually refutes the former while entailing the latter.

The trouble here is that $\mathcal{L}^{LNT}$ is a bad language for representing the evidential structure of our sea creatures situation—it was deliberately constructed to obscure patterns in the evidence. For any evidence and pair of hypotheses, it will be possible to construct a language that obscures the favoring relations among them. I will refer to such a language as an "inadequate" language, and refer to a language that lacks such obscuring features as an **adequate** language for the three relata. (When context makes the hypotheses and evidence in question clear, I'll sometimes just say that a particular language is "adequate.") I do not have a formal definition of language adequacy to offer, but I do have two conditions that I think any plausible account of adequacy will meet.

First, let's look at what's wrong with language $\mathcal{L}^{LNT}$. When we reasoned through the favoring relation between e', Tf, and $\sim Tf$, we took advantage of some entailment relations between the properties N, T, and L (for example, the fact that Lf entails $\sim Nf$). Yet these entailment relations are not reflected in the syntactic structure of $\mathcal{L}^{LNT}$; there is no way for a formal theory working just with sentences in that language to know that Lf entails $\sim Nf$. This is because the syntactic consequence relation in language $\mathcal{L}^{LNT}$ does not match up with the entailments between the propositions the language's sentences represent. To prevent this problem, we will require adequate languages to meet the following condition:

Faithfulness: A language $\mathcal{L}$ is faithful when for any $x, y \in \mathcal{L}$, $x \vdash y$ just in case the proposition represented by x entails the proposition represented by y.

Among other things, faithfulness requires a language's atomic sentences (Fa, Gb, etc.) to represent propositions that are logically independent: atomic sentences have no syntactic consequence relations among them, so any entailment relations among the propositions they represent would cause a failure of faithfulness.[9] (This is why language $\mathcal{L}^{LNT}$ fails to be faithful.)[10]

Our definition of faithfulness yields another fact as well:

Equivalence Condition: For any e, h_1, h_2, h_1', h_2', and e' in a faithful language $\mathcal{L}$, if $h_1 \dashv\vdash h_1'$, $h_2 \dashv\vdash h_2'$, and $e \dashv\vdash e'$ then $h_1 \succ_e h_2$ just in case $h_1' \succ_{e'} h_2'$.

Assuming $\mathcal{L}$ is faithful, $h_1 \dashv\vdash h_1'$ tells us that the proposition represented by h_1 is logically equivalent to the proposition represented by h_1'; taking propositions to be sets of possible worlds, this means that h_1 and h_1' express the same proposition. Similarly, h_2 and h_2' express the same proposition and e and e' express the same proposition. Thus the proposition expressed by e favors the proposition expressed

[9] Goodman (1946, n. 2) notes that Hempel, Carnap, and company all recognized that formal languages used to analyze evidential favoring relations must have atomic sentences representing logically independent propositions.

[10] Faithfulness says that syntactic consequence among sentences holds just when semantic entailment among the represented propositions does. The need to keep syntactically independent sentences representing semantically independent propositions explains why semantic entailment should hold only when syntactic consequence does. But why do we require syntactic consequence to obtain only when semantic entailment does?

In a moment we'll be looking at intuitive conditions on the evidential favoring relation. Some of these depend on entailment facts among evidence and hypothesis propositions. A formal theory that operates on a language and must work with consequence facts within that language can capture these conditions on evidential favoring only if the consequence facts reflect the entailment facts among the propositions.

by h_1 over the proposition expressed by h_2 just in case the proposition expressed by e' favors the proposition expressed by h'_1 over the proposition expressed by h'_2.

To generate our next condition on language adequacy, we have to consider another stupid way in which a language could represent your sea creature information. Suppose we have a language $\mathcal{L}^{FG}$ with one name t representing the sextuple consisting of sea creatures a through f. This language has two predicates, F and G. The property represented by F applies to a sextuple just in case its first and fourth elements have a right and left fin, its second, third, and fifth elements have no right or left fin, and its sixth element has a left fin. The property represented by G applies to a sextuple just in case its sixth element has a right fin. This language is faithful; there are no logical dependencies between propositions expressible using F and propositions expressible using G. Moreover, it is capable of expressing the evidence we expressed earlier using the sentence e. In language $\mathcal{L}^{FG}$, this evidence becomes

$$Ft$$

As we saw earlier, the evidential proposition represented by this sentence favors the proposition that creature f has a right fin over the proposition that f lacks one. So $Gt \succ_{Ft} \sim Gt$. Yet this language has completely obscured the evidential structure that leads to this favoring relation; there's no way a formal theory could look at the sentences Ft and Gt and discern that $Gt \succ_{Ft} \sim Gt$.

The trouble with language $\mathcal{L}^{FG}$ is that it is not expressive enough to reveal the substructure of our evidence and hypotheses. $\mathcal{L}^{FG}$ has been designed with just enough names and predicates to be able to represent the evidence and hypotheses, but it is incapable of expressing the subparts of those propositions that create the evidential favoring relation. To remedy this, a plausible notion of language adequacy would have to put some floor on how expressive a language needs to be to be adequate for a particular evidence and set of hypotheses. For instance, in our sea creatures example an adequate language would have to be capable of expressing facts about whether each individual creature has a left fin or a right fin.

I have very little to say about how this expressiveness floor should work, or how we can tell when a language is expressively adequate for some particular evidence and pair of hypotheses. But I want to point out that the concern is about a *floor* for expressiveness; adding expressive capability to a language (as long as the result is faithful) should never take it from adequate to inadequate. Put more precisely, if we are given faithful languages $\mathcal{L}$ and $\mathcal{L}'$ such that every proposition expressible in $\mathcal{L}$ is expressible in $\mathcal{L}'$ and $\mathcal{L}$ is adequate for some particular evidence and pair of propositions, a plausible notion of adequacy will deem $\mathcal{L}'$ adequate for those relata as well.

Let's illustrate with our sea creatures example. We've already seen adequate language $\mathcal{L}^{LR}$. Now consider faithful language $\mathcal{L}^{LS}$, whose predicates are L for the property of having a left fin and S for the property of being symmetrically finned. $\mathcal{L}^{LS}$ is capable of expressing every proposition expressible in $\mathcal{L}^{LR}$; the expression of L-facts is obvious, while any proposition in $\mathcal{L}^{LR}$ of the form Rx can be represented in $\mathcal{L}^{LS}$ as $Lx \equiv Sx$. Represented in $\mathcal{L}^{LS}$, the evidence represented by e becomes

$$\begin{array}{ccccccccccc} La & \& & \sim Lb & \& & \sim Lc & \& & Ld & \& & \sim Le & \& & Lf & \& \\ Sa & \& & Sb & \& & Sc & \& & Sd & \& & Se & & & & \end{array}$$

This representation certainly reveals the pattern in your evidence—every sea creature whom you've examined on both sides has been symmetrically finned. (It arguably reveals the evidential pattern even better than the representation in $\mathcal{L}^{LR}$!) $\mathcal{L}^{LS}$ is clearly an adequate language for this evidential situation.

To sum up: our two conditions on a plausible notion of adequacy are first, that faithfulness is necessary for adequacy, and second, that any faithful language capable of expressing every proposition expressed in an adequate language is itself adequate.

Having considered the properties and availability of languages, let's turn our attention to intuitively plausible conditions on the evidential favoring relation. First, the favoring relation induces a **strict ordering** on hypotheses; that is, for any e, $\succ_e$ is antisymmetric ($h_1 \succ_e h_2$ entails $h_2 \not\succ_e h_1$) and transitive ($h_1 \succ_e h_2$ and $h_2 \succ_e h_3$ entail $h_1 \succ_e h_3$). However, we will not assume that evidential favoring induces a **total** ordering; we will not assume that for any h_1, h_2, and e either $h_1 \succ_e h_2$ or $h_2 \succ_e h_1$.[11] This is to allow for the possibility of incommensurate favorings. For example, my current total evidence might favor the proposition that the Celtics will win the NBA championship over the proposition that the Lakers will win, and it might favor the proposition that Sarah Palin will run for President in 2012 over the proposition that she won't. But there may be no fact of the matter about whether my current total evidence favors the proposition that the Celtics will win the championship over the proposition that Palin will run for President. To allow for the possibility of incommensurate favorings, we will not assume that $h_1 \not\succ_e h_2$ entails $h_2 \succ_e h_1$, nor will we assume that $h_1 \not\succ_e h_2$ and $h_2 \not\succ_e h_1$ entails that the evidence represented by e favors hypotheses h_1 and h_2 *equally.*

Next we assume that the favoring relation satifies this condition:

> **Substantivity:** There exists at least one adequate language $\mathcal{L}$ and set of relata $h_1, h_2, e \in \mathcal{L}$ such that $h_1 \succ_e h_2$ while e, h_1, and h_2 are logically independent.

I can readily grant that when deductive relations get involved—for example, when e entails h_1 but refutes h_2—evidence may objectively favor one hypothesis over another. But as Hume taught us, if evidential relations are to be anything like those we rely on in daily life, they must obtain in a myriad of cases in which no special entailment relations are involved. For example, your current total evidence favors the proposition that the sun will come over the horizon tomorrow morning over the proposition that a giant Cadillac will.[12] For a more precise example, if your evidence is that a number between 1 and 10 was drawn at random and the result was between 1 and 5, this favors the hypothesis that the number is odd over the hypothesis that it is between 4 and 7.[13] These examples and countless others

[11]Our conclusions will of course be *consistent* with evidential favoring's inducing a total ordering; they just won't *require* that it do so.

[12]I'm assuming it's logically compatible with your evidence that both will occur. . . .

[13]Mentioning "randomness" in the evidence might beg the question in some way, so technically the evidence in this example should be something like "I drew a ball from a thoroughly-shaken urn containing balls numbered 1 through 10 and it came up between 1 and 5" and the hypotheses should be something like "I drew a ball from a thoroughly-shaken urn containing balls numbered 1 through 10 and it came up odd" and "I drew a ball from a thoroughly-shaken urn containing balls numbered 1 through 10 and it came up between 4 and 7." (This example works on both "firmness" and "increase in firmess" conceptions of favoring—see note 5 above.)

like them demonstrate that the evidential favoring relation must be **substantive**; it must go beyond special cases of logical dependence among evidence and hypotheses.

To this point we have assumed that there is such a thing as a three-place evidential favoring relation, we have talked about representing evidence and hypotheses in languages, and we have articulated plausible conditions on that favoring relation. Now let's further suppose that the evidential favoring relation is the sort of relation that can be entirely captured by a theory operating on formal patterns among the sentences that express evidence and hypotheses. If the theory of favoring is working only with *formal* (or *logical*) patterns in the hypotheses and evidence, it will pay no attention to which particular predicates play which roles in those propositions. For a favoring relation to be susceptible to this type of formal analysis, it must satisfy this condition:

Identical Treatment of Predicate Permuations: Given a language $\mathcal{L}$ adequate for $e, h_1, h_2 \in \mathcal{L}$, if a permutation of $\mathcal{L}$'s predicates converts e to e', h_1 to h_1', and h_2 to h_2' then $h_1 \succ_e h_2$ just in case $h_1' \succ_{e'} h_2'$.

To the sea once more: We have already seen that the evidence

$$\begin{array}{ccccccccccccc} La & \& & {\sim}Lb & \& & {\sim}Lc & \& & Ld & \& & {\sim}Le & \& & Lf & \& \\ Ra & \& & {\sim}Rb & \& & {\sim}Rc & \& & Rd & \& & {\sim}Re & & & & \end{array}$$

favors Rf over ${\sim}Rf$. Now imagine that the evidence

$$\begin{array}{ccccccccccccc} Ra & \& & {\sim}Rb & \& & {\sim}Rc & \& & Rd & \& & {\sim}Re & \& & Rf & \& \\ La & \& & {\sim}Lb & \& & {\sim}Lc & \& & Ld & \& & {\sim}Le & & & & \end{array}$$

favored ${\sim}Lf$ over Lf. This is of course possible; the evidential favoring relation might be arranged so as to treat facts about right fins differently from facts about left fins. But an evidential favoring relation arranged in that fashion could not be captured by a formal theory of favoring. Such a theory operates only on the *structure* of evidence and hypotheses as expressed in an adequate language, and these two cases have exactly the same structure. So if evidential favoring is susceptible to formal analysis, it must treat predicate permutations identically.

We now have all the conditions needed for our general result. The result itself is simple:

General Result: The availability of independent properties, substantivity, and the identical treatment of predicate permutations cannot all hold.

This is *very* bad news for formal theories of three-place evidential favoring. If the evidential favoring relation is to capture anything like our normal notion of evidence, it must be substantive—we make correct evidential inferences all the time that do not involve entailments. Yet this general result tells us that if the evidential favoring relation is substantive (and independent properties are always available), it cannot treat predicate permutations identically and therefore cannot be captured by a formal theory of evidence.

2. What This Tells Us About Evidential Favoring

So much the worse for formal theories of evidence, you might say. After all, evidential relations are enormously complex—perhaps it's not so surprising that their vast subtleties cannot be adequately captured by a theory that pushes symbols around, no matter how inventively it does so. As Goodman concluded after introducing his grue problem, "Lawlike or projectible hypotheses cannot be distinguished on any merely syntactical grounds." (1979, p. 83)[14]

But I think the general result doesn't just tell us about the evidential favoring relation's prospects for formal representation—I think it also tells us something about that relation itself, as it exists out among the propositions. Our general result tells us not just that a three-place evidential favoring relation must depend on the semantics of hypotheses and evidence, but which *portions* of the semantics it must depend on. After all, entailment relations are a part of semantics, and those relations are fully reflected in the syntax of a (faithful) formal language. Yet a theory that operates solely on syntax seems unable to capture a substantive favoring relation.

To understand what the general result is telling us about evidential favoring, let's look back at the conditions we needed for the theorem. The availability of independent properties concerns what properties there are, while substantivity is an intuitively required feature of the favoring relation itself. The only time we invoked the project of characterizing favoring with a formal theory was in defending the identical treatment of predicate permutations. If we give up the formal project, we can no longer require evidential favoring to satisfy this condition. But we can ask what it tells us that identical treatment of predicate permutations is incompatible with a substantive favoring relation.

What kind of favoring relation treats predicate permutations identically, and what kind of favoring relation fails to do so? A favoring relation that fails to treat predicate permutations identically *plays favorites* among properties. That is, it responds differently to evidence involving one property than it does to evidence that is identical except that it involves a different property. We saw this already in our last sea creatures example. There, we imagined one case in which every observed left-finned creature has also been right-finned and creature f is known to be left-finned, and another case in which every observed right-finned creature has been left-finned and creature f is known to be right-finned. Suppose that as the evidential favoring relation actually stands, the evidence in the first case favors creature f's having a right fin over creature f's lacking one, but the evidence in the second case favors creature f's lacking a left fin over creature f's having one. This would violate the identical treatment of predicate permutations. It would also mean that it was somehow a built-in feature of the evidential favoring relation that it treated left-finnedness differently than right-finnedness; one of those properties would have some sort of *special status* from the point of view of evidential favoring.

Our general result tells us that identical treatment of predicate permutations and substantivity are incompatible. A substantive evidential favoring relation must play favorites among properties. This explains why formal theories of evidence kept running into language dependence difficulties. From the point of view of a theory working with the syntax of formal languages, working within a particular language is a way of privileging particular properties, such as the ones that are represented by predicates in the language. To take a silly example, you might have a formal theory holding that hypotheses that affirm more predicates are always

[14]The discussion of language dependence with regards to evidential favoring has in many ways paralleled the language dependence discussion in the versimilitude literature. (Miller 2005, Ch. 11) provides a nice summary, and approvingly quotes the comment in (Niiniluoto 1987) that language dependence hits only "essentially syntactic [or]... 'linguistic' definitions of truthlikeness." (I am grateful to Lloyd Humberstone for directing me to Miller's book.)

favored over hypotheses that affirm fewer. Applying this theory to a language with a predicate O for "open" would favor hypotheses asserting the existence of open doors over hypotheses suggesting closed ones; applying the theory to a language with a predicate C for "closed" would have the opposite effect. Our general result tells us that in order to capture a substantive favoring relation, formal theories need a designated list of "special" languages to which they are to be applied—those languages that represent the underlying "special" properties in the right way. But being purely formal, the theories in question have no way to separate the good languages from the bad; every language with the same syntactical structure looks the same to a syntactical theory.

Yet ultimately this is not a point about formal theories of favoring. What we have learned is that it takes a *combination* of a body of evidence *and* a preferred property set (or preferred language, if you like) to substantively favor some hypotheses over others. By itself, the informational content of a body of evidence (understood as a proposition) is insufficient to yield a substantive evidential favoring relation.

We can also (somewhat tentatively) characterize our result in terms of a supervenience thesis. Roughly speaking, a substantive evidential favoring relation cannot supervene on the *general logical relations* among evidence and hypotheses. This is a rough formulation because I don't yet have a precise account of what I mean by "general logical relations." The idea is that the following would all count as general logical relations among our three relata:

- The first hypothesis entails the second.
- The evidence refutes the first hypothesis but not the second.
- The first hypothesis is a tautology, while the second is not.
- The hypotheses are not mutually exclusive.

while this would not:

- The first hypothesis entails "Grass is green" while the second does not.

This last relation is in a sense a logical relation, but it is not a *general* logical relation. It relates the hypotheses to a particular proposition that we have singled out for our own extraneous purposes.

To make this supervenience interpretation work, we would need an account of general logical relations and then (hopefully) a proof that began with our general result and showed that a substantive evidential favoring relation could not supervene on general logical relations. The best I can do right now is to suggest a place to start: perhaps the general logical relations are those that can be expressed in a metalanguage that contains quantifiers and variables ranging over object-language sentences but whose only constants are names for our three relata and whose only relational symbol is one for object-language syntactical consequence.

3. Why Natural Properties Won't Help

To some, all of this will be old hat. We've understood this problem for a long time (they'll say), and we already know what the solution is. There *are* in fact special properties; following Lewis (1983) (and borrowing the relevant adjective from Quine (1969)) we can call them the "natural properties." These natural properties have a special metaphysical status and play a distinctive role in evidential favoring. For example, when some evidence indicates that all the observed objects have displayed a particular natural property, that evidence favors the hypothesis that the next object will display that property over the hypothesis that it won't. In Goodman's terminology, the natural properties are "projectible."

I can't deny that this is possible. However, it does create a problem. Presumably the contents of the natural properties list is an *a posteriori* fact. (I'll consider the possibility that it's not in Section 4.) How can we go about determining the contents of that list?

Consider an agent (call him Pedro) who wants to determine whether his total evidence favors one particular hypothesis over another. Pedro has absorbed the lessons of our general result, but believes that the evidential favoring relation works with a combination of one's total evidence and the list of natural properties. Still, he knows he won't be able to draw any conclusions from his evidence until he has the list of natural properties. So he sets out to determine the list of natural properties. Since the contents of that list is an *a posteriori* fact, he decides to determine the list using his total evidence. But he knows he can't draw any conclusions from his evidence until he has the list of natural properties. . . .[15]

Perhaps there is some way out of Pedro's predicament—perhaps there is some bootstrapping procedure by which we can use our total evidence to give us both a list of natural properties *and* the evidential favoring relation generated by that list. Unfortunately, we can use our general result to show that this is impossible.

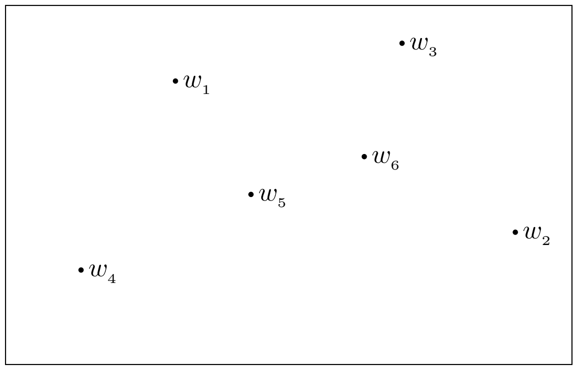
Let's imagine that given a pair of hypotheses and some evidence, our bootstrapping procedure proceeds in three steps. First, it uses the evidence to generate a list of natural properties—call that list $np(e)$. Second, it determines the evidential favoring relation to which that list of natural properties gives rise. We'll call that $\overset{np(e)}{\underset{e}{\succ}}$. Finally, it determines which hypothesis is favored by the evidence according to that favoring relation. In other words, it determines whether $h_1 \overset{np(e)}{\underset{e}{\succ}} h_2$.

The trouble is, $\overset{np(e)}{\underset{e}{\succ}}$ is still a favoring relation and so is susceptible to our general result. Assuming independent properties are always available, if $\overset{np(e)}{\underset{e}{\succ}}$ is substantive it must play favorites among properties. If $\overset{np(e)}{\underset{e}{\succ}}$ is substantive, some step in the process that generates it will handle some properties differently than it handles others; identically-structured bodies of evidence will be treated differently either in the creation of the natural properties list or in the determination of an evidential favoring relation from that list. To yield a substantive relation, the process that generates $\overset{np(e)}{\underset{e}{\succ}}$ will have to have a propensity to treat some properties different than others antecedent to the evidence's playing any role.

I've described the bootstrapping process as if it works sequentially, first determining the natural properties list from the evidence and then determining a favoring relation from that. But this was simply an illustrative device; it makes no difference if the process works "all at once," taking evidence and a pair of hypotheses and yielding a favoring judgment in one fell swoop. However the process works—whether it generates the simplest possible theory consistent with the evidence and evaluates the hypotheses according to that; whether it infers the best possible explanation of the evidence and goes from there; whatever subtle process from epistemology, statistics, the philosophy of science, or wherever else it employs—ultimately the

[15]It's consistent with our general result that conclusions can be drawn from evidence without the list of natural properties when those conclusions are not logically independent of the evidence—for example, when they are entailed by the evidence. But the contents of the natural properties list is not *entailed* by our evidence; the list of natural properties has to be *drawn out* from our observations of the natural world.

Figure 1



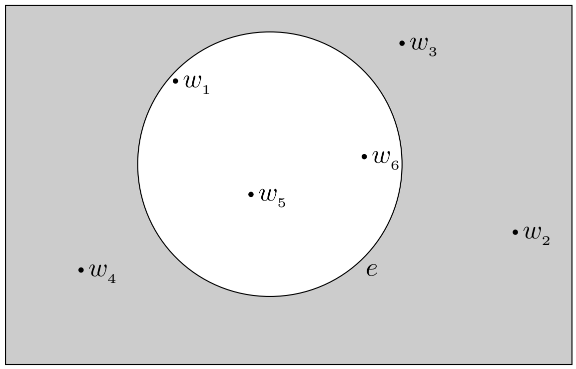
process will have to take a pair of hypotheses and some evidence as inputs and produce a favoring judgment as output. And here the *generality* of our result bites: the net effect of that process, captured as a relation $\overset{np(e)}{\underset{e}{\succ}}$, will satisfy our conditions and so be susceptible to our theorem. If favoring is substantive, the process will have to play favorites among properties, and therefore involve a list of special properties supplied by something beyond the evidence. But that was the role the natural properties were supposed to fill. To play its role in creating an evidential favoring relation, the list of natural properties cannot be determined from our evidence.

Proponents of a three-place evidential favoring relation often advise an agent to be guided solely by his evidence in choosing which hypotheses to believe. Our general result tells us that a body of evidence cannot substantively favor one hypothesis over another without the help of an additional element (a list of natural properties, a preferred language, etc.). That additional element cannot be determined by the evidence; it needs to be supplied from outside the agent's evidence, by something else entirely. Put another way: an evidential proposition does not have enough information content in itself to favor one hypothesis over another. To paraphrase Gertrude Stein, there's not enough there there.[16]

I'll conclude this section by illustrating the point with a diagram. Figure 1 depicts the space of all possible worlds as a rectangle; I've marked some of those possible worlds with names for convenience. Figure 2 shows the effects of a particular evidential proposition, depicted as a circle. What the evidence does is rule out some possible worlds (those outside the circle) while leaving others in contention. The crucial thing we learn from our general result is that that's *all* the evidence can do. The evidence does not first rule some of the possible worlds out and then

[16]There are by now a number of formal measures of the information content of a proposition, and of course a body of evidence can have an arbitrarily large information content as calculated by one of those measures. When I say that a body of evidence does not have enough information content to favor one hypothesis over another, I mean roughly what a judge means when he says, "There's not enough here for an indictment." It does no good for the prosecutor to respond, "But Your Honor, I've supplied over a thousand pages of evidence, each of which contains multiple logically independent facts!"

Figure 2



rank those that remain within the circle. To the extent that some of the remaining possibilities bear a special relation to others once the evidence arrives, that is only because they bore a special relation before the evidence was introduced. To continue the geometric metaphor, we might think that when the circle appears the remaining possibilities that are closer to its center (such as w_5) are somehow rendered more probable than the ones near the edge (like w_1). But proximity to the edge of the circle is determined not just by the circle, but also by the geometric arrangement of the worlds before the circle was introduced. Shuffle the original arrangement, and this "favoring" relation would be different entirely.

4. Responses and Alternatives

To this point, we have argued first that if there is a substantive three-place evidential favoring relation, it must be determined by something like a list of special properties, and second that such a list of special properties cannot be uncovered *a posteriori*. I now want to consider two responses on behalf of the *a posteriori* natural properties view, and then two alternatives to it.

First, it might be suggested that Pedro's predicament can be solved by a feature David Lewis attributed to natural properties: their so-called "reference magnetism."[17] Natural properties played a role not only Lewis's metaphysics, but also in his philosophy of language. Briefly, Lewis thought that among the considerations that determine the meanings of terms in a language is a preference for assigning natural properties as the extensions of predicates. The thought would then be that given reference magnetism, a predicate like "green" is likely to pick out a natural property, and since part of the metaphysical role of natural properties is to play a part in physical laws, we can be confident that "green" will be projectible as well.

I've explained Lewis's view very quickly, without taking into account complexities of his position such as the suggestion that naturalness may come in degrees. Still, I think that even if we grant Lewis both his metaphysics of natural properties and his philosophy of language, and build in all the complexities he proposes,

[17] I am grateful to Andrew McGonigal and Jonathan Schaffer for discussing this suggestion with me.

reference magnetism will not solve our epistemic problem.[18] There are two basic difficulties. The first is that a speaker of a language needs some way to determine which of that language's predicates are going to be assigned the more natural properties. Suppose Nigel speaks a language we'll call ENGLISH, that sounds exactly like English and is even spelled identically except that all the letters are capitalized. Nigel needs to figure out whether "GREEN" or "GRUE" is assigned a more natural property under the theory of reference magnetism. According to Lewis, Nigel can determine this by seeing which is more simply definable in the terms employed by the best-supported scientific theory. The trouble is, Nigel then needs to determine which of the scientific theories currently available is best supported by his (or his society's) total evidence. And according to our general result, in order to do *that* he first has to know which predicates in his language pick out natural properties.

The second difficulty is a bit more subtle. Suppose that Nigel somehow manages to figure out that "GREEN" picks out a more natural property than "GRUE." How does he know whether that natural property is greenness or grueness? That is, how can he figure out if the next GREEN object he observes will look like previous GREEN objects or like previous BLUE objects?[19] This sort of question can have tremendous practical import. For instance, Nigel may have managed to project that if he steps out in front of a bus tomorrow he will wind up DEAD. But he will also want to know whether being DEAD tomorrow is more like being DEAD today or being ALIVE today—presumably he's trying to avoid a condition like the former!

There is a second response available to the natural properties theorist—call it the "hard externalist" response. The hard externalist maintains that there is an objective evidential favoring relation determined by a metaphysically privileged set of natural properties. (This view may be combined with reference magnetism or not.) He grants my argument that we cannot discern the list of natural properties from our evidence, and grants therefore that we cannot determine when the favoring relation holds between two hypotheses relative to some evidence. But he maintains that despite all that, there are still facts of the matter about what favors what and about what each of us ought to believe given our total evidence; it's just that these facts are *inaccessible* to us.

This strikes me as a seriously unattractive position. Evidential favoring is supposed to play a crucial role in *guiding* our theoretical deliberations (and thereby indirectly portions of our practical deliberations as well). It helps us figure out what we should conclude from our evidence, and grounds judgments we make correctly all the time about whether others have drawn appropriate conclusions from their

---

[18]Here I think it's significant that despite the fact that Lewis published multiple articles describing the philosophical "work" he thought natural properties could accomplish (such as (Lewis 1983) and (Lewis 1984)), he never once suggested they would solve Goodman's projectibility problem, even though green/grue was one of his favorite examples of a natural/unnatural property contrast. Multiple philosophers who studied with Lewis have suggested to me that he might well have agreed with this essay's main claims.

[19]Of course, he knows that future GREEN objects will look LIKE previous GREEN objects, but that isn't the question here. Notice that these sorts of questions are much easier to articulate when talking about someone else's language than when talking about our own. van Fraassen (1997) uses this fact to argue that "Putnam's Paradox"—the original problem to which Lewis's reference magnetism theory was meant to be a solution—can be dissolved without the use of natural properties.

own evidence. If the favoring relation is in principle inaccessible to us, I can't see how it could play these guiding roles.[20]

As these difficulties mount for an *a posteriori* natural properties view, the obvious alternative is that the list of natural properties (or special properties in some other sense) can be determined *a priori*.[21] This alternative is not so implausible; after all, it seems intuitive that there is something wrong with predicates like "grue" and that such predicates cannot be projected from past observations onto future predictions. Perhaps our intuition enables us to rule out gruesome predicates as unnatural; or perhaps we can just perceive that sharing particular features make some objects more similar than others; or perhaps the light of reason reveals to us that some possible arrangements of the universe are more uniform and therefore more likely than others.

I can't definitively rule out this alternative. However, I want to make clear how difficult it would be to carry through as a viable philosophical theory and what obstacles would have to be surmounted along the way. First, while it may seem that our intuition immediately rejects "disjunctive" or "logically complex" predicates like "grue,"[22] our sea creatures example has already shown that logically complex predicates may sometimes be perfectly acceptable. Having symmetrical fins, the property that language $\mathcal{L}^{LS}$ represented with predicate S, may be defined as "has a left fin just in case it has a right fin." Yet scientists trade in this sort of predicate all the time, and as we saw above language $\mathcal{L}^{LS}$ may do a better job of revealing evidential structure than the language $\mathcal{L}^{LR}$ whose predicates are "logically simple."

It's sometimes suggested that intuition rejects logically complex predicates whose definition involves a spatio-temporal property in a disjunctive way. Yet for many centuries the dominant Aristotelian physics held that the natural (*phusei*) motion of an inanimate object was linear if it was in the sublunar realm or circular if the object dwelt in the heavens. Some who take the Bible literally hold that "snake" applied to legged creatures before the Fall but has applied to legless creatures ever since. The predicate "at home" applied for me to a particular location during my first year of graduate school, then to a different location for the next few years, then finally to a third location near the end of my studies. Yet it was highly projectible that at the end of each day I could be found at home.

Moreover, the property picked out by "grue" isn't really disjunctive. Assuming (as Lewis does) that for every set of objects there exists a property of belonging to that set, "grue" picks out a set that is no more or less "disjunctive" than any other set of objects. What's disjunctive is the definition of "grue" in terms of our more standard color predicates, but that's only relevant if those predicates are already privileged for other reasons. Even more importantly, ruling out "grue" and its other quirky kin is not really what the project of explaining *a priori* property prioritization should be about. "Grue" was a counterexample constructed by Goodman to demonstrate in an expedient fashion that the formal evidential theories available in his day lacked a particular general feature. Some philosophers respond to

---

[20] Here it's instructive to compare the hard externalist view with a "factualist" view on which one hypothesis is favored over another just in case the former is true and the latter false.

[21] I am grateful to David Chalmers for discussion of this alternative.

[22] If we are observing objects that may only be green or blue, "x is grue" can be defined as "x is green just in case x has been observed already."

Goodman's puzzle (usually in private) by saying that we need not worry, because in real life there's no threat that any of us is going to try to project a predicate like "grue." But that's like responding to Gödel's First Incompleteness Theorem by saying that we weren't really worried about whether things like the Gödel sentence could be proved within our formal systems. In both cases, the point of the admittedly odd counterexample is that it exemplifies the lack of a general feature we might have otherwise assumed obtained. In the case at hand, predicates like "grue" demonstrate that a three-place favoring relation requires a list of preferred properties and that that list cannot be determined from our evidence. This opens up a broad challenge of explaining how any property does or doesn't get placed on the list on an *a priori* basis, not just how properties like grueness can be excluded.[23]

And this is a serious challenge, whatever one's theory of the *a priori*. A list of preferred properties is not going to be achieved simply through analysis of the concepts "evidence," "similarity," or "uniformity." The *a priori* theorist needs to explain not only the mechanism by which we assemble a special properties list, but also why that mechanism has epistemic validity. After all, explaining how and why we intuitively take some properties to be more natural than others does not explain whether or why our intuition is thereby latching on to properties that are projectible.

For example, it might be suggested that some objects just obviously go together phenomenologically, and that the favoring relation should be analyzed using a language whose predicates capture that phenomenological grouping. In response to this suggestion, we should first note that it's unclear whether any real human languages respect phenomenology all that closely—every evening as the light fades green objects come to look very different than they did during the day. Second, many features of our phenomenology are determined by contingent facts about our sensory faculties. The way humans group objects by color has a great deal to do with particular facts about our retinas and visual processing units, including the fact that we cannot perceive most wavelengths of light at all! To take another example, the predicate "odorless" might for a great deal of human history have seemed intuitively to express a natural, projectible property. But the set of chemicals that can be detected by human olfaction proves to be a hodgepodge grouping with almost no underlying scientific unity.

Now it may simply be a nonstarter to try to understand "true" similarity in terms of human phenomenology. But even when we turn to apparently *a priori* similarity judgments made by our higher rational faculties, we should remember that many of these have changed over the course of human history. Nowadays it would seem silly to group two objects as belonging to a fundamental kind on the grounds that they occupy some particular spatial or temporal region. But a good explanation of Aristotle's "disjunctive" theory of natural motion for inanimate objects is that he thought each kind of object had a proper motion and that objects farther from Earth than the Moon were of a different *kind* than objects closer in.[24] Scientists (such as Newton in the Western tradition) had to *discover* that the same physical laws apply throughout the universe, and that spatio-temporal location is irrelevant to grouping into physical kinds.

[23]Compare Goodman's discussion at (1979, p. 80).

[24]See (Aristotle 1984, Bk. 1, Ch. 3). I am grateful to Jessica Gelber for discussion of this point and help with the reference.

Again, my point is not that it is *impossible* that we have the ability to discern *a priori* the list of special properties that underwrites the evidential favoring relation. It's just that anyone who believes there is an objective fact of the matter about which hypotheses are favored by which evidence is thereby committed to a very strong conception of the *a priori*. After all, the proposal is that given two situations with *identical* logical form we can tell *a priori* that one provides evidential favoring in a particular direction and the other does not. I think that philosophers often interpret Goodman's problem as showing that there's something difficult and complex we do with our evidence, and that it's very mysterious how we manage to do it. The thought is that by some subtle and abstract yet still perfectly general procedures, we manage to draw out (at least reasonably well) from our evidence which are the projectible properties. Yet the current result shows that this isn't what we do at all; if we do somehow determine the list of projectible properties, we do it not by relying on our evidence but by some process that is capable of ranking particular properties over others entirely prior to that evidence.[25]

There is a tempting fallback position available to the philosopher who thinks that our intuitions recommend some properties as being more natural than others. That is to maintain not that our intuitions grab onto some objective truth about a list of special properties that underlies the evidential favoring relation, but instead that those intuitions give us a starting point and provide us with warrant for reasoning in particular ways on the basis of our evidence as we work to improve our understanding of the world. For example, it might be suggested that in their scientific investigation of the world humans were initially justified in grouping objects by phenomenological similarity, and that our subsequent investigations have altered those groupings over time.

But to go in this direction is to adopt a different alternative altogether. Really, it is to drop the idea of a three-place favoring relation and admit that an agent's theoretical reasoning is normatively governed by favoring relations that are relative not only to his evidence but also to his current personal outlook on the world. That personal outlook may depend on his society, his upbringing, his biology, his native language, or any of a number of other contingent factors. In the course of a particular inquiry, personal outlook may bring certain pragmatic interests to bear, may highlight certain questions as more important than others, or may prioritize the ruling out of certain alternatives. At any given time an agent's personal outlook must conform to particular constraints—say, constraints of internal consistency—and if it changes over time in response to new evidence there will be constraints on how it may do so. But the crucial point is that which hypothesis an agent's evidence favors at a given time will always depend at least in part on *subjective* factors about the agent, factors that are not entirely determined by the agent's total evidence.

The relevance of such subjective factors is sometimes tacitly acknowledged by proponents of various formal theories of evidence. For example, philosophers who

---

[25]In calling our natural property-detection procedures "general," I mean to suggest that they work at a level of abstraction from which all properties look the same, not that they are necessarily capturable using general *principles*. Suppose you're a particularist who holds that agents determine the list of natural properties from their evidence using a faculty of judgment whose operations are not expressible in principles. Our general result shows that this judgment faculty must have a disposition to treat some properties differently than others antecedent to its encountering any evidence.

urge scientists to employ tools like maximum entropy in ranking hypotheses will respond to complaints that those tools are language dependent by saying that the scientists should just use "whatever scientific language they are working with." On the present view that is entirely right, but one needs to recognize that which language a scientist is working with is not a matter entirely determined by past experiments. Moreover, one needs to understand the costs of a view that relativizes evidential favoring to a fourth, subjective relatum. For instance, one consequence of this view may be that two agents with the same total evidence can be rationally permitted to have different beliefs.[26] In fact, if their outlooks differ enough they may be rationally *required* to have different beliefs.[27]

This subjectivist position may prompt us to rethink our view of scientific reasoning.[28] Carnap thought of scientific reasoning as being like an argument in the sense of "argument" we teach in deductive logic classes: flowing from a set of premises to a conclusion. He imagined that a complete scientific novice with substantial powers of reasoning could be given the results of all the experiments conducted up to the present and discern from them which scientific theory was favored. But the present position suggests that we think of science as being more like an "argument" in the vernacular sense: an activity that carries on through time, over the course of which positions change, information is introduced, and new views develop. Understanding science that way, our total evidence at a time doesn't tell us where we should *be*; instead, each piece of evidence as it accumulates tells us, given where we are *now*, where we should go *next*.

Of course, the subjective element in this view—what I've been calling an agent's "personal outlook"—need not be a language or a list of preferred properties. We can, for instance, obtain a substantive evidential favoring relation by supplementing a total evidence set with a probability distribution. The view known as "Subjective Bayesianism" models an agent as assigning an initial numerical distribution of credences over propositions (called a "prior") that is not determined by his evidence. This distribution is subject to particular constraints such as the laws of probability. The view then requires the agent to modify his credence distribution over time according to certain rules as more evidence comes in. At any given time, favoring relations are determined by the agent's current credences. The constraints on credences supervene on general logical relations among evidence and hypotheses, and for any given prior distribution they guarantee such features as the equivalence condition and language invariance. Whether "green" or "grue" is more projectible then becomes a question of what sorts of credence distributions allow green evidence to favor green hypotheses.[29]

---

[26]The view may thus deny what White (2005) calls the Uniqueness Thesis: "Given one's total evidence, there is a unique rational doxastic attitude that one can take to any proposition." (White, in turn, attributes the thesis to (Feldman 2007).)

[27]"Subjective" is therefore being used in the present discussion in both the vernacular sense (on which something is "subjective" if it varies from person to person) and the more traditional philosophical sense (on which something is "subjective" if it depends in part on features of the analyzer/perceiver and not just on features of the object being analyzed/perceived).

[28]I am grateful for a conversation with Kevin Kelly on this point.

[29](Fitelson and Hawthorne ms) pursues this sort of question for the Paradox of the Ravens. Their work suggests the kind of thing a Subjective Bayesian should do to respond to the grue problem.

Adopting the subjectivist position may require a further shift in our philosophical outlook. In certain areas of philosophy it has recently become popular to think of norms as originating in facts about reasons: what an agent has *reason* to believe given his evidence, what he has *reason* to do given his evidence and his intentions, etc.[30] This parallels a recent shift towards thinking of normativity in terms of objective rather than subjective "oughts."[31] Yet if our argument is correct, at least when it comes to theoretical reason it is very difficult to maintain that there are any objective facts of the matter about what an agent's total evidence gives him reason to believe.[32] Instead, our story about what an agent ought to believe involves rationality constraints explaining how his current subjective position should generate a response to new evidence. The norms of Subjective Bayesianism, for instance, certainly seem to be best understood as norms of internal consistency, which are paradigmatic norms of rationality.

As we mentioned in the discussion of reference magnetism, all of this is perfectly compatible with a *metaphysics* that asserts the existence of special, "natural" properties. What it challenges is an *epistemology* on which those natural properties give rise to an objective evidential favoring relation that dictates what we ought to believe. One might maintain that an objective favoring relation exists and the subjective rationality constraints I have been discussing are derived from it or its general features, in the way it is sometimes suggested that the injunction against inconsistent beliefs derives from the fact that the set of true propositions is logically consistent. On this view the subjective norms provide a guide as we struggle to discern the truth about the natural properties and thereby learn what we objectively have reason to believe. The view must concede that while we're still in the dark, rationality will sometimes require us to believe what we have objective reason not to believe. But some talk about an agent's doing what seems objectively required "by his lights" at the time might manage this bit of awkwardness.

Now it may be the case—in fact, I hope it's the case—that obeying rationality constraints has various epistemic benefits such as helping us discover true facts about the universe over time. (Various Subjective Bayesian theorems showing that in the long run differences in priors "wash out" as rational agents converge on the truth are meant to demonstrate exactly this type of feature.) But I think we should take the work of Kolodny (in his (2005), (2008), and elsewhere) as a cautionary tale about the difficulties of trying to derive subjective norms from objective. In the meantime, our argument shows that absent a very strong theory of the *a priori*, everyday theoretical inquiry is guided by evidential relations that supervene on elements of an agent's subjective personal outlook. The resulting subjective constraints are real, are normative, and are the only things that can

---

[30](Scanlon 1998) is a prominent recent example of this reasons-based approach to normativity; the first chapter begins, "I will take the idea of a reason as primitive." Contemporary reasons theorists owe a great deal to the work of Joseph Raz, who rekindled the tradition in the 1970s.

[31]To use the example made famous by (Williams 1981), an agent faced with a glass of petrol that he thinks is a glass of gin subjectively ought to drink what's in the glass (or at least is permitted to drink what's in the glass) while objectively ought not.

[32]If evidential favoring can vary from person to person, why does our reasons *talk* often sound so objective? One possibility is that in most conversations the agents have so much in common among their personal outlooks (or priors) that a wide variety of favoring relations are effectively objective among them.

guide us given our lack of access to the list of natural properties. The burden is on the objective oughts theorist to explain where objective reasons fit into this picture.

We might think of the twentieth-century Objective Bayesians who tried to give recipes for calculating probability distributions licensed by particular bodies of evidence as trying to work out the details of a view on which there is an objective fact of the matter about what any agent has reason to believe given his evidence. Understood that way, I don't think the dramatic shift towards Subjective Bayesianism among philosophers of science in the late twentieth century was a coincidence. I think that as they worked through the details of the objectivist project, Bayesians came to sense that there simply isn't enough information in an evidential proposition to support something as strong as a probability distribution over a hypothesis space.[33] The present essay began in my attempts to determine whether there might be an objective way to achieve something weaker, like a partial ordering over the hypotheses. It turns out there isn't enough information in a body of evidence to do even that.

A key has no logic to its shape. Its logic is: it turns the lock.
—G.K. Chesterton

References

Aristotle (1984). *De Caelo* (On the heavens). In J. Barnes (Ed.), *The Complete Works of Aristotle*, Volume I, pp. 447–511. Princeton: Princeton University Press.

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

Carnap, R. (1952). *The Continuum of Inductive Methods*. Chicago: University of Chicago Press.

Feldman, R. (2007). Reasonable religious disagreements. In L. M. Antony (Ed.), *Philosophers without Gods: Meditations on Atheism and the Secular Life*, pp. 194–214. Oxford: Oxford University Press.

Fitelson, B. (2006). Logical foundations of evidential support. *Philosophy of Science 73*, 500–512.

Fitelson, B. and J. Hawthorne (ms.). How Bayesian confirmation theory handles the paradox of the ravens. In E. Eells and J. H. Fetzer (Eds.), *Probability in Science*. Open Court.

Goodman, N. (1946). A query on confirmation. *The Journal of Philosophy 43*, 383–385.

Goodman, N. (1979). *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.

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

Jaynes, E. T. (1957a). Information theory and statistical mechanics I. *Physical Review 106*, 620–630.

Jaynes, E. T. (1957b). Information theory and statistical mechanics II. *Physical Review 108*, 171–190.

Kolodny, N. (2005). Why be rational? *Mind 114*, 509–563.

Kolodny, N. (2008). Why be disposed to be coherent? *Ethics 118*, 437–463.

[33] As a number of people I know like to say, if Carnap couldn't make it work then we aren't going to either.

Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy 61*, 343–377.

Lewis, D. (1984). Putnam's paradox. *Australasian Journal of Philosophy 62*, 221–236.

Miller, D. (2005). *Out of Error: Further Essays on Critical Rationalism.* Burlington, VT: Ashgate.

Niiniluoto, I. (1987). *Truthlikeness.* Dordrecht: D. Reidel Publishing Company.

Quine, W. (1969). Natural kinds. In *Ontological Relativity, and Other Essays*, pp. 114–138. New York: Columbia University Press.

Scanlon, T. (1998). *What We Owe to Each Other.* Cambridge, MA: Harvard University Press.

Seidenfeld, T. (1986). Entropy and uncertainty. *Philosophy of Science 53*, 467–491.

van Fraassen, B. C. (1997). Putnam's paradox: Metaphysical realism revamped and evaded. *Noûs 31*, 17–42.

White, R. (2005). Epistemic permissiveness. *Philosophical Perspectives 19*, 445–459.

Williams, B. (1981). Internal and external reasons. In *Moral Luck*, pp. 101–113. Cambridge: Cambridge University Press.

Williamson, T. (2000). *Knowledge and Its Limits.* Oxford: Oxford University Press.

Appendix A. Proof of the General Result

I'll begin by re-stating the definitions and conditions needed for our result, in some cases expressing them more precisely than in Section 1. As in Section 1, I will **bold** a term when it is introduced in this discussion.

A **language** $\mathcal{L}$ will be a first-order language with no variables or quantifiers. $P(\mathcal{L})$ will be the number of predicates in $\mathcal{L}$; $N(\mathcal{L})$ will be the number of names. We will limit ourselves to languages with finite $P(\mathcal{L})$ and $N(\mathcal{L})$ values. An **atomic sentence** applies a predicate to a name; a **literal** is an atomic sentence or its negation; a **state description** is a conjunction of literals in which each predicate is applied to each name exactly once. The number of atomic sentences in $\mathcal{L}$ is $P(\mathcal{L})\cdot N(\mathcal{L})$; this is also the number of conjuncts in a state description. We will imagine that each language $\mathcal{L}$ comes with a total ordering on its atomic sentences (not too difficult given that the number of atomic sentences is finite); a state description whose conjuncts appear in order is a **canonical** state description. We will also imagine that $\mathcal{L}$ comes with a total ordering on its canonical state descriptions (again not difficult because there are finitely many of them); a sentence $x \in \mathcal{L}$ that is a disjunction of canonical state descriptions in which no state description appears more than once and all the state descriptions appear in order will be said to be in **disjunctive normal form**. A standard result in sentential logic demonstrates that any non-contradictory sentence in $\mathcal{L}$ has a unique equivalent in disjunctive normal form.

Each language also comes with an interpretation that maps its sentences onto propositions. I will write μx for the proposition represented by sentence $x \in \mathcal{L}$, μF for the property represented by predicate F, and μa for the object represented by name a.[34] Given languages $\mathcal{L}$ and $\mathcal{L}'$ with sentence $x \in \mathcal{L}$ and sentence $x' \in \mathcal{L}'$, I will say that x' is a **synonym** of x just in case $\mu x = \mu x'$. If every sentence $x \in \mathcal{L}$ has a synonym in $\mathcal{L}'$, I will say that $\mathcal{L}'$ **expresses all of** $\mathcal{L}$.

Each language also comes with a syntactic consequence relation. Language $\mathcal{L}$ is **faithful** just in case for any $x, y \in \mathcal{L}$, $x \vdash y$ if and only if $\mu x \vDash \mu y$. We will assume that $\vDash$ is the classical entailment relation and that $\vdash$ interacts with the sentential connectives in the way standardly taught in introductory logic classes.

We can now precisely state our condition on the plenitude of properties:

> **Availability of Independent Properties:** Given any faithful language $\mathcal{L}$, there exists another faithful language $\mathcal{L}'$ such that: $\mathcal{L}'$ has the same names as $\mathcal{L}$ representing the same objects, $P(\mathcal{L}') = P(\mathcal{L})+1$, and for each predicate $F_i \in \mathcal{L}$ there exists a predicate $F'_j \in \mathcal{L}'$ such that $\mu F'_j = \mu F_i$.

As we discussed in Section 1, the predicates of faithful languages express logically independent properties, so the extra predicate in $\mathcal{L}'$ expresses a property logically independent of all the properties expressed in $\mathcal{L}$.

When we work with some evidence and a pair of hypotheses (each understood as a proposition), these three relata will be represented by the sentences e, h_1, and h_2 (respectively) in a language $\mathcal{L}$. As discussed in Section 1, given a particular pair of hypotheses and evidence we can divide the set of languages into those that are **adequate** for these relata and those that are not. (Again, when the relata under discussion are clear I will sometimes simply refer to adequate or inadequate languages.) I have no formal definition of adequacy, but the following two conditions are assumed to hold:

(1) Faithfulness is necessary for adequacy.
(2) If language $\mathcal{L}$ is adequate for a particular set of relata and faithful language $\mathcal{L}'$ expresses all of $\mathcal{L}$, $\mathcal{L}'$ is adequate for those relata as well.

Result 1. Suppose we have a pair of hypotheses, some evidence, and an adequate language $\mathcal{L}$. Then there exists a language $\mathcal{L}'$ adequate for that evidence and pair of hypotheses with $N(\mathcal{L}') = 1$.

Proof. Suppose the predicates of $\mathcal{L}$ are $F_1, F_2, \ldots, F_{P(\mathcal{L})}$ and the names of $\mathcal{L}$ are $a_1, a_2, \ldots, a_{N(\mathcal{L})}$. Language $\mathcal{L}'$ is constructed so as to have one name, a', such that $\mu a'$ is the $N(\mathcal{L})$-tuple $(\mu a_1, \mu a_2, \ldots, \mu a_{N(\mathcal{L})})$. The predicate letters of $\mathcal{L}'$ are double-indexed (e.g. $F'_{3,4}$) with the first index running from 1 through $P(\mathcal{L})$ and the second index running from 1 through $N(\mathcal{L})$. We assign $\mu F'_{j,k} a' = \mu F_j a_k$ for every $1 \leqslant j \leqslant P(\mathcal{L})$ and $1 \leqslant k \leqslant N(\mathcal{L})$.

We have now identified a synonym in $\mathcal{L}'$ for every atomic sentence of $\mathcal{L}$. Given any other sentence $x \in \mathcal{L}$, we construct its synonym in $\mathcal{L}'$ by replacing each of the atomic sentences in x with its synonym in $\mathcal{L}'$. Clearly $\mathcal{L}'$ expresses all of $\mathcal{L}$.[35] □

We now introduce the relation $\succ_e$. $h_1 \succ_e h_2$ holds just in case the proposition represented by h_1 is favored over the proposition represented by h_2 relative to the proposition represented by e.

---

[34] Suppose we have two languages $\mathcal{L}$ and $\mathcal{L}'$ containing the predicates F and F' respectively. When I write $\mu F = \mu F'$, I mean that for any names $a \in \mathcal{L}$ and $a' \in \mathcal{L}'$ such that $\mu a = \mu a'$, $\mu F a = \mu F' a'$.

[35] [Note: Here I need to fill in a proof of the faithfulness of $\mathcal{L}'$, thereby demonstrating that $\mathcal{L}'$ is adequate for the hypotheses and evidence.]

Result 2. $\succ_e$ is **language invariant**. That is, given any $h_1, h_2, e \in \mathcal{L}$ and $h'_1, h'_2, e' \in \mathcal{L}'$ such that $\mu h_1 = \mu h'_1$, $\mu h_2 = \mu h'_2$, and $\mu e = \mu e'$, $h_1 \succ_e h_2$ just in case $h'_1 \succ_{e'} h'_2$.[36]

Proof. By definition, $h_1 \succ_e h_2$ just in case μh_1 is favored over μh_2 relative to μe. Since $\mu h_1 = \mu h'_1$, $\mu h_2 = \mu h'_2$, and $\mu e = \mu e'$, this occurs just in case $\mu h'_1$ is favored over $\mu h'_2$ relative to $\mu e'$. By definition, this in turn occurs just in case $h'_1 \succ_{e'} h'_2$. □

Result 3. $\succ_e$ satisfies the **equivalence condition**. That is, for any faithful $\mathcal{L}$ and $e, h_1, h_2, e', h'_1, h'_2 \in \mathcal{L}$ such that $h_1 \dashv\vdash h'_1$, $h_2 \dashv\vdash h'_2$, and $e \dashv\vdash e'$, $h_1 \succ_e h_2$ just in case $h'_1 \succ_{e'} h'_2$.

Proof. Since $\mathcal{L}$ is faithful, $h_1 \dashv\vdash h'_1$ entails $\mu h_1 \;=\!\!\!|\!\!\models \mu h'_1$. Understanding propositions as sets of possible worlds, this entails that $\mu h_1 = \mu h'_1$. Similarly, $\mu h_2 = \mu h'_2$ and $\mu e = \mu e'$. Thus μh_1 is favored over μh_2 relative to μe just in case $\mu h'_1$ is favored over $\mu h'_2$ relative to $\mu e'$, so $h_1 \succ_e h_2$ just in case $h'_1 \succ_{e'} h'_2$. □

We next introduce a sequence of conditions on $\succ_e$. First, given any language $\mathcal{L}$ and any sentence $e \in \mathcal{L}$, $\succ_e$ induces a **strict ordering** on the sentences of $\mathcal{L}$. That is, $\succ_e$ is transitive and antisymmetric.[37] We do not require that $\succ_e$ induce a total ordering on the sentences of $\mathcal{L}$.

Our other two conditions on $\succ_e$ are:

Substantivity: There exists some language $\mathcal{L}$ adequate for the propositions expressed by $e, h_1, h_2 \in \mathcal{L}$ such that $h_1 \succ_e h_2$ but e, h_1, and h_2 are logically independent.

Identical Treatment of Predicate Permuations: Suppose we have a language $\mathcal{L}$ adequate for $e, h_1, h_2 \in \mathcal{L}$, and a permutation π of the predicates of $\mathcal{L}$. If $\pi(e) = e'$, $\pi(h_1) = h'_1$, and $\pi(h_2) = h'_2$, then $h_1 \succ_e h_2$ just in case $h'_1 \succ_{e'} h'_2$.

This last condition gives us:

Result 4. Suppose $\succ_e$ treats predicate permutations identically. Then for any adequate language $\mathcal{L}$ and $e, h_1, h_2 \in \mathcal{L}$, if there exists a permutation π of the predicates of $\mathcal{L}$ such that $\pi(e) = e$, $\pi(h_1) = h_2$, and $\pi(h_2) = h_1$, then $h_1 \not\succ_e h_2$.

Proof. Suppose $\succ_e$ treats predicate permutations identically, and suppose we have adequate language $\mathcal{L}$, $e, h_1, h_2 \in \mathcal{L}$, and a permutation π of the predicates of $\mathcal{L}$ such that $\pi(e) = e$, $\pi(h_1) = h_2$, and $\pi(h_2) = h_1$. Suppose further for *reductio* that $h_1 \succ_e h_2$. Since $\succ_e$ treats predicate permutations identically, we would also have $h_2 \succ_e h_1$. But that violates the antisymmetry of $\succ_e$, so we have a contradiction. □

We can now state our main result:

Result 5 (The General Result)**.** The availability of independent properties, substantivity, and the identical treatment of predicate permutations cannot all hold.

Overview of the Proof: The proof of our general result proceeds by *reductio*. We suppose that independent properties are available and that $\succ_e$ treats predicate permutations identically. We then suppose that (as substantivity requires) there

[36] In the statement of this result, the second sentence states precisely what I mean by the first. I hope the way of talking exemplified by the first sentence is a useful enough rough shorthand that it excuses my metalogically abusive use of "e" as a sort of variable there.

[37] We actually will not require the transitivity of $\succ_e$ for our proof.

exists an adequate language $\mathcal{L}$ with $e, h_1, h_2 \in \mathcal{L}$ such that $h_1 \succ_e h_2$ and e, h_1, and h_2 are logically independent. By the equivalence condition, we can assume without loss of generality that e, h_1, and h_2 are in disjunctive normal form. We also use Result 1 to make our lives easier by assuming (without loss of generality) that $\mathcal{L}$ has only one name.

The goal of the proof is to construct a language $\mathcal{L}^*$ containing sentences e^*, h_1^*, and h_2^* that express the same propositions as e, h_1, and h_2 respectively. By language invariance, $h_1^* \succ_{e^*} h_2^*$. Yet there is a permutation π of the predicates of $\mathcal{L}^*$ that maps h_1^* to h_2^*, h_2^* to h_1^*, and e^* to itself. $\mathcal{L}^*$ is adequate for e^*, h_1^*, and h_2^*, so by Result 4 we have $h_1^* \not\succ_{e^*} h_2^*$. This yields the desired contradiction.

e^*, h_1^*, and h_2^* are all in disjunctive normal form; π achieves the mappings we want by treating the disjuncts of these sentences in very particular ways. For example, a canonical state description that is a disjunct of e^* but of neither of the other two is mapped by π to itself. On the other hand, each canonical state description that is a disjunct of h_1^* but neither of the other two is mapped to a distinct canonical state description that is a disjunct of only h_2^* (and *vice versa*).

In order for this mapping scheme to work, some very specific relations have to hold between the numbers of disjuncts that are of particular types. For example, the number of h_1^*-only disjuncts has to match the number of h_2^*-only disjuncts. To achieve these numerical relationships, our proof works in two steps.

In Step 1, we construct a language $\mathcal{L}'$ from $\mathcal{L}$. $\mathcal{L}'$ contains disjunctive normal form synonyms e', h_1', and h_2' for e, h_1, and h_2 (respectively). More importantly, the numbers of disjuncts shared by e', h_1', and h_2' are exactly what they need to be for our mapping scheme to work. In Step 2, we then construct $\mathcal{L}^*$ from $\mathcal{L}'$ by taking each canonical state description that appears as a disjunct of e', h_1', or h_2' and giving it a synonym that is a canonical state description of $\mathcal{L}^*$. The numerical relationships we're interested in are preserved, and the synonyms in $\mathcal{L}^*$ are chosen so as to make a predicate permutation π achieving the desired mappings available.

How does the move from $\mathcal{L}$ to $\mathcal{L}'$ work? The logical independence of e, h_1, and h_2 guarantees that disjuncts of each type are available—so we know, for instance, that there will be at least one h_1-only disjunct and at least one h_2-only disjunct. But it still may be the case that, say, the number of h_2-only disjuncts exceeds the number of h_1-disjuncts by 2. We then give each h_2-only disjunct a synonym that is a canonical state description of $\mathcal{L}'$, but we give one of the h_1-only disjuncts a synonym that is the disjunction of *three* canonical state descriptions of $\mathcal{L}'$. (This means that $\mathcal{L}'$ must have more state descriptions than $\mathcal{L}$, but the availability of faithful languages larger than $\mathcal{L}$ is guaranteed by the availability of independent properties.)

Splitting one h_1-only disjunct into three h_1'-only disjuncts means that while there are exactly as many h_2'-only disjuncts as there are h_2-only disjuncts, there are 2 more h_1'-only disjuncts than there are h_1-only disjuncts. And this in turn means that there are exactly as many h_1'-only disjuncts as there are h_2'-only disjuncts. The sought-after numerical relationship has been achieved.

The full proof is presented below. It is entirely constructive; given a language $\mathcal{L}$ and an $e, h_1, h_2 \in \mathcal{L}$, it gives an algorithm for constructing a language $\mathcal{L}^*$ that witnesses the incompatibility of substantivity and the identical treatment of predicate permutations.

Proof of Result 5. Suppose for *reductio* that all three conditions hold. That is, suppose that independent properties are available and that $\succ_e$ treats predicate permutations identically. Suppose also that there exists a language $\mathcal{L}$ adequate for the propositions expressed by $e, h_1, h_2 \in \mathcal{L}$, that $h_1 \succ_e h_2$, and that e, h_1, and h_2 are logically independent. If $N(\mathcal{L}) > 1$, Result 1 allows us to construct a single-name language $\mathcal{L}'$ containing the synonyms e', h_1', and h_2' for e, h_1, and h_2 respectively. By language invariance we will have $h_1' \succ_{e'} h_2'$, and by the faithfulness of $\mathcal{L}'$ our three synonyms will be logically independent. So to simplify matters we will assume without loss of generality that the language $\mathcal{L}$ we are given has only one name. The equivalence condition also allows us to assume without loss of generality that e, h_1, and h_2 are in disjunctive normal form.

It will help to introduce some notation at this point. We will be dividing the canonical state descriptions of $\mathcal{L}$ up by type, according to which of our three sentences of interest they are disjuncts of. For example, a canonical state description of $\mathcal{L}$ will be described as an $e\overline{h_1}h_2$-disjunct if it is a disjunct of e and of h_2 but not of h_1. (Put another way, an $e\overline{h_1}h_2$-disjunct is a canonical state description that entails e and h_2 but does not entail h_1.) $\#(e\overline{h_1}h_2)$ will be the number of distinct $e\overline{h_1}h_2$-disjuncts. The logical independence of e, h_1, and h_2 guarantees that there is at least one disjunct of each type.

Step 1: Construct language $\mathcal{L}'$. $\mathcal{L}'$ will have the same single name as $\mathcal{L}$ (call it a) with the same referent; we will give the precise number of predicates in $\mathcal{L}'$ below. We specify an interpretation for $\mathcal{L}'$ by first assigning each canonical state description in $\mathcal{L}$ a synonym in $\mathcal{L}'$. The synonyms will all be sentences of $\mathcal{L}'$ in disjunctive normal form; in many cases the synonyms will have only a single disjunct. Most importantly, each canonical state description of $\mathcal{L}'$ will appear exactly once as a disjunct of a synonym of a canonical state description of $\mathcal{L}$.

Step 1a: Assign each eh_1h_2-, $\bar{e}h_1h_2$-, and $e\overline{h_1h_2}$-disjunct of $\mathcal{L}$ a synonym that is a canonical state description of $\mathcal{L}'$.

Step 1b: If $\#(e\overline{h_1}h_2) \geqslant \#(eh_1\overline{h_2})$, skip to Step 1c. Otherwise, assign one of the $e\overline{h_1}h_2$-disjuncts (we've seen that there is at least one) a synonym that is the disjunction of $\#(eh_1\overline{h_2}) - \#(e\overline{h_1}h_2) + 1$ canonical state descriptions of $\mathcal{L}'$. Assign the rest of the $e\overline{h_1}h_2$-disjuncts synonyms that are single canonical state descriptions of $\mathcal{L}'$. Also assign each $eh_1\overline{h_2}$-disjunct a synonym that is a single canonical state description of $\mathcal{L}'$. Then skip to Step 1d.

Step 1c: Assign one of the $eh_1\overline{h_2}$-disjuncts (we've seen that there is at least one) a synonym that is the disjunction of $\#(e\overline{h_1}h_2) - \#(eh_1\overline{h_2}) + 1$ canonical state descriptions of $\mathcal{L}'$. Assign the rest of the $eh_1\overline{h_2}$-disjuncts synonyms that are single canonical state descriptions of $\mathcal{L}'$. Also assign each $e\overline{h_1}h_2$-disjunct a synonym that is a single canonical state description of $\mathcal{L}'$. Then go to Step 1d.

Step 1d: If $\#(\overline{eh_1}h_2) \geqslant \#(\bar{e}h_1\overline{h_2})$, skip to Step 1e. Otherwise, assign one of the $\overline{eh_1}h_2$-disjuncts (we've seen that there is at least one) a synonym that is the disjunction of $\#(\bar{e}h_1\overline{h_2}) - \#(\overline{eh_1}h_2) + 1$ canonical state descriptions of $\mathcal{L}'$. Assign the rest of the $\overline{eh_1}h_2$-disjuncts synonyms that are single canonical state descriptions of $\mathcal{L}'$. Also assign each $\bar{e}h_1\overline{h_2}$-disjunct a synonym that is a single canonical state description of $\mathcal{L}'$. Then skip to Step 1f.

Step 1e: Assign one of the $\bar{e}h_1\overline{h_2}$-disjuncts (we've seen that there is at least one) a synonym that is the disjunction of $\#(\overline{eh_1}h_2) - \#(\bar{e}h_1\overline{h_2}) + 1$ canonical state descriptions of $\mathcal{L}'$. Assign the rest of the $\bar{e}h_1\overline{h_2}$-disjuncts synonyms that are single canonical state descriptions of $\mathcal{L}'$. Also assign each $\overline{eh_1}h_2$-disjunct a synonym that is a single canonical state description of $\mathcal{L}'$. Then go to Step 1f.

Step 1f: Assign each of the $\overline{eh_1h_2}$-disjuncts (we've seen that there is at least one) a synonym that is a disjunctive normal form $\mathcal{L}'$ sentence. It doesn't matter which canonical state descriptions of $\mathcal{L}'$ become disjuncts of which synonym, as long as every canonical state description not assigned as a disjunct of some synonym in Steps 1a through 1e is assigned to some synonym in this step.

To make all these steps possible, $\mathcal{L}'$ must have at least a certain number of canonical state descriptions available. Since the number of canonical state descriptions in $\mathcal{L}'$ is $2^{P(\mathcal{L}')}$, we will set $P(\mathcal{L}')$ equal to the smallest integer satisfying

$$
\begin{aligned}
2^{P(\mathcal{L}')} \geqslant \\
& \#(eh_1h_2) + \#(\bar{e}h_1h_2) + \#(e\overline{h_1h_2}) + && \text{(for Step 1a)} \\
& \#(e\overline{h_1}h_2) + \#(eh_1\overline{h_2}) + |\#(e\overline{h_1}h_2) - \#(eh_1\overline{h_2})| + && \text{(for Steps 1b and 1c)} \\
& \#(\overline{eh_1}h_2) + \#(\bar{e}h_1\overline{h_2}) + |\#(\overline{eh_1}h_2) - \#(\bar{e}h_1\overline{h_2})| + && \text{(for Steps 1d and 1e)} \\
& \#(\overline{eh_1h_2}) && \text{(for Step 1f)}
\end{aligned}
$$

Now assign a synonym in $\mathcal{L}'$ to every disjunctive normal form sentence $x \in \mathcal{L}$ as follows: the synonym of x is the disjunctive normal form sentence of $\mathcal{L}'$ that is the disjunction of the synonyms of the disjuncts of x. (The rest of the interpretation of $\mathcal{L}'$ can be filled in in any way that doesn't prevent $\mathcal{L}'$ from being faithful.) Of particular interest will be $e', h_1', h_2' \in \mathcal{L}'$, the disjunctive normal form $\mathcal{L}'$ synonyms of e, h_1, and h_2 respectively. Our construction of $\mathcal{L}'$ has been designed so that the following hold:

$$
\begin{aligned}
\#(e'\overline{h_1'}h_2') &= \#(e'h_1'\overline{h_2'}) \\
\#(\overline{e'h_1'}h_2') &= \#(\overline{e'}h_1'\overline{h_2'}) \\
\#(\overline{e'h_1'h_2'}) &> 0
\end{aligned}
$$

Note that by language invariance, $h_1' \succ_{e'} h_2'$. Also, $\mathcal{L}'$ is adequate for the propositions expressed by e', h_1', and h_2'.[38]

Step 2: Now construct a language $\mathcal{L}^*$. $\mathcal{L}^*$ will have the same single name (a) as $\mathcal{L}$ and $\mathcal{L}'$ with the same referent. The predicates of $\mathcal{L}^*$ will be $F^*, G^*, B_1^*, B_2^*, \ldots, B_n^*$. (The precise value of n will be given below.) A state description of $\mathcal{L}^*$ is canonical if it first affirms or negates F^*a, then affirms or negates G^*a, then affirms or negates of a some pattern of all the B^*-predicates with those predicates appearing

[38][Note: Here I need to fill in a proof of the adequacy of $\mathcal{L}'$ just asserted. This proof relies on the availability of independent properties.]

in numerical order by subscript. Also define a permutation π of the predicates of $\mathcal{L}^*$ that takes F^* to G^*, G^* to F^*, and each B^*-predicate to itself.

We will specify an interpretation for $\mathcal{L}^*$ by first assigning each canonical state description in $\mathcal{L}'$ a synonym in $\mathcal{L}^*$. The synonyms will all be sentences of $\mathcal{L}^*$ in disjunctive normal form; in many cases the synonyms will have only a single disjunct. Most importantly, each canonical state description of $\mathcal{L}^*$ will appear exactly once as a disjunct of a synonym of a canonical state description of $\mathcal{L}'$.

Step 2a: Assign each $e'h_1'h_2'$-, $\overline{e'}h_1'h_2'$-, and $e'\overline{h_1'h_2'}$-disjunct of $\mathcal{L}'$ a synonym that affirms both F^*a and G^*a, then affirms and denies of a some pattern of the B^*-predicates. (It doesn't matter what pattern of B^*-predicates is assigned, as long as no canonical state description of $\mathcal{L}^*$ is assigned more than once.) Note that π will map each of these assigned state descriptions of $\mathcal{L}^*$ to itself.

Step 2b: Pair off each $e'\overline{h_1'}h_2'$-disjunct with an $e'h_1'\overline{h_2'}$-disjunct. (The equalities above guarantee that they will pair off without remainder.) Each $e'\overline{h_1'}h_2'$-disjunct receives a synonym in $\mathcal{L}^*$ that affirms F^*a, denies G^*a, then affirms and denies of a some pattern of the B^*-predicates. (It doesn't matter what pattern of B^*-predicates is assigned, as long as no two $e'\overline{h_1'}h_2'$-disjuncts receive the same pattern.) The $e'h_1'\overline{h_2'}$-disjunct paired with that $e'\overline{h_1'}h_2'$-disjunct receives a synonym in $\mathcal{L}^*$ that denies F^*a, affirms G^*a, then affirms and denies of a the same pattern of B^*-predicates as its mate. This means that applying π to the synonym of the $e'\overline{h_1'}h_2'$-disjunct will map it to the synonym of the $e'h_1'\overline{h_2'}$-disjunct, and *vice versa*.

Step 2c: Pair off each $\overline{e'h_1'}h_2'$-disjunct with an $\overline{e'}h_1'\overline{h_2'}$-disjunct. (The equalities above guarantee that they will pair off without remainder.) Each $\overline{e'h_1'}h_2'$-disjunct receives a synonym in $\mathcal{L}^*$ that affirms F^*a, denies G^*a, then affirms and denies of a some pattern of the B^*-predicates. (It doesn't matter what pattern of B^*-predicates is assigned, as long as that pattern hasn't already been used for an $e'\overline{h_1'}h_2'$-disjunct or another $\overline{e'h_1'}h_2'$-disjunct.) The $\overline{e'}h_1'\overline{h_2'}$-disjunct paired with that $\overline{e'h_1'}h_2'$-disjunct receives a synonym in $\mathcal{L}^*$ that denies F^*a, affirms G^*a, then affirms and denies of a the same pattern of B^*-predicates as its mate. This means that applying π to the synonym of the $\overline{e'h_1'}h_2'$-disjunct will map it to the synonym of the $\overline{e'}h_1'\overline{h_2'}$-disjunct, and *vice versa*.

Step 2d: The inequality above guarantees that there is at least one $\overline{e'h_1'h_2'}$-disjunct. Assign each $\overline{e'h_1'h_2'}$-disjunct an $\mathcal{L}^*$ synonym in disjunctive normal form. It doesn't matter which canonical state descriptions of $\mathcal{L}^*$ become disjuncts of which synonym, as long as every canonical state description not assigned as a synonym in Steps 2a through 2c is assigned to some synonym in this step.

The number of available B^*-predicate patterns is 2^n. So to make the steps above possible, we set n to the smallest integer meeting all of the following conditions:

$$\begin{aligned} 2^n &\geqslant \#(e'h_1'h_2') + \#(\overline{e'}h_1'h_2') + \#(e'\overline{h_1'h_2'}) && \text{(for Step 2a)} \\ 2^n &\geqslant \#(e'\overline{h_1'}h_2') + \#(\overline{e'h_1'}h_2') && \text{(for Steps 2b and 2c)} \\ n &\geqslant P(\mathcal{L}') - 2 \end{aligned}$$

The last of these inequalities ensures that $P(\mathcal{L}^*) \geqslant P(\mathcal{L}')$, so there are at least as many canonical state descriptions of $\mathcal{L}^*$ as there are of $\mathcal{L}'$. This allows each canonical state description of $\mathcal{L}'$ to have a distinct synonym in $\mathcal{L}^*$.

Now assign a synonym in $\mathcal{L}^*$ to every disjunctive normal form sentence $x' \in \mathcal{L}'$ as follows: the synonym of x' is the disjunctive normal form sentence of $\mathcal{L}^*$ that is the disjunction of the synonyms of the disjuncts of x'. (The rest of the interpretation of $\mathcal{L}^*$ can be filled in in any way that doesn't prevent $\mathcal{L}^*$ from being faithful.) Of particular interest will be $e^*, h_1^*, h_2^* \in \mathcal{L}^*$, the disjunctive normal form $\mathcal{L}^*$ synonyms of e', h_1', and h_2' respectively. By language invariance, $h_1^* \succ_{e^*} h_2^*$. Also, $\mathcal{L}^*$ is adequate for the propositions expressed by e^*, h_1^*, and h_2^*.[39]

By our construction, π maps $e^*h_1^*h_2^*$-, $\overline{e^*}h_1^*h_2^*$-, and $e^*\overline{h_1^*h_2^*}$-disjuncts to themselves. π maps distinct $e^*\overline{h_1^*}h_2^*$-disjuncts to distinct $e^*h_1^*\overline{h_2^*}$-disjuncts and *vice versa*, while $\#(e^*\overline{h_1^*}h_2^*) = \#(e^*h_1^*\overline{h_2^*})$. Also, π maps distinct $\overline{e^*h_1^*}h_2^*$-disjuncts to distinct $\overline{e^*}h_1^*\overline{h_2^*}$-disjuncts and *vice versa*, while $\#(\overline{e^*h_1^*}h_2^*) = \#(\overline{e^*}h_1^*\overline{h_2^*})$. Putting these facts together, we can see that $\pi(e^*) = e^*$, $\pi(h_1^*) = h_2^*$, and $\pi(h_2^*) = h_1^*$. By Result 4, $h_1^* \nsucc_{e^*} h_2^*$. But now we have a contradiction.

□

Appendix B. Responses to Technical Objections

I haven't written this section yet. Here are some of the concerns/objections to which I intend to respond once this section is written:

- Respond to concerns that a finite, first-order language without variables or quantifiers is insufficient to properly represent evidential situations.
- Respond to concerns that a formal language without fancy do-dads like modal structures is insufficient to properly represent evidential situations. (Some overlap with the previous item.)
- Respond to worries that hypotheses and evidence shouldn't be understood as propositions, or that propositions shouldn't be understood as sets of possible worlds.
- Respond to worries about replacing multi-name languages with single-name languages. Perhaps some people think that the universe breaks down into properties of single objects, and that moving to a single name that represents a tuple allows us to construct predicates representing properties of that tuple that don't nicely factor into properties of individuals. (For example, we wind up with a predicate that holds of tuple t just in case F holds of a, or G holds of b, or H doesn't hold of c, etc.)
- Untangle the sorts of inter-language predicate translations given rise to by the constructions in the proof of our general result.
- There might be a worry that the predicate translations we wind up with from one language to another mention particular individuals. So, for instance, we wind up with something that says "F' in language $\mathcal{L}'$ agrees with F in $\mathcal{L}$ on objects a, c, d, and g and disagrees with F on b, e, and f."
- Respond to the worry that the way I've set up my framework (with the equivalence condition especially) means that there can be no interesting favoring relations among logical truths.

[39][Note: Here I need to fill in a proof of the adequacy of $\mathcal{L}^*$ just asserted. This proof relies on the availability of independent properties.]

- Respond to concerns about the availability of independent properties.
- Respond to the claim that we wouldn't have these problems if we worked with infinite languages.
- If evidential favoring is contrastive, it may only apply to pairs of hypotheses that are mutually exclusive. The substantivity requirement that evidence and hypotheses all be mutually logically independent is therefore too strong.

  Response: In the proof of our general result, the logical independence of e, h_1, and h_2 is initially used to guarantee that disjuncts of all types are available in $\mathcal{L}$. However, the rest of the proof doesn't actually require the availability of *all* types of disjuncts; it only requires the availability of disjuncts of types $e\overline{h_1}h_2$, $eh_1\overline{h_2}$, $\overline{eh_1}h_2$, $\overline{e}h_1\overline{h_2}$, and $\overline{eh_1h_2}$. These disjunct types can all be available even if h_1 and h_2 are mutually exclusive (or are mutually exclusive given e). So the proof will be able to demonstrate a contradiction even if the full logical independence of e, h_1, and h_2 is not assumed and $\succ_e$ is stipulated to hold only when h_1 and h_2 are mutually exclusive.