<u>**Not Enough There There**</u>
Evidence, Reasons, and Language Independence

Michael G. Titelbaum

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

The philosophical significance of language dependence results (à la Goodman's "grue" problem) has been seriously underestimated. In this article I argue that:

- The trouble caused by language dependence is not confined to *formal* analyses of evidential support; language dependence concerns apply to any story about how we should draw conclusions from evidence.
- This is because language dependence reveals a deep problem: evidential propositions do not possess enough information to substantively favor some conclusions over others.
- Introducing a distinction between natural and nonnatural properties does not alleviate this problem.
- The problem is not equally a problem for everyone—it makes some views of evidential favoring much harder to maintain than others.

The special victims of the problem I will discuss posit an *evidential favoring relation*: an objective three-place relation between two hypotheses and a body of evidence that obtains when the evidence favors one hypothesis over the other. By "objective" I mean that the relation does not vary with particular facts about agents. Subjective facts about an agent may determine what counts as his evidence, but our question is whether, once the evidence is established, what that evidence favors depends on subjective considerations.[1]

A wide variety of contemporary philosophical views maintain the existence of an evidential favoring relation, though they often don't call it by that name. Some epistemologies make it an objective fact that a particular body of evidence *supports* one hypothesis over another, or *provides more justification to believe* one hypothesis. Epistemologists may make this more precise by suggesting that each hypothesis has a specific *evidential*, *logical*, or *objective* probability relative to a given body of evidence.[2] The evidence then favors (or *confirms*) one hypothesis over another just in case it renders the former more probable than the latter.[3]

In a different area of philosophy, metaethicists consider whether an agent's evidence provides more *reason* to believe one hypothesis than another. A widespread

[1] I am interested in evidential favoring that 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.

[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] One version of this view follows Carnap in maintaining that the objective probability of a hypothesis on a body of evidence equals the *degree of belief* a rational agent would assign that hypothesis were the evidence his total evidence. Objective probability defenders are not required to accept this claim, though.

I should also note that at this point I am following (the early) Carnap in discussing a "firmness" approach to confirmation rather than an "increase in firmness" approach. The arguments to follow, however, apply equally well to both. (For further discussion of the firmness/increase in firmness distinction and references to Carnap, see (Fitelson 2006).)

metaethical view maintains that there are objective reason relations, which in the epistemic case would yield objective facts about which evidential propositions provide reasons to believe which hypotheses.[4] We could then say that a body of evidence favors one hypothesis over another when it provides reason to believe the former and disbelieve the latter.[5]

What Goodman's grue problem makes vivid is that if there is such an evidential favoring relation, and if it goes beyond simply saying which hypotheses are logically *entailed* by our evidence, it must treat some properties differently from others—that is, it must treat some properties as "special." This opens up the following *reductio* against evidential favoring relations:

(1) There is an evidential favoring relation that holds in at least some cases in which the evidence and hypotheses are logically independent. *(premise)*
(2) In at least some of these logically independent cases, agents can determine that the body of evidence favors the one hypothesis over the other. *(premise)*
(3) Agents can detemine that a body of evidence favors one logically independent hypothesis over another only if they can determine that particular properties are special (or something equivalent).
(4) Hypotheses about which properties are special are empirical hypotheses that must be determined from an agent's evidence and are logically independent of that evidence. *(premise)*
(5) So agents can determine that particular properties are special (or something equivalent) only if they can determine that a particular body of empirical evidence favors one logically independent hypothesis over another.
(6) Because of the circularity in (3) and (5), such a determination cannot be made.
(7) So there are no cases in which a body of evidence favors one logically independent hypothesis over another and agents can determine this fact. This contradicts the conjunction of (1) and (2) above.

This *reductio* is hardly the first challenge that has been raised to evidential favoring views such as the objective reasons theory. Philosophers have raised both metaphysical questions (What *is* this objective reasons relation?) and epistemic questions (By what *faculty* do we discern when it obtains? Why should we think that faculty is *reliable*?).[6] But the argument above, though epistemic, is still new: Setting aside questions of what faculty would detect evidential favoring, it focuses on what that faculty would have to be able to *do* to determine that a body of evidence favored one hypothesis over another.

---

[4]Adherents of objective-reasons views concerning *practical* reasoning include Broome (2004), Dancy (1993), Kolodny (2005), Parfit (2006), Raz (1999), Scanlon (1998), and Shafer-Landau (2009). Among these, Broome, Kolodny, and Scanlon are explicitly committed to an objective-reasons view of *theoretical* reasoning.

[5]I have decided to work with a three-place evidential favoring relation ("the evidence favors this hypothesis over that one") instead of a two-place relation ("the evidence favors this hypothesis") to accommodate positions (such as Likelihoodism) according to which evidential favoring is always contrastive. Views about two-place relations can be worked into this scheme in the manner just demonstrated for the two-place relation of evidence's providing reason to believe a hypothesis: we use the two-place facts to generate a three-place relation.

[6]See, for example, (Parfit 2006, p. 330), (Scanlon 1998, Ch. 1, Section 11), and (Enoch 2010).

The argument derives a contradiction from three premises. The first half of this article shows that the premises are genuinely contradictory by defending steps (3) and (6) above.[7] Section 1 describes a technical result that generalizes Goodman's grue problem. It demonstrates that any process that detects evidential favoring among logically independent evidence and hypotheses must have a bias towards certain properties (or something equivalent) that *precedes* the influence of the evidence. This technical result (which I prove in Appendix A and then defend from objections in Appendix B) establishes step (3) in the argument above. Section 2 then uses the result to generate the circularity problem in step (6): Agents cannot discern evidential favoring unless they can tell that certain properties are special, but they cannot tell that certain properties are special without discerning evidential favoring. Section 3 defends step (6) from claims that this circularity problem can be solved by Lewisian "reference magnetism" or that the circle is virtuous rather than vicious.

Once we have shown that the three premises are contradictory, we must choose one to discard. The latter half of this article works through each of the options. Since our evidence does not *entail* that particular properties are special, the only way to deny step (4) is to deny that hypotheses about the specialness of properties are empirical. So in Section 4 I consider a rationalism that rejects step (4) by claiming that agents can discern the specialness of propreties (or something equivalent) *a priori*. In Section 5, I consider an externalism that rejects step (2) by holding that while evidential favoring can obtain between logically independent evidence and hypotheses, agents can never determine in a particular instance that this is the case. Finally, in Section 6 I suggest that the trouble is right at the beginning—in step (1). We should abandon the notion of objective, three-place evidential favoring that goes beyond entailment and acknowledge that in most cases in which we select among hypotheses on the basis of evidence, the favoring relations to which we appeal must be relative to subjective agential features.

1. The General Result

Goodman's "grue" lecture (Goodman 1979) is by far the best-known philosophical discussion of language dependence, but for present purposes it has a number of shortcomings. First, Goodman doesn't actually *demonstrate* anything nearly as general as he claims. Goodman concludes his discussion by claiming that "lawlike or projectible hypotheses *cannot* be distinguished on *any* merely syntactical grounds." (pp. 81-2, emphases mine) But in the course of that discussion he examines only a handful of formal confirmation theories—all drawn from his contemporaries[8]—and gives little indication how to generalize his approach. Moreover, Goodman leaves it unclear exactly what the problem with those formal theories is supposed to be.[9] In Appendix A I address these shortcomings by proving a result much more general than anything Goodman provides. That result defines properties one would want

[7] I take it that steps (5) and (7) obviously follow once these steps and the premises are in place.

[8] In the first sentence of his (1946) Goodman lists as his targets Hempel, Carnap, Oppenheim, and Helmer.

[9] For example, Goodman claims Hempel's theory yields "the intolerable result that anything confirms anything." (Goodman 1979, p. 75) Yet Hempel *proves* that on his theory of confirmation the same evidence cannot confirm contradictory hypotheses (see the discussion of the "Consistency Condition" in (Hempel 1945)), so clearly something is wrong with Goodman's argument against Hempel. (For extensive analysis see (Hooker 1968).)

an evidential favoring relation to have, then shows that these properties cannot coexist. I will explain the result informally here, then assess its broader significance.

Suppose there is an objective, three-place favoring relation that holds of two hypotheses and a body of evidence just in case the evidence favors the first hypothesis over the second. We will take the hypotheses to be propositions and the evidence to be a proposition as well.[10] In order to work with these propositions, we must represent them in some language. If h_1 is a sentence in a language representing the first hypothesis, h_2 a sentence in that language representing the second hypothesis, and e a sentence representing the evidence, we will write $f(h_1, h_2, e)$ when the evidence favors the first hypothesis over the second.[11]

What should we assume about the relation f? We will assume that relative to a given e the relation is antisymmetric: for any h_1, h_2, and e, if $f(h_1, h_2, e)$ then it's not the case that $f(h_2, h_1, e)$.[12] We will not, however, assume that relative to a given e the relation f introduces a *total ordering* over hypotheses. That is, we will not assume that for any e, h_1, and nonequivalent h_2 it's the case that either $f(h_1, h_2, e)$ or $f(h_2, h_1, e)$. We want to allow the possibility that evidence can support nonequivalent hypotheses equally, or that there may be incommensurate favorings. My current total evidence might favor the proposition that the Lakers will win this year's NBA championship over the proposition that the Celtics will, 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 total evidence favors the proposition that the Lakers will win the championship over the proposition that Palin will run for President.[13]

The defender of objective evidential favoring relations should agree to a further requirement on f. I cannot dispute that there are some trivial cases in which a body of evidence supports one hypothesis over another—for instance cases in which the evidence deductively entails one hypothesis while refuting the other. But as Hume taught us, the evidential favorings that arise most frequently in our everyday lives do

[10]For those who prefer to think of a body of evidence as a set of propositions, it will make no difference if we let the third relatum be the proposition that is the conjunction of that set. It will also make no difference if we add a requirement that the evidential proposition be *true*, as people will who require reasons to be facts.

[11]Note that "f" is not a part of the imagined language—for instance, if the language in question is a formal language, "f" will not be one of its symbols. "$f(h_1, h_2, e)$" is something we say in the *meta*language to indicate that the proposition represented by e favors the proposition represented by h_1 over the proposition represented by h_2. (Here "h_1", "h_2", and "e" are being used as metalinguistic variables.)

[12]This means that our results may not apply if $f(h_1, h_2, e)$ represents a notion of evidential favoring like "the proposition represented by e provides *some* reason to believe the proposition represented by h_1 over the proposition represented by h_2," because it's possible that the same body of evidence could provide *some* reason to believe one hypothesis over another while also providing *some* reason to believe the latter over the former. Our results will apply to antisymmetric favoring notions such as "the proposition represented by e provides *all-things-considered* reason to believe the proposition represented by h_1 over the proposition represented by h_2."

[13]In fact, we won't assume that relative to a given e the relation f introduces any kind of ordering at all. That's because we won't assume that evidential favoring by a given body of evidence is transitive. (We won't assume that for any e, h_1, h_2, and h_3 if $f(h_1, h_2, e)$ and $f(h_2, h_3, e)$ then $f(h_1, h_3, e)$.) Our result is *consistent* with the possibility that f introduces an ordering (or even a total ordering) over hypotheses, but doesn't *require* that. Similarly, while our result is *consistent* with the possibility that evidential favoring comes in precise numerical degrees, it doesn't *require* that there be anything like quantitative "degrees" of favoring. (For Sorites-style concerns about transitivity and favoring relations, see (Novack 2010).)

not involve entailment relations. 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, even though it is logically independent of both those propositions. For a more precise example, if your evidence is that a number between 1 and 10 was drawn in a perfectly random fashion 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.[14] These examples and countless others like them demonstrate that if there is an evidential favoring relation, it must be *substantive*; there must be at least some cases in which evidence favors one hypothesis over another even though the three relata are logically independent.[15]

The conditions listed so far come from thinking just about the evidential favoring relation ("f") itself. Now suppose (just for a moment—we'll be lifting this supposition shortly) that some formal theory can detect when the evidential favoring relation obtains. The idea would be that the theory could take sentences h_1, h_2, and e expressed in some *formal* language and discern just from the arrangement of the symbols in those sentences whether $f(h_1, h_2, e)$. The theory would not have to invoke anything about the meanings of the language's predicates or the referents of its constants; it would work on syntactic considerations alone.

Even for a defender of formal theories of evidential favoring, the job description in the last paragraph is a bit too demanding. The theory cannot be expected to pick up on evidential favoring relations when hypotheses and evidence are expressed in just *any* formal language. For instance, we might have a propositional language in which the hypotheses and evidence are represented with three atomic sentences. (Imagine a language that represented "The first 1000 emeralds have been green," "the next emerald will be green," and "the next emerald will have feathers" with the atomic sentences "p", "q", and "r" respectively.) Even if an evidential favoring relation held among the hypotheses and evidence, their representations in this language would not reveal enough of their internal structure for the formal theory to properly analyze. So we should demand only that a formal theory detect evidential favoring when hypotheses and evidence are represented in an *adequate* language—a language rich enough to express the structural features of those propositions relevant to any favoring relations between them.

To get the result we want we need not say much about which languages are adequate for which hypotheses and evidence. All we need is that the adequacy concern is a concern about representational *paucity*—a concern that a language will not have *enough* formal structure to reveal evidential relations. We will assume that if one language is at least as expressive as another, the former is adequate for any hypotheses and evidence that the latter is adequate for.

[14]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" approaches to favoring—see note 3 above.)

[15]In Appendix B I discuss loosening the substantivity requirement to accommodate the position (maintained by some philosophers of science) that evidential favoring occurs only between mutually exclusive hypotheses. Our general result can still be proven in this case.

Finally, if the theory in question works purely syntactically—purely on the arrangement of symbols in sentences—it must treat predicate permutations identically. That is, if in an adequate language we can swap predicates around so that h_1 becomes h_2 and h_2 becomes h_1 while e continues to represent the same proposition, then our theory must treat h_1 and h_2 identically. For example, suppose that in some adequate language h_1 is "$Fa \vee Ha$", h_2 is "$Ga \vee Ha$", and e is "$Ha \vee Ia$". The permutation (swapping) that exchanges "F" and "G" will convert h_1 into h_2 and *vice versa*, while leaving e unaffected. In this case, for a theory of favoring to treat h_1 differently from h_2 relative to e (for instance, for the theory to assert that e favors h_1 over h_2) would be for the theory to pick up on something more than these sentences' syntactical form.[16] So if h_1 and h_2 are predicate permutations no formal, syntactical theory will suggest that e favors h_1 over h_2. And since we are supposing that some formal theory can accurately detect the presence of evidential favoring, this will in turn mean that $f(h_1, h_2, e)$ does not hold.

To recap: We have supposed that there is an evidential favoring relation, and we have described some features of that relation. Most importantly, we have assumed that the evidential favoring relation is substantive, meaning that it obtains for at least some logically independent evidence and hypotheses. We have also supposed that the relation's presence can be detected by a formal theory. This entails the identical treatment of predicate permutations: for any h_1, h_2, and e in an adequate language, if a predicate permutation interchanges h_1 and h_2 while leaving e intact then it's not the case that $f(h_1, h_2, e)$. Appendix A proves the following simple result:

General Result: Substantivity and the identical treatment of predicate permutations are inconsistent.

The proof proceeds by showing that given any h_1, h_2, and e in an adequate language with no special entailment relations, there exists an adequate language in which the same propositions are expressed by an h'_1, h'_2, and e' such that a predicate permutation swaps h'_1 with h'_2 while leaving e' equivalent. By the identical treatment of predicate permutations, e' cannot favor h'_1 over h'_2. But h'_1, h'_2, and e' represent arbitrarily selected propositions with no special entailment relations. So there can be no evidential favoring among such propositions, and the evidential favoring relation cannot be substantive.

---

[16] At this point one might worry—and an audience at the University of Sydney did—that I have moved from talking about a theory that is purely *syntactical* to talking about a theory that works solely with syntactical *form*. For instance, a theory that said that sentences containing the predicate symbol "F" are always favored over sentences containing the symbol "G" would still be *syntactical*, even though it would treat h_1 and h_2 differently in the example above. But as I clarify in Appendix A, the definitions given so far entail that f is *language invariant*: if $f(h_1, h_2, e)$ and h'_1, h'_2, and e' re-express h_1, h_2, and e (respectively) in a different adequate language, then $f(h'_1, h'_2, e')$. In the case under consideration, if we had a formal theory that treated hypotheses containing the symbol "F" differently than it treated hypotheses containing the symbol "G", then translating the propositions expressed by h_1 and h_2 into a different adequate language in which "F" was used to express the property "G" had been used for previously (and *vice versa*) would reverse the formal theory's judgments about favorings. Since f is language invariant, such a formal theory would not be accurately tracking the presence of the evidential favoring relation it was designed to detect. So the imagined formal theory would be incapable of doing its job. Similar language invariance arguments undermine any theory that picks up on syntactical features other than syntactical form.

The proof in Appendix A is constructive; it shows exactly how to build the required language given any h_1, h_2, and e. And unlike proofs that are sometimes used to refute Principles of Indifference and the like, it doesn't rely on any curious features of infinite sets; all the languages involved have finite atomic sentence sets.[17]

This result 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 our general result says that if the evidential favoring relation is substantive, it cannot treat predicate permutations identically and therefore cannot be captured by a formal theory of evidence.

But maybe formalizing evidential support was an unattractive prospect to begin with. Evidential relations are enormously complex—we might have suspected from the start that their vast subtleties cannot be adequately captured by a theory that pushes symbols around, no matter how inventively it does so. Many philosophers think that grue's significance lies entirely in its confirmation of this suspicion—as we already saw, Goodman's takeaway was a condemnation of theories that operate on "merely syntactical grounds." But this is another important shortcoming of Goodman's discussion; there is much more to be learned from our general result than just a lesson about formal theorizing.[18]

We arrived at most of the conditions described above by thinking simply about evidential favoring. The only place where we invoked the idea that favoring could be detected by a formal theory was in motivating the identical treatment of predicate permutations. But if we look at that condition itself (setting aside its motivation), the condition doesn't mention formal theories at all. It says that if two hypotheses and a body of evidence can be represented in an adequate language such that a predicate permutation interchanges the hypotheses while leaving the evidence intact, then the body of evidence doesn't favor one of the hypotheses over the other. Our general result tells us that if the evidential favoring relation is substantive, it cannot satisfy this condition.

What kind of favoring relation treats predicate permutations identically, and what kind of favoring relation does not? 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. For instance, suppose we have a piece of evidence that mentions greenness and grueness in exactly the same ways, but that evidence favors a hypothesis involving the property of being green over a hypothesis that involves the property of being grue in structurally identical ways. If the evidential favoring relation behaves in this way, it fails to treat predicate permutations identically. And notice that this property favoritism *precedes the influence of the evidence.* It's not that the difference occurs because the evidence indicates that greenness is a property worthy of special evidential

[17]While our result is much more general than Goodman's, it falls short of *perfect* generality by following Goodman in considering only hypotheses and evidence that can be adequately represented in first-order languages (without variables or quantifiers). Appendix B discusses how this limitation affects the lessons we can draw from the result.

[18]Goodman was, of course, just the first of many authors to take too narrow a lesson from "grue". To pick an example at random of someone who tows the Goodmanian line, Earman (1992) discusses a grue-style predicate then concludes, "an adequate account of confirmation must be sensitive to semantics." (p. 69)

consideration; we stipulated that the evidence says exactly the same things about (or using) greenness that it says about (/using) grueness. If we could behold the $f(\cdot,\cdot,\cdot)$ relation itself before any evidence had been plugged in, we could already see that plugging in evidence and hypotheses involving certain properties would cause it to react differently than plugging in evidence and hypotheses that differed only in the properties that appeared.[19]

Incidentally, this explains why formal theories of evidential favoring repeatedly ran into language-dependence problems over the course of the 20th century.[20] From the point of view of a theory working with the syntax of sentences in formal languages, working within a particular language is a way of privileging particular properties, such as the ones that are represented by predicates in that language. To take a silly example, you might have a formal theory holding that hypotheses that affirm more predicates are always favored over hypotheses that affirm fewer. Applying this theory to a language with a predicate-letter "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-letter "C" for closed would have the opposite effect. In general, whether the sentence representing a particular proposition has the features emphasized by a particular syntactical theory will be an artifact of language choice.

Our general result tells us that in order to capture a substantive favoring relation, a formal theory must privilege certain properties in some way, and an easy way to do this would be for the theory to apply only to designated "special languages"—those languages that represent the underlying "special" properties in the right way. The trouble is that designating special languages is a non-syntactical affair. 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 our general result is not about formal theories of favoring. What we have learned is that a body of evidence is not sufficient in itself to substantively favor some hypotheses over others. It must work in concert with something (a preferred language, a property list, etc.) that does the work of privileging some

[19] Once we turn our attention from what the general result says about *formal theories* of evidential support to what it says about the favoring relation *itself*, we might wonder what to make of the emphasis in that result on adequate languages. Adequacy still matters because even if we want a favoring relation that is unbiased among properties, it's too much to expect favoring to be preserved under predicate permutations in *every* language. (Consider again our earlier example about emeralds, greenness, and feathers.)

[20] Goodman's "grue" problem is the best-known example of these problems for theories such as Hempel's and Carnap's. But there have been other formal theories of favoring, such as Jaynes's "maximum entropy approach," and each has been shown to have language-dependence problems of its own. (The maximum entropy approach was introduced in (Jaynes 1957a) and (Jaynes 1957b); for language dependence criticisms see (Seidenfeld 1986).) To my knowledge the result presented here is the first that applies in general to all possible formal analyses of evidential favoring.

properties over others.[21] By itself, the informational content of a body of evidence is insufficient to yield a substantive evidential favoring relation.

2. 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 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."[22]

I can't deny that this is possible. But even if there are natural properties, and even if they are the ones that get special attention from the evidential favoring relation, we are left with an epistemic problem: How are agents to determine which properties are natural?

On the view that takes natural properties to be projectible, determining the natural properties list is central to the project of determining which bodies of evidence favor which hypotheses. The natural properties list (we suppose for now) is an empirical fact about our world; it must be determined from our evidence. And yet our evidence doesn't carry the natural properties list on its sleeve; nothing we observe ever *entails* that such-and-such is a natural property while such-and-such is not. So selecting the correct natural properties list is a matter of determining which of a number of hypotheses is favored by our evidence.

Now consider an agent—call him Pedro—who has some hypotheses, none of which is entailed or refuted by his evidence. He 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 determine any favorings 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 empirical fact, he decides to determine the list from his total evidence. But he knows he can't determine what his evidence favors until he has the list of natural properties. . . .

We can make this circularity problem more precise. Imagine that somehow there were a process (or set of processes) by which Pedro could take his evidence, work

---

[21] It is sometimes suggested that contextual factors determine how strongly evidence has to support a hypothesis before it counts as favoring that hypothesis. But even if evidential favoring is made relative to particular contextual factors for this purpose, it will need to be *further* relativized to some element (contextual or otherwise) that privileges certain properties over others. For let f represent the relation generated by taking the contextual evidential favoring relation and holding the contextual factors in question fixed at some "value." Presumably we will want this f to be substantive, so our general result will apply.

[22] Depending on one's metaphysical theory of the natural properties, it may be possible that there are projectible properties and there are natural properties, but the two lists of properties don't line up. However, since natural properties are the most popular candidate for the projectible properties, I am going to move freely between talking about "special properties," "projectible properties," and "natural properties." What I say below about natural properties will typically apply to other theories of the projectible properties as well.

out the list of natural properties from it, and then determine (relative to that set of natural properties) which of two hypotheses was favored by that same evidence. We can capture the outputs of this process by a three-place relation: we will say that $np(h_1, h_2, e)$ holds just in case the natural properties list Pedro's process works out from the evidence represented by e generates a favoring relation on which e favors the hypothesis represented by h_1 over the hypothesis represented by h_2.

I've described Pedro's 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. It also makes no difference whether the process employs a formal theory, or a group of informal heuristics; whether it works with general principles or is unavoidably particularistic; whether it works through conscious, deliberate reasoning or whether it relies on cognitive virtues and judgment calls. The process might take into account a variety of considerations: it might generate the simplest possible theory consistent with the evidence and evaluate hypotheses according to that; it might infer the best possible explanation of the evidence and go from there; it might use any subtle process you like from epistemology, statistics, the philosophy of science, or wherever else. However the process works, ultimately it will have to take a pair of hypotheses and some evidence as inputs and produce a favoring judgment as output.

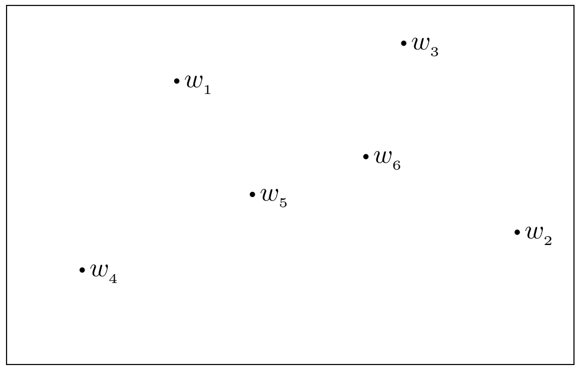
And here the *generality* of our result kicks in: the net effect of that process, captured as the relation $np(\cdot, \cdot, \cdot)$, will satisfy the conditions we listed earlier for $f(\cdot, \cdot, \cdot)$ and so be susceptible to our theorem. np will be antisymmetric for a given e. And we certainly want it to be substantive—if Pedro's process works it should give him more than just what he could get from strict entailment relations. But our general result tells us that if np is substantive, it does not treat predicate permutations identically. So it displays a bias towards certain properties that is prior to and independent of the influence of any evidence. But supplying such a bias was the job the natural properties were supposed to do! In order for the list of natural properties to play its envisioned role in shaping the evidential favoring relation, it cannot be determinable from an agent's evidence.[23]

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 preferred language, a special properties list, 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.[24]

[23]Given Pedro's particular epistemic project, not every item on the natural property list will be relevant to determining whether his evidence favors particular hypotheses over others. But the argument just provided shows that by itself Pedro's evidence won't be sufficient to determine even that *portion* of the natural properties list that underlies the substantive evidential favoring relations in which he's interested.

[24]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 any of those measures. When I say that a body of evidence does not have enough information

Figure 1



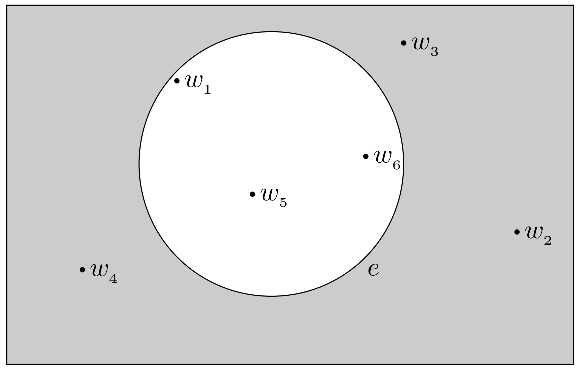
I'll conclude this section by illustrating the point with a diagram. Historically, one popular strategy (epitomized by Carnap's approach) for analyzing evidential favoring has been to suppose that a body of evidence favors certain possible worlds over others—perhaps by making them more *probable* than the others. Once a favoring relation has been established over the possible worlds, determining which propositions composed of those possible worlds are favored by the evidence is a relatively straightforward affair.

In Figure 1 I have depicted the space of all possible worlds as a rectangle; I've marked some of those possible worlds with dots and 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 out some possible worlds and then 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)—the ones that were "closest to being ruled out by the evidence." 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.[25]

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!"

[25]One might think there is an obvious geometrical arrangement for the possible worlds: one that makes possible worlds close to each other in proportion to their similarity. (Among philosophers "close" has become virtually synonymous with "similar" when applied to possible worlds.) Yet determining evidential favoring on the basis of possible-world proximity would generate yet another circle: to determine which worlds are similar, one must determine the list of special properties such

Figure 2



3. Reference Magnets and Virtuous Circles

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 on an empirical basis. In this section I will examine two positions that try to undermine that argument; in the sections to follow I will examine positions that accept the argument but nevertheless maintain the existence of an objective, three-place evidential favoring relation.

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." Natural properties played a role not only in 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 letting natural properties determine 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. But even if we grant Lewis both his metaphysics of natural properties and his philosophy of language, and build in all the complexities he proposes, reference magnetism will still not solve our epistemic problem.[26] The basic issue is that a speaker of

---

that worlds are similar by virtue of having those properties in common; to determine that list of special properties, one must determine what conclusions are favored by one's evidence; and to determine what conclusions are favored by evidence, one must determine which worlds are similar.

It's no coincidence that the discussion of language dependence and evidential favoring in the wake of "grue" parallels 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 would suggest that the problem goes farther than Niiniluoto sees.

[26]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

a language needs some way to determine which of that language's predicates are 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 at all times. 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 determine which predicates in his language pick out natural properties.[27]

One might think that Nigel should simply determine which predicate is more entrenched in his language[28] —perhaps "GRUE" was very recently introduced by troublesome philosophers—and assume that by the magic of reference magnetism that predicate identifies a natural property and so is projectible. But the well-established predicate need not be the one that picks out the natural property. Consider Lavoisier's late-18$^{\text{th}}$ century introduction of the term "oxygen" as part of a theory of heat meant to displace the phlogiston theory. A chemist of the time might have genuinely wondered whether "phlogiston" or "oxygen" carved nature at its joints, featured in genuine scientific laws, and could be reliably employed in making projections about future experiments. Here it would do no good to point out that "phlogiston" had been in the lexicon longer. The crucial issue would be whether "phlogiston" or "oxygen" played a central role in the scientific theory best supported by the available evidence. But if determining which theory was best supported by the evidence depended on ascertaining the correct list of natural properties, our chemist would be stuck. Even if he were somehow provided with (and believed) the full Lewisian story about reference magnetism, he would be unable to determine to which of the predicates in his language that story applied.

Setting reference magnetism aside, one might decline to participate in our argument's "cyclophobia":[29] In philosophy some circles are virtuous, and the circle involved in Pedro's process might just be one of those. The seven-step *reductio* in the introduction claimed that an agent can't determine a favoring relation unless he can determine a special properties list, and he can't determine a special properties list unless he can determine a favoring relation. But maybe there is a way for Pedro to determine both at the same time.[30] After all, Pedro (or any of us) will not start from scratch in trying to determine which properties are preferred and where the evidential favoring relations lie. He will have some set of properties he tends to

1983) and (Lewis 1984)), and despite the fact that green/grue was one of his most frequent examples of a natural/unnatural property contrast, he never once suggested that natural properties would solve Goodman's projectibility problem.

[27]The problem is particularly stark in Lewis, for whom scientific laws are regularities in the systematization of nature that best balances simplicity and strength, and simplicity is measured in a language whose primitive predicates refer to natural properties. (See (Lewis 1983, pp. 367-8).) How can Nigel hope to determine the natural properties list from scientific theory when candidate scientific theories must be evaluated in a language that expresses natural properties?

[28]Compare (Goodman 1979, Lecture IV).

[29]Mark Colyvan introduced me to this term; he attributed it to Adam Elga.

[30]This suggestion denies step 6 of the *reductio*.

think of as fundamental, or he will speak some natural language that employs certain predicates instead of others. He will also have opinions about which bodies of evidence favor which hypotheses over which others. Perhaps he can engage in some sort of reflective equilibrium process by which he develops an evidential favoring relation from his list of preferred properties, compares it to what he thinks he knows about the true evidential favoring relation, then adjusts his preferred properties list, adjusts his opinions about favoring, etc. The idea would be that by a process of mutual adjustment Pedro could somehow bootstrap his way to a properties list and evidential favoring relation that line up correctly.[31]

It's important to notice that we are *not* discussing here the kind of process (familiar to us) by which an agent gains evidence over time and gradually adjusts his views of the world to fit. There are reasons to think that processes like *that* can gradually hone in on accurate results in the long run,[32] even if they start from a point that's far from the truth. But in the equilibrium process we're imagining Pedro keeps his total evidence fixed throughout; the only things that change over the course of the process are his special properties list and his views on evidential favoring.

The key question to ask about this equilibrium process is whether it is guaranteed to wind up at the same point no matter where it starts. That is, would an agent who employed the process starting from any property list and any set of assumptions about evidential favoring ultimately wind up with the same relation? If not, then there are two possibilities: either (i) there is a true, unique evidential favoring relation and the equilibrium process is not a means of determining its extension; or (ii) the equilibrium process is picking up on the truth about evidential favoring, but whether an agent's total evidence favors one hypothesis over another depends not just on the evidence and hypotheses but also on whatever factors supply the agent with his bootstrapping starting points, such as the native language with which he was raised. To embrace the latter possibility is to deny that there is an objective three-place evidential favoring relation at all.

This last option brings up an important consequence of our discussion of objective favoring. Consider Roger White's "Uniqueness Thesis":

> **Uniqueness:** Given one's total evidence, there is a unique rational doxastic attitude that one can take to any proposition.[33]

If there is an objective, three-place evidential favoring relation this will not guarantee that Uniqueness is true. We have already admitted that the evidential favoring relation may not introduce a total ordering among hypotheses relative to a given body of evidence, in which case rationality may allow for differing doxastic attitudes where there are gaps. And even if the evidential favoring relation does introduce a

[31] For an idea of the kind of process I'm talking about, see the discussion of correcting errors in pretheoretical judgments via bootstrapping at (Bealer 1992, p. 106).

[32] See, for instance, the overview of convergence results in (Earman 1992, Chapter 6).

[33] (White 2005). White, in turn, attributes the thesis to (Feldman 2007), though Feldman's Uniqueness Thesis is slightly different. Uniqueness is also linked to forms of Evidentialism. Kelly (2008), for instance, writes, "Some philosophers hold that what one is justified in believing is entirely determined by one's evidence. This view—which sometimes travels under the banner of 'Evidentialism'—can be formulated as a supervenience thesis, according to which normative facts about what one is justified in believing supervene on facts about one's evidence. . . . Thus, according to the Evidentialist, any two individuals who possessed exactly the same evidence would be exactly alike with respect to what they are justified in believing about any given question."

total ordering, it may be that only an ideal reasoner would be capable of governing his doxastic attitudes perfectly by that ordering. The first objection most people have to Uniqueness is that it just seems too demanding to maintain that if two agents with identical evidence disagree on anything, at least one of them is failing to be rational. An everyday agent might make a few mistakes here and there, diverge from what the ideal reasoner would do given his evidence, and still clear the bar for rationality.

So the existence of an objective, three-place evidential favoring relation does not entail Uniqueness. On the other hand, if we *reject* an evidential favoring relation on the grounds that evidential favoring must be relative to some feature that varies from agent to agent, this seems fatal to Uniqueness. It must always be at least rationally *permissible* for an agent to align his doxastic attitudes with his evidence and the favoring relation that applies to him. If favoring depends on subjective features that do not supervene on one's total evidence, there will be cases in which two agents with the same evidence vary in those subjective features, making it rationally permissible for them to take different attitudes towards particular hypotheses (because the evidence favors different hypotheses for one agent than it does for the other).

Now suppose on the other hand that the equilibrium process under consideration ends up at the same evidential favoring relation and list of special properties no matter where it begins. Then we can show that the process must treat some properties differently than others antecedent to the influence of the agent's initial property list, proposed favoring relation, and total evidence. The proof appears as a corollary to our general result in Appendix A, but the rough idea is this: Take a case in which we have two agents. One agent applies the equilibrium process to a proposed favoring relation, a body of total evidence, and a property list that includes green but not grue. The other agent applies that process to the same relation, evidence, and list except that grue has been uniformly substituted for green throughout. Under our current supposition both agents will settle in equilibrium on the same favoring relation—namely the true, objective three-place evidential favoring relation. But we know that that relation treats green differently than it treats grue.[34] Since we started out with inputs that were perfectly symmetrical with respect to green and grue, something in the equilibrium process itself must have broken the symmetry by the time we reached our favoring-relation output. In other words, something in the equilibrium process itself must treat green differently from grue.

The response here is very similar to our response to the proposal that Pedro could squeeze a favoring relation between evidence and hypotheses out of his evidence alone. In both cases a favoring relation (and therefore a preferred property list) is supposed to arise out of particular materials, but in order for this to work the process applied to those materials must already prefer particular properties before the materials are introduced. As we try to find a way for agents to discern which properties are natural, we learn that they cannot determine this either from their evidence or from the proposed favoring relation and property list they contingently

---

[34]The proof doesn't actually assume that green and grue provide an example of properties treated differentially by the evidental favoring relation. It assumes only that the evidential favoring relation is substantive, which (by our general result) means it must treat *some* properties differently than others. Whatever the differentially treated properties actually are, they could be substituted for "green" and "grue" in the discussion above.

have on hand. This leads to the suggestion that properties can be determined to be natural *a priori*.

4. *A Priori* Special Properties

In its broadest outlines, our epistemic problem for the evidential favoring relation suggests that something we thought we could determine by reasoning from empirical evidence cannot be determined in that way after all. At this level of description our problem resembles a number of familiar philosophical skepticisms (concerning induction, the existence of an external world, the existence of other minds, etc.) and it is tempting to answer our problem using what have become familiar responses to skepticism: one argues either that agents can acquire what is sought entirely *a priori*, or that agents need not acquire it at all (the externalist response). We will take up these responses in turn.

The first thing to note is that our problem is *not* general inductive skepticism.[35] It has recently become more acceptable to suggest that agents have *a priori* warrant to believe that the universe is regular and that some properties are projectible.[36] Our concern is that even if *a priori* warrant to believe in a general regularity principle is available, that does no good helping agents discern *which* are the projectible properties. As Goodman put it in distinguishing his new problem of induction from the old, "To say that valid predictions are those based on past regularities, without being able to say *which* regularities, is thus quite pointless." (1979, p. 82)

In the metaethics literature, most of those who hold that there is an objective reasons relation assume that its extension can be determined *a priori*.[37] I have no way of showing that that is impossible. But our general result shows how difficult a position this is to maintain, at least for the case of theoretical reasoning. Our arguments to this point indicate that there can be a single, objective evidential favoring relation whose extension is determinable by agents only if it is possible to work out a list of preferred properties *a priori*.[38]

Now it certainly may feel like an *a priori* fact 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; perhaps we can just perceive that sharing

[35]It is also not the problem of underdetermination of theory by evidence. In an underdetermination of theory by evidence setup, you give me a body of evidence and a hypothesis and I construct another, incompatible hypothesis that is equally well-supported by the evidence as yours. Our general result, on the other hand, is set up such that you give me a body of evidence and two hypotheses, and I demonstrate that neither of the hypotheses *you gave me* is favored over the other by that evidence alone.

[36]See, for instance, (Wright 2004) and the final chapter of (Fumerton 1995). Wright acknowledges the point I am about to make at p. 185, n. 17.

[37]See, for instance, Scanlon's comparison between judgments about reasons and judgments about arithmetic at his (1998, Ch. 1, Section 11).

[38]One might object that in order to determine what his evidence favors an agent need not form beliefs about certain properties' having special status; he need only draw conclusions about his evidence that match up with the biases that would follow from a particular property list. (Keep in mind that if the evidential conclusions drawn are substantive, our general result tells us that there must *be* such biases in the favorings.) But presumably if an agent can determine which evidence favors which hypotheses he is at least capable of *noticing* that these favorings are treating some properties differently from others. Thus if it is possible to determine evidential favoring relations *a priori* it is also possible to work out on an *a priori* basis which properties are preferred. This justifies the claim in the main text.

particular features makes 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 probable than others. But any attempt to fill out these proposals even slightly runs into immediate problems, and it's worth reiterating a few well-worn points about where those problems lie.

First, while it may seem that our intuition immediately rejects "disjunctive" or "logically complex" predicates like "grue,"[39] we are happy to work with predicates like "bilaterally symmetrical" whose definition in terms of "has a fin on the left" and "has a fin on the right" is logically complex. It is sometimes suggested that intuition rejects logically complex predicates whose definition involves a *spatio-temporal* property in a disjunctive way. Yet for 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 creatures with full legs before the Fall but has applied to creatures with only vestigial legs 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 of graduate school I could be found at home.

Moreover, the property picked out by "grue" isn't really disjunctive. Assuming (with Lewis) 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.[40] What's disjunctive is the definition of "grue" in terms of our more standard color predicates, but that's relevant only 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 *a priori* property prioritization is about. "Grue" was Goodman's expedient counterexample intended to demonstrate that the formal evidential theories of his day lacked a particular general feature. Some philosophers respond to 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 in 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" help establish that a three-place favoring relation requires a list of preferred properties and that that list cannot be determined from our evidence.[41] This opens up a broad challenge of explaining how any property

[39] 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 was first observed before t" for some specific time t.

[40] There are those who hold that the term "property" should only properly speaking be used for the subset of properties we have described as "natural." Throughout this article I have followed Lewis in being liberal about properties (for instance, I assume that for every set there is a property of belonging to that set) and then designating some of them as "natural" or "special." To avoid terminological conflict one could replace all my uses of "property" as the sort of item denoted by a predicate with uses of "set."

[41] Hand-tailored, "gruesome" predicates play an important role in the proof of our general result in Appendix A.

does or doesn't get placed on the list on an *a priori* basis, not just how properties like grueness may be excluded.[42]

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," "regularity," "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 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.[43] Newton and others had to *discover* that the same physical laws apply throughout the universe, and that spatio-temporal location is irrelevant to grouping into physical kinds.[44]

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 a determinable fact of the matter about which hypotheses are favored by which evidence is thereby committed to a very strong conception of the *a priori*.[45] After all, the proposal is that given two

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

[43]See (Aristotle 1984, Bk. 1, Ch. 3).

[44]*A priori* judgments about the uniformity of various possible arrangements of the universe are also suspect. It might seem intuitive that the most uniform possible universe would be one in which all particles had identical velocities and were exactly the same distance from their nearest neighbor. But apparently *a priori* considerations from thermodynamics deem this a highly unlikely arrangement.

[45]Earlier I described our general result as showing that a body of evidence does not contain enough information to substantively favor one hypothesis over another. Those who believe that the projectible property list can be discerned *a priori* may take issue with this characterization,

situations in which one's total evidence says identical things *about* two predicates and says identical things *using* those predicates we may nevertheless be able to determine on *a priori* grounds that in one situation the predicate is projectible while in the other it is 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. But our general result shows that this isn't what we do at all; any such general procedures would wind up in Pedro's predicament. 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.[46] No one has ever put forward a serious story about how it is we might do this. Moreover, to the extent that any philosophers have offered theories of what the "specialness" of projectible properties consists in, those theories (such as Lewis's "naturalness" account) have all made the identity of the special properties an *a posteriori* fact.

Of course, we still have the datum that our natural faculties and cultural heritage emphasize certain properties over others. It may be that while our intuitions fall short of identifying some absolute truth about a list of special properties underlying an objective evidential favoring relation, we nevertheless have *a priori* warrant to take them as a *starting point* for reasoning on the basis of our evidence as we work to improve our understanding of the world. For example, it may 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 justifiably altered those groupings over time. But this is another view that makes evidential favoring relative to a subjective factor (the agent's starting point), as is revealed by the way it violates Uniqueness. For instance, consider a rational species with a different phenomenology from our own.[47] Early in this species's scientific development, its members will still categorize the world in a very different way than we do, and so the same evidence will favor different hypotheses (and justify different doxastic attitudes) for them than it will for us.

---

depending on how they understand "informational content." If anything that can be drawn out from a proposition by *a priori* means counts as part of the informational content of that proposition, then perhaps the apriority of the projectible properties list entails that a body of evidence contains enough information to favor one logically independent hypothesis over another. But this strikes me as a terminological issue, having to do with (for instance) whether one wants to say that all the informational content of the arithmetic of natural numbers is contained within the Peano axioms.

[46]In calling our procedures for detecting natural properties "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 expressible using general *principles.* As we noted in Section 2, our arguments apply just as well to particularists who hold 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.

[47]Or different subsets of early humans, if their phenomenology was not uniform.

5. Hard Externalism

We have established that if there is an objective, three-place evidential favoring relation that goes beyond entailments, agents will be able to determine when it obtains only if they can determine a list of special properties (or something equivalent) *a priori*. In the last section, I argued that it is very implausible that they can do so. But perhaps agents cannot, and need not, determine when the evidential favoring relation holds. Perhaps it is enough that there *is* such a relation out there in the world, and that it is possible to reason in ways that match up with what it recommends.

To be clear: the position we will consider in this section concedes that it is impossible *even in principle* for agents to work out that their evidence favors one hypothesis over another (except in special cases of entailment). While agents may regularly transition from belief in a body of evidence to belief in a hypothesis that that evidence in fact favors, it won't be possible for them in general to access facts about whether those transitions are in accord with the true favoring relation and they will not in general be justified in forming beliefs about those facts.[48] This position, which I call "hard externalism," is extreme even among the externalist views one currently finds in the literature. Hard externalism goes beyond the claim that agents may be unable to discern what their evidence *is* to the further claim that even if they can discern what their evidence is they may be unable to discern what that evidence *favors*.[49] It winds up denying commonly held views about reasons, for instance that reasons facts (that is, favoring facts) are necessary. For the hard externalist, what favors what depends on an empirical list of natural properties, and so varies from world to world.[50]

My guess is that many philosophers who believe in objective evidential favoring will find the prospect of favoring facts inaccessible to agents immediately unpalatable—but those philosophers will tend to be internalists on other fronts already. Committed externalists, on the other hand, may be glad to find yet one more area of epistemology that they can be externalists about. I know from experience presenting this material that there is little I can say to dissuade externalists from hard externalism. But in case anyone remains undecided between internalism and externalism, I will now explain what makes me uncomfortable about hard externalism.

Consider the debates between Copernican and Ptolemaic astronomers about their respective models of the universe. At some point proponents of the Copernican view judged that a heliocentric theory was favored by the available evidence over the Ptolemaic options, and we think the Copernicans were correct in their assessment of

[48] I realize there is an externalist reading of "access" on which agents *will* have access to the favoring facts I am discussing—there are externalist readings of terms like "discern," "determine," and "guide" as well. However, what I say should be read as using these terms in the way epistemological internalists use them. There must be such an internalist reading of "access" (for instance) so that we can distinguish internalists from externalists by saying that the former require agents to have access to certain epistemic properties of their beliefs when those properties obtain.

[49] Our paragon contemporary externalist Williamson, for instance, explicitly endorses the former limitation on discernment in his (2000), but I have never seen him endorse the latter.

[50] In this sense, hard externalists may think of evidential favoring as a four-place relation, with the fourth relatum being the projectible properties list in a particular world. Notice, though, that this sort of relativization leaves evidential favoring objective and maintains the Uniqueness Thesis, at least among agents in the same world.

that evidence.[51] Yet both parties in that debate had serious misconceptions about the fundamental properties and categories of astronomy.[52] On the externalist story about evidential favoring, we should try to understand what the evidence actually favored in the Ptolemaic/Copernican debate by consulting *our* list of what the relevant natural properties are, determining what evidential favoring relation those properties give rise to, then judging whether the Copernicans had the evidence on their side according to that relation. But this is to indulge in anachronism: it seems to me that the Copernicans not only *had* the evidence on their side, but also were able to *determine* that the evidence was on their side—a determination they did not make using an evidential relation derived from the modern natural property list.[53]

Hard externalism introduces a discontinuity between our view of reasoning in times past and our view of reasoning in our own. Possessed of a natural properties list that we take to be correct, we think that while *they* in the past proceeded with a false understanding of what their evidence favored, *we* (because we have the true property list) are under no such illusions. *Our* evidential reasoning is correct, while theirs just seemed to be so. But our natural properties list is not perfect either. We need nothing as strong as a pessimistic meta-induction to convince us of that; in a number of contemporary empirical sciences (subatomic physics, phylogeny, etc.) current categorization schemes yield enough inconsistencies and failures to explain the data that practitioners recognize those schemes cannot be entirely right. So reasoners a short time from now are apt to think about us just the way we think about the Copernicans; perhaps they will even conclude that we were not justified in concluding that the Copernicans' evidence supported heliocentrism. It *seemed* to the Copernicans that it did, it *seems* to us that it did, and we can be confident only that God (in possession of the actual natural properties list) has access to the truth of the matter.

While I have no conclusive argument against this view, I can't get myself to believe it. Past some point the astronomical evidence favored heliocentrism over the Ptolemaic theory; the Copernicans believed that, and they were justified in that belief by considerations to which they had access. Moreover, I am justified in making the assertions in the previous sentence even though my own property list remains flawed (and I am justified in saying that I am so justified). The fact that the Copernicans' evidence favored heliocentrism plays an important role in explaining why astronomy proceeded the way it did. An objective evidential favoring relation that we seek yet do not find threatens to become explanatorily and normatively inert. At each historical stage agents reckon what they have reason to believe by their current understanding of the world. At the current stage, the only natural properties list we can appraise evidence and hypotheses with respect to is the list we have before us (not some objective list to which we lack access). And this is

[51]Copernicus himself probably did not possess sufficient evidence to favor the heliocentric approach over the Ptolemaic options; by "the Copernicans" I mean to include proponents of Copernicus's view (such as Kepler and Galileo) who succeeded him in the history of astronomy. (For details about the development of the evidence, see (Kuhn 1957, Ch. 6).

[52]Copernicus, for example, treated the sun very differently from the other stars and took all of the latter to dwell on a single "sphere".

[53]If you object to this example because you think the natural property list in use at the time was *close enough* to correct that the Copernicans could be on to the real evidential favoring relation, feel free to substitute the presocratic example we'll get to a little bit farther on.

the list we *should* use. While there may be an objective evidential favoring relation somewhere out there of the sort the hard externalist envisions, that relation cannot guide or explain the ongoing story of human evidential reasoning.

It may be suggested that these complaints can be accommodated by borrowing a distinction from another area of philosophy—the distinction between subjective and objective "oughts." The suggestion will be that agents *subjectively* ought to reason according to the favoring relation dictated by materials to which they have access, while they *objectively* ought to reason according to the favoring relation generated by the true natural properties list.

In response to this suggestion, consider Thales and Anaximenes' debate about whether the original element from which all matter was constituted was water or air.[54] In the course of this debate, Anaximenes adduced evidence involving condensation in support of his view. Personally, I cannot find any good sense of "ought" in which Thales and Anaximenes *ought* to have adjudicated their debate, and assessed the significance of the condensation evidence, in light of the true quantum mechanical story concerning matter's natural properties. Moreover, it is hard for me to see why a theory that postulated objective "oughts" applying even to cases like these would be better motivated than a theory that simply said we ought objectively to believe whatever's true.[55]

Finally, let me point out that hard externalism elevates a particular set of empirical facts to a peculiar place in our epistemology. Typically with an empirical proposition, an agent who has a justified yet false belief about the proposition is justified in shaping his other views in line with that belief. But according to hard externalism, an agent who has justified false beliefs about the natural properties is not justified in drawing hypotheses from evidence in line with what those beliefs would suggest. Whether evidence provides reason to believe particular hypotheses over others is dictated by the true natural properties list, not by the agent's justified beliefs about the contents of that list. This has strange side effects, for instance a version of Jonathan Vogel's (2000) bootstrapping problem for externalism.

Consider Roxy, a scientist in the time of Lavoisier whose total evidence favors neither the phlogiston theory nor the oxygen theory over the other. Roxy imagines an experiment for which the two theories would predict different results. Then she projects that the result predicted by the oxygen theory would occur. Since oxygen is in fact a natural kind, oxygen claims are indeed projectible, Roxy has reason to believe the predicted result would occur, and she is justified in believing that it would. Roxy then imagines a number of such experiments, in each case drawing the justified conclusion that the experiment would come out as the oxygen theory would predict. Once she has enough such justified beliefs about how particular types of experiments would come out, Roxy is (by the hard externalist's lights) justified in concluding that the oxygen theory is correct, because it makes so many correct predictions about experiments. This despite the fact that (by stipulation) at the

[54]For a brief historical overview and references, see (Curd 2008).

[55]It's worth noting that authors (from (Williams 1981) through (Kolodny and MacFarlane 2010)) typically introduce objective "oughts" using a story (such as an episode of advice-giving) that gets traction from the fact that one party is aware of how things actually stand while the other is informationally impaired. Once we are working with a hard externalist account on which it is impossible even in principle for either party to have access to the objective truth about evidential favoring, it is unclear what role objective "oughts" can play in our deliberations, conversations, etc. and what evidence we have for their existence.

beginning of our story Roxy's evidence favored neither theory over the other, and the fact that over the course of the story she has imagined a number of experiments but not conducted any of them! While this type of maneuver would be impossible for most empirical beliefs, the special role given truths about the natural properties in the hard externalist's view allows agents to justify beliefs about the natural properties without gathering evidence in their favor.

As I said at the outset, I doubt that any of this will convince the committed externalist to turn away from hard externalism. While hard externalism gives special status to natural property facts, this is not much different from the special status externalists typically give to empirical facts about the reliability of our faculties. By now externalists have many replies prepared to well-known problems with their view like Vogelian bootstrapping. So I will simply leave the hard externalist with a question. Through a combination of evolutionary winnowing, cultural change, and scientific progress, it is very likely that 21st-century humans employ a natural property list that hews very closely to the true list of projectible properties in the world.[56] If so, our practices of projection are highly reliable, and the externalist may say that this reliability is crucial to the fact that our projections are justified. But our practices are not perfect, and the question is: In cases in which our generally reliable current understanding of the natural properties leads us to project a property that does not match the true list, are we justified in the conclusions we draw from our evidence?[57] A "no" answer remains faithful to the hard externalist position. But if we answer "yes," we are admitting that what an agent has reason to believe can vary with a subjective factor—the understanding of natural properties with which his evolutionary and scientific history has endowed him. According to this answer two species (or human cultures) with natural properties lists that slightly diverged but were both very close to the truth could disagree on which hypotheses a given body of evidence favored and both be justified. Uniqueness would be denied once more.

6. Subjectivism

Philosophers often think of "grue" the way they think of the Liar Paradox: as something no one has managed to resolve satisfactorily and so is equally a problem for everyone. Somehow the *ubiquity* of the problem reduces its significance. This stance allows proponents of, for instance, *a priori*-discernible objective reason relations to beg off of questions about *how* the faculties in question might work with a "We don't know how they do it—but they must do it somehow!" Yet the problems indicated by language dependence do not cut equally against all views. I have argued that the view on which agents determine objective evidential favoring relations *a priori*—a view that is the default in many metaethical discussions—requires an implausibly strong assessment of human reasoning capacities. Externalists, on the other hand, may accommodate the informational deficiencies of evidential propositions by extending their view to a hard externalism about relations between hypotheses and evidence. This creates new difficulties for their view, but

[56]The suggestion that our species could not have survived for long without a roughly correct grasp of which properties are projectible dates back at least to (Quine 1969).

[57]Compare: In the rare cases in which a generally reliable faculty of perception goes wrong, are we justified in the beliefs that result?

those difficulties are similar to bullets externalists have been willing to bite for some time.

There is, however, a view that not only *accommodates* but in fact *predicts* the fact that a body of evidence alone cannot substantively favor one hypothesis over another. The crux of this view, which I call "subjectivism," is to deny that there is objective, three-place evidential favoring in cases without entailment and admit that in those cases evidential favoring is relative to a fourth relatum: a subjective factor beyond the agent's evidence that plays a role in favoring some properties over others. That subjective factor may be a subject's theory of natural properties, it may be the predicates contained in the subject's native language, or it may be something else.[58] It may depend on the subject's society, his upbringing, his biology, or any of a number of contingent factors.[59] It might work by bringing certain pragmatic interests to bear, by highlighting certain questions as more important than others, or by prioritizing the ruling out of certain alternatives. The key point is that it be *accessible* to agents, so that an agent does not have the externalist's problem of being unable to determine when a body of evidence favors a particular hypothesis for him.

One well-known subjectivist view in epistemology is Subjective Bayesianism. That view certainly has its flaws, and I've done nothing to argue for it here, but it provides a nice example of the sort of structure I'm talking about. Subjective Bayesianism models an agent as assigning an initial numerical distribution of credences over propositions (called a "prior") that is not determined by his total evidence. Like a natural properties list or a preferred language, a prior does the work of making some properties more projectible than others. Priors vary from agent to agent, but there are rules governing which bodies of evidence favor which hypotheses *relative* to a given prior, and these rules yield substantive favoring relations. Since each agent can discern his own prior, he can access the evidential favoring facts that apply to him.

Notice that Subjective Bayesianism retains an important objective factor. *Relative* to an agent's prior, evidential favoring relations among hypotheses and evidence are completely specified. Whether "green" or "grue" is more projectible for an agent becomes a question of whether a particular subset of his credences meet various *a priori*, mathematical conditions that allow green evidence to favor green hypotheses differently than grue evidence favors grue hypotheses.[60] Thus there are

---

[58]Notice that it *cannot* just be a further set of propositions expressed in the same observational language as the evidence. Our general result could quickly be extended to show that if a substantive evidential favoring relation is determined relative to an extra set of propositions representable in the formal language containing h_1, h_2, and e, that relation must privilege a set of properties *prior* to the influence of either the evidence or that extra set of propositions. (To get a shortcut sense of why, imagine that the extra propositions functioned by being added to the agent's total evidence set.) Compare the discussion of "background theories" in (Sober 1988, Section 2.5).

[59]As Michael Bratman pointed out to me, the fourth relatum need not be an aspect of the agent whose evidence we are considering. For instance, when one agent assesses which hypotheses are favored by another agent's evidence, the fourth relatum may be determined by features of the assessing agent. The fourth relatum may also be be determined by a broader community of people around the agent possessing the evidence. All of these options, however, contribute a "subjective" factor in the sense I defined in the introduction—instead of evidential favoring's being purely a relation among evidence and hypothesis *propositions*, that relation varies with facts about *agents*—so I will count them as "subjectivist" approaches.

[60]For a precise set of such mathematical conditions see (Eells 1982).

still standards for good reasoning—biologists, for example, are perfectly rational to go on developing phylogenetic trees according to the categories that seem perspicuous to us and are deeply rooted in our way of understanding the world—it's just that subjects shouldn't think what they're doing is driven *entirely* by their evidence.[61]

Other formal theories of evidential support tend not to be so explicitly subjectivist, but their proponents will tacitly admit that subjective factors play a role. For example, philosophers who endorse scientists' ranking hypotheses using maximum entropy will respond to complaints of language-dependence by saying that 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. As we saw in the oxygen/phlogiston example, the language that best carves nature at its joints is often not a given but is instead part of the scientific problem investigators are trying to solve.[62] One also needs to understand the costs of a view that relativizes evidential favoring to a fourth, subjective relatum. As we have seen, such a view violates Uniqueness: two agents with the same total evidence may be rationally permitted to have different beliefs. In fact, if the relevant subjective factors (such as their priors) differ in the right ways, they may be rationally *required* to have different beliefs.

Subjectivism encourages us to abandon a view of science that was discarded some time ago by most philosophers of science but lingers on among many epistemologists. The early Carnap thought of scientific reasoning as 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 subjectivism recommends that we think of science as being more like an "argument" in the everyday sense: an activity that carries on through time, over the course of which positions change, information is introduced, and new views develop. It may be that relative to the course of scientific inquiry at a particular stage in history, or to a particular set of questions on which that inquiry has focused, or to a particular set of hypotheses under consideration, there are objective requirements on what a scientist should conclude from a body of evidence. But our evidence alone 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*.

Subjectivism also has implications for how we think about normativity. Objective reasons theorists tend to take facts about reasons as conceptually or explanatorily prior to rational constraints on agents' thought processes. Yet the norms of subjecitivst views such as Subjective Bayesianism are most straightforwardly read

[61]I mentioned earlier that the failure of evidential propositions to substantively favor hypotheses is a *prediction* of subjectivist theories. In Subjective Bayesianism it's easy to show that any favoring (or "confirmation") relations among logically independent hypotheses and evidence will depend on values of the prior, which are not responsive in any way to an agent's evidence.

[62]I want to emphasize that nothing in this article is meant to deny the metaphysical thesis that there are such things as natural properties or the epistemological thesis that determining the natural property list is an important part of the scientific enterprise. I have argued only that the mere *existence* of natural properties does not automatically solve the epistemic problem presented to an agent trying to work out which empirical hypotheses are favored by his evidence.

as constraints of internal consistency or coherence among an agent's attitudes. This suggests that our theory of normativity should move *from* a story about rationality *to* a story about reasons;[63] to the extent there is something interesting to say about reasons it will arise out of an account of what it is rational for an agent to believe given his subjective point of view.[64]

Historically, Subjective Bayesianism was a response to Objective Bayesian theories such as those of Keynes (1929) and Carnap (1950). We might think of the Objective Bayesians' formal recipes for calculating probabilities of hypotheses relative to evidence as attempts to work out objective facts about what an agent's evidence gives him reason to believe. 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 understand that there simply isn't enough information in an evidential proposition to support something as strong as a probability distribution over a hypothesis space. This article began in my attempts to determine whether we might objectively obtain something weaker, like a partial ordering over the hypotheses. It turns out there isn't enough information in a body of evidence to give us even that.[65]

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.

Bealer, G. (1992). The incoherence of empiricism. *Proceedings of the Aristotelian Society, Supplementary Volumes 66*, 99–143.

Broome, J. (2004). Reasons. In R. J. Wallace, P. Pettit, S. Scheffler, and M. Smith (Eds.), *Reasons and Value: Themes from the Moral Philosophy of Joseph Raz*, pp. 28–55. Oxford: Oxford University Press.

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

Chandler, J. (ms). Contrastive support: Some competing accounts. Forthcoming in *Synthese*.

[63]Think here of the difference between a traditional Kantian story about normativity and that of, say, Joseph Raz.

[64]One might think that we need objective reason facts to explain the normativity of rational requirements. But here Kolodny's recent work (in his (2005), (2008), and elsewhere) offers a cautionary tale. A defender of objective reasons, Kolodny cannot find any way for them to generate subjective constraints and so argues that the latter must be abandoned. But if subjectivism is correct and subjective rationality constraints are crucial to the story we tell about everyday evidential relations, Kolodny's work dims the prospects for fitting objective reason relations into that story at all. (On this point see also (Jackson 1991).)

[65]For advice, feedback, and conversation I am grateful to audiences at the Australian National University, the University of Melbourne, the University of Sydney, the University of Minnesota Duluth, Northwestern University, the first St. Louis Annual Conference on Reasons and Rationality, and the 2010 Formal Epistemology Workshop; to my Fall 2009 graduate seminar at the University of Wisconsin-Madison, and especially to Matt Kopec and Casey Helgeson for checking the proofs; and to Fabrizio Cariani, David Chalmers, Andy Egan, Branden Fitelson, Malcolm Forster, Jessica Gelber, Lloyd Humberstone, Kevin Kelly, Andrew McGonigal, Jake Ross, Jonathan Schaffer, Russ Shafer-Landau, Ted Sider, Elliott Sober, Joel Velasco, and Peter Vranas.

Curd, P. (2008). Presocratic philosophy. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2008 ed.).

Dancy, J. (1993). *Moral Reasons*. Oxford: Blackwell.

Earman, J. (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: The MIT Press.

Eells, E. (1982). *Rational Decision and Causality*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press.

Enoch, D. (2010). The epistemological challenge to metanormative realism: how best to understand it, and how to cope with it. *Philosophical Studies 148*, 413–438.

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.

Fumerton, R. A. (1995). *Metaepistemology and Skepticism*. Lanham, MD: Rowman & Littlefield.

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 (II). *Mind 54*, 97–121.

Hooker, C. A. (1968). Goodman, 'grue' and Hempel. *Philosophy of Science 35*, 232–247.

Jackson, F. (1991). Decision-theoretic consequentialism and the nearest and dearest objection. *Ethics 101*, 461–482.

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.

Kelly, T. (2008). Evidence. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2008 ed.).

Keynes, J. M. (1929). *Treatise on Probability*. London: MacMillan and Co., Limited.

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

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

Kolodny, N. and J. MacFarlane (2010). Ifs and oughts. Forthcoming in *The Journal of Philosophy*.

Kuhn, T. S. (1957). *The Copernican Revolution*. New York: MJF Books.

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.

Novack, G. (2010). A defense of the principle of indifference. Forthcoming in *Journal of Philosophical Logic*.

Parfit, D. (2006). Normativity. In R. Shafer-Landau (Ed.), *Oxford Studies in Metaethics*, Volume 1, pp. 325–380. Oxford: Clarendon Press.

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

Raz, J. (1999). *Engaging Reason.* Oxford: Oxford 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.

Shafer-Landau, R. (2009). A defence of categorical reasons. *Proceedings of the Aristotelian Society 109*, 189–206.

Sober, E. (1988). *Reconstructing the Past: Parsimony, Evolution, and Inference.* Cambridge, MA: The MIT Press.

Vogel, J. (2000). Reliabilism leveled. *The Journal of Philosophy 97*, 602–623.

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.

Wright, C. (2004). Warrant for nothing (and foundations for free)? *Supplement to the Proceedings of the Aristotelian Society*, 167–212.

Appendix A. Proof of the General Result

We'll begin with some definitions, with technical terms marked in **bold** when they're introduced. A **language** $\mathcal{L}$ will be a first-order language containing only constants, one-place predicates, and the standard sentential connectives (no variables or quantifiers). $P(\mathcal{L})$ will be the number of predicates in $\mathcal{L}$; $C(\mathcal{L})$ will be the number of constants. We will limit ourselves to languages with finite $P(\mathcal{L})$ and $C(\mathcal{L})$ values. An **atomic sentence** applies a predicate to a constant; a **literal** is an atomic sentence or its negation. We will imagine that each language $\mathcal{L}$ comes with a total ordering on its atomic sentences (not too difficult considering that the number of atomic sentences is finite).

A **state description** is a conjunction of literals in which each atomic sentence appears exactly once (perhaps negated) and the atomic sentences appear in order. We will often abbreviate "state description" as "**sd**". The number of atomic sentences in $\mathcal{L}$ is $P(\mathcal{L}) \cdot C(\mathcal{L})$; this is also the number of conjuncts in a state description. The number of state descriptions in $\mathcal{L}$ is $2^{P(\mathcal{L}) \cdot C(\mathcal{L})}$. We will imagine that $\mathcal{L}$ also comes with a total ordering on its state descriptions (again not difficult because there are finitely many of them). A sentence $x \in \mathcal{L}$ that is a disjunction of state descriptions in order in which no state description appears more than once is in **disjunctive normal form**, or just "**dnf**".

Each language will also come with an **interpretation** that maps its sentences onto propositions. Propositions will be understood as sets of possible worlds. I will write μx for the proposition represented by sentence $x \in \mathcal{L}$. 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 $\mu x = \mu x'$. (Note that synonymy is transitive.) 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 symbolized $x \vdash y$.[66] We will assume that $\vdash$ interacts with the sentential connectives in the way typically taught in introductory logic classes. A standard result from such classes shows that for any non-contradictory sentence in $\mathcal{L}$, there is a unique dnf sentence of $\mathcal{L}$ that is syntactically equivalent.[67]

We will say that language $\mathcal{L}$ is **faithful** just in case the following four conditions are met:

(1) For each possible world β, there exists exactly one sd $s_i \in \mathcal{L}$ such that $\beta \in \mu s_i$.
(2) For each sd $s_i \in \mathcal{L}$, there exists at least one possible world β such that $\beta \in \mu s_i$.
(3) Each non-contradictory sentence $x \in \mathcal{L}$ represents the union of the propositions represented by the sds occurring in x's dnf equivalent.
(4) Each contradictory sentence in $\mathcal{L}$ represents the empty proposition.

I call such languages "faithful" because their syntax faithfully represents imporant aspects of the semantics of the propositions represented (assuming a classical "$\vDash$" relation). For instance, we have:

Result 1. A sentence in a faithful language represents the empty proposition just in case it is a contradiction.

Proof. Since the fourth faithfulness condition guarantees that all contradictions in a faithful language represent the empty proposition, we need only demonstrate that if $x \in \mathcal{L}$ is not a contradiction then μx is non-empty. So suppose x is not a contradiction. By the third faithfulness condition, μx is the union of the propositions represented by the sds in x's dnf equivalent. Since x is not a contradiction, there will be at least one such sd. By the second faithfulness condition, that sd will represent a proposition containing at least one possible world. So x is nonempty. □

Result 2. A faithful language $\mathcal{L}$ possesses the **entailment property**: for any $x, y \in \mathcal{L}$, $x \vdash y$ just in case $\mu x \vDash \mu y$.

Proof. **Forward Direction:** Suppose $\mathcal{L}$ is faithful and $x \vdash y$. If neither x nor y is a contradiction, then every sd appearing in the dnf equivalent of x will appear in the dnf equivalent of y. By the third faithfulness condition, this means that $\mu x \subseteq \mu y$, so $\mu x \vDash \mu y$. If x is a contradiction, μx is the empty proposition by the fourth faithfulness condition, so $\mu x \vDash \mu y$. If y is a contradiction and x is not, $x \nvdash y$, which contradicts our supposition.

Reverse Direction: Suppose $\mathcal{L}$ is faithful and $\mu x \vDash \mu y$. Suppose further that neither x nor y is a contradiction. Finally, suppose for *reductio* that $x \nvdash y$. Then there's an sd in the dnf equivalent of x that is not in the dnf equivalent of y—call it s_i. By the second faithfulness condition, there exists a possible world $\beta \in \mu s_i$, and by the first faithfulness condition β does not appear in a proposition represented by any other sd of $\mathcal{L}$. So since s_i does not appear in y's dnf equivalent, $\beta \notin \mu y$.

[66] I apologize for leaving it largely to the reader to disambiguate use from mention, predicates and constants from metavariables, etc. throughout these appendices.

[67] Notice that what I call a "language" is actually a package of three things: (i) the set of formal sentences usually called a "language" in formal logic; (ii) an interpretation assigning propositions to those sentences; and (iii) a syntactic consequence relation.

But $\beta \in \mu x$, so $\mu x \nsubseteq \mu y$, so $\mu x \not\models \mu y$ and we have a contradiction. So we must have $x \vdash y$.

Now suppose x is a contradiction. Then $x \vdash y$. Finally, suppose y is a contradiction but x is not. By the fourth faithfulness condition, μy is the empty proposition. By Result 1, μx is nonempty. This means $\mu x \not\models \mu y$, which contradicts our supposition. So this case is impossible. ☐

Result 3. In a faithful language $\mathcal{L}$, $x \dashv\vdash y$ just in case $\mu x = \mu y$.

Proof. By the entailment property, $x \dashv\vdash y$ just in case $\mu x \mathrel{=\!\mid\models} \mu y$. $\mu x \mathrel{=\!\mid\models} \mu y$ means that every possible world in μx is in μy, and *vice versa.* This occurs just in case μx and μy are the same proposition. ☐

I will leave it to the reader to demonstrate that a faithful language matches syntax with semantics in other nice ways, for instance:

- For any $x \in \mathcal{L}$, $\mu({\sim}x)$ is the complement of μx.
- For any $x, y \in \mathcal{L}$, $\mu(x \vee y)$ is the union of μx and μy.
- For any $x, y \in \mathcal{L}$, $\mu(x \,\&\, y)$ is the intersection of μx and μy.

Notice that by Result 1, no sd of a faithful language represents the empty proposition. So each sd in a faithful language represents a logically possible proposition. This means that the properties represented in a faithful language's predicates are logically independent over the set of objects named by that language's constants; that is, for any object named it is logically possible that that object have any combination of the properties represented in the language.

We will assume that logically independent properties are plentiful in the following sense:

Availability of Independent Properties: Given any faithful language $\mathcal{L}$, there exists another language $\mathcal{L}'$ such that:
(1) $C(\mathcal{L}') = C(\mathcal{L})$ and $P(\mathcal{L}') = P(\mathcal{L}) + 1$.
(2) There exists a one-to-one mapping c from the constants of $\mathcal{L}$ to the constants of $\mathcal{L}'$ and a one-to-one mapping p from the predicates of $\mathcal{L}$ to a set consisting of all but one of the predicates of $\mathcal{L}'$ such that for any constant a and predicate F in $\mathcal{L}$, if $a' \in \mathcal{L}'$ is $c(a)$ and $F' \in \mathcal{L}'$ is $p(F)$ then $\mu F'a' = \mu Fa$.
(3) $\mathcal{L}'$ is faithful.

The idea here is that given any faithful language $\mathcal{L}$, there exists another language $\mathcal{L}'$ that represents all the same objects and properties plus one additional property that is logically independent of all the properties represented in $\mathcal{L}$ (over the set of objects named in $\mathcal{L}$).

Since we are working with languages with finite sets of constants and predicates, the Availability of Independent Properties is a reasonable assumption. It requires the supply of logically independent properties to be inexhausible. This seems obvious to me, but in case you need a bit of convincing, consider a case in which the constants of $\mathcal{L}$ represent concrete objects and the properties represented in $\mathcal{L}$'s predicates say nothing about the masses of those objects. We could create the language $\mathcal{L}'$ required by the Availability of Independent Properties by adding an additional predicate F to those in $\mathcal{L}$ such that Fa represents the proposition that a's mass in grams has a 3 in its first decimal place. We could then create another language $\mathcal{L}''$ by adding a predicate G to those in $\mathcal{L}'$ such that Ga represents a proposition about the mass's *second* decimal place. Clearly we could keep adding predicates expressing logically independent properties in this fashion and never exhaust our supply.[68] If our original $\mathcal{L}$ *does* contain a predicate expressing something about the mass of an object, we could always introduce new predicates concerning the object's momentum, or its charge, or.... In extreme cases, we could forgo properties of the object itself and describe properties of an object 10 meters to the north, or 11 meters, or....

The following result allows us to build one faithful language from another in another way:

Result 4. Suppose we have a faithful language $\mathcal{L}^1$ and we construct a language $\mathcal{L}^2$ meeting the following conditions:

(1) $\mathcal{L}^2$ has no more sds than $\mathcal{L}^1$ has.
(2) The sds of $\mathcal{L}^1$ are partitioned into $2^{P(\mathcal{L}^2)\cdot C(\mathcal{L}^2)}$ sets such that each $\mathcal{L}^1$ sd belongs to exactly one such set. Each set is assigned to a different sd of $\mathcal{L}^2$. Each sd of $\mathcal{L}^2$ is then interpreted as the proposition that is the union of the propositions represented by the $\mathcal{L}^1$ sds in its designated set.
(3) Each non-contradictory, non-sd sentence $x \in \mathcal{L}^2$ represents the union of the propositions represented by the $\mathcal{L}^2$ sds in x's dnf equivalent.
(4) Each contradictory sentence in $\mathcal{L}^2$ represents the empty proposition.

Then $\mathcal{L}^2$ will be faithful as well.

Proof. Trivial from the definition of faithfulness. □

We now introduce the three-place relation $f(\cdot,\cdot,\cdot)$ whose relata are sentences in our formal languages. $f(h_1, h_2, e)$ 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. Keep in mind that f is a symbol in our metalanguage—not a predicate in one of our object languages—indicating when a certain relation holds among object-language sentences (and more particularly among the propositions they represent). We will assume that for any given $e \in \mathcal{L}$, $f(\cdot,\cdot,e)$ is antisymmetric. That is, if $f(h_1, h_2, e)$ then it's not the case that $f(h_2, h_1, e)$. The following further facts about $f(\cdot,\cdot,\cdot)$ follow from our definition:

Result 5. The relation $f(\cdot,\cdot,\cdot)$ 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 $\mu e = \mu e'$, $f(h_1, h_2, e)$ just in case $f(h_1', h_2', e')$.

Proof. By definition, $f(h_1, h_2, e)$ 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 $f(h_1', h_2', e)$. □

Result 6. $f(\cdot,\cdot,\cdot)$ 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 $e \dashv\vdash e'$, $f(h_1, h_2, e)$ just in case $f(h_1', h_2', e')$.

[68] Even if quantum-mechanical requirements limit how many decimal places a mass in grams may have, these are not *logical* requirements, so they will not affect the *logical* independence of the properties represented by the predicates introduced.

Proof. Since $\mathcal{L}$ is faithful, $h_1 \dashv\vdash h_1'$ gives us $\mu h_1 = \mu h_1'$ by Result 3. 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 $f(h_1, h_2, e)$ just in case $f(h_1', h_2', e')$. □

As discussed in Section 1 of this article, it would be too demanding to expect a formal theory to detect evidential favoring among the sentences of just *any* formal language, for the language might not be expressive enough to capture the substructure of the evidence and hypotheses that caused the favoring relation to hold. So given a particular set of hypotheses and evidence, we restrict our attention to languages that are **adequate** for those relata. (When the relata under discussion are clear I will sometimes say just that a language is "adequate".) We have no formal definition of adequacy, but we do have a necessary condition on such a definition:

> If language $\mathcal{L}$ is adequate for a particular set of relata and language $\mathcal{L}'$ expresses all of $\mathcal{L}$, then $\mathcal{L}'$ is adequate for those relata as well.

I also believe it is unreasonable to expect a formal theory to pick up on evidential relations when our relata are expressed in a formal language that obscures entailment relations among the propositions represented by its atomic sentences. That is, I think we should work with formal languages whose atomic sentences represent logically independent propositions. For this reason we will restrict our attention not just to adequate languages but to faithful adequate languages. In Appendix B I will discuss this requirement further.

Result 7. Suppose we have a pair of hypotheses, some evidence, and a faithful language $\mathcal{L}$ adequate for those three relata. Then there exists a faithful language $\mathcal{L}'$ that is adequate for those relata with $C(\mathcal{L}') = 1$.

Proof. Suppose the predicates of $\mathcal{L}$ are $F_1, F_2, \ldots, F_{P(\mathcal{L})}$ and the constants of $\mathcal{L}$ are $a_1, a_2, \ldots, a_{C(\mathcal{L})}$. Language $\mathcal{L}'$ is constructed so as to have one constant a' (which we can think of as representing the $C(\mathcal{L})$-tuple consisting of the objects represented by $a_1, a_2, \ldots$). 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 $C(\mathcal{L})$. We assign $\mu F'_{j,k} a' = \mu F_j a_k$ for every $1 \leqslant j \leqslant P(\mathcal{L})$ and $1 \leqslant k \leqslant C(\mathcal{L})$.

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 $\mathcal{L}$. Moreover, the interpretation of $\mathcal{L}'$ will allow it to play the role of $\mathcal{L}^2$ to $\mathcal{L}$'s $\mathcal{L}^1$ in Result 4 (with a simple one-to-one mapping between sds of $\mathcal{L}$ and sds of $\mathcal{L}'$). So since $\mathcal{L}$ is faithful, $\mathcal{L}'$ is as well. □

We now define two conditions that $f(\cdot, \cdot, \cdot)$ may or may not possess:

Substantivity: There exists some faithful language $\mathcal{L}$ and $h_1, h_2, e \in \mathcal{L}$ such that:

(1) $\mathcal{L}$ is adequate for the propositions expressed by h_1, h_2, and e.
(2) $f(h_1, h_2, e)$.
(3) Neither the conjunction $h_1 \,\&\, h_2 \,\&\, e$ nor any conjunction that results from negating some of the conjuncts in that conjunction is a contradiction in $\mathcal{L}$.

(Since $\mathcal{L}$ is faithful, this last clause and Result 1 guarantee that h_1, h_2, and e represent logically independent propositions.)

For any $x \in \mathcal{L}$ and any permutation π of the predicates in language $\mathcal{L}$ (that is, any one-to-one mapping from the set of predicates in $\mathcal{L}$ to itself), we will let $\pi(x)$ denote the sentence in $\mathcal{L}$ that results from replacing each predicate symbol in x with its image under π. With this notation in place we define:

Identical Treatment of Predicate Permuations: For any e, h_1, and h_2 in adequate, faithful language $\mathcal{L}$ and any permutation π of the predicates of $\mathcal{L}$, $f(h_1, h_2, e)$ implies $f(\pi(h_1), \pi(h_2), \pi(e))$.

The following result draws out the implications of this last condition:

Result 8. Suppose $f(\cdot,\cdot,\cdot)$ treats predicate permutations identically. Then for any adequate, faithful 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 $\pi(e) \dashv\vdash e$, $\pi(h_1) \dashv\vdash h_2$, and $\pi(h_2) \dashv\vdash h_1$, then it's not the case that $f(h_1, h_2, e)$.

Proof. Suppose $f(\cdot,\cdot,\cdot)$ treats predicate permutations identically, and suppose we have e, h_1, h_2 in adequate faithful language $\mathcal{L}$ and a permutation π of the predicates of $\mathcal{L}$. Suppose further that $\pi(e) \dashv\vdash e$, $\pi(h_1) \dashv\vdash h_2$, and $\pi(h_2) \dashv\vdash h_1$. Finally, suppose for *reductio* that $f(h_1, h_2, e)$. Since $f(\cdot,\cdot,\cdot)$ treats predicate permutations identically, we have $f(\pi(h_1), \pi(h_2), \pi(e))$. By the equivalence condition (Result 6), we also have $f(h_2, h_1, e)$. But that violates the antisymmetry of $f(\cdot,\cdot,e)$, so we have a contradiction. □

We can now state our main result:

Result 9 (The General Result)**.** Assuming the availability of independent properties, the relation f cannot both be substantive and treat predicate permutations identically.

Overview of the Proof: The proof of our general result proceeds by *reductio*. We suppose that $f(\cdot,\cdot,\cdot)$ treats predicate permutations identically. We then suppose that (as substantivity requires) there exists a faithful, adequate language $\mathcal{L}$ with $e, h_1, h_2 \in \mathcal{L}$ such that $f(h_1, h_2, e)$ and e, h_1, and h_2 represent logically independent propositions. 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 7 to make our lives easier by assuming (without loss of generality) that $\mathcal{L}$ has only one constant.

The goal of the proof is to construct a faithful, adequate 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, $f(h_1^*, h_2^*, e^*)$. Yet there is a permutation π of the predicates of $\mathcal{L}^*$ that maps h_1^* to an equivalent of h_2^*, h_2^* to an equivalent of h_1^*, and e^* to an equivalent of itself. $\mathcal{L}^*$ is faithful and adequate for e^*, h_1^*, and h_2^*, so Result 8 tells us it's not the case that $f(h_1^*, h_2^*, e^*)$. 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 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 state description that is a disjunct of h_1^* but of neither of the other two is mapped to a distinct 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 construct $\mathcal{L}^*$ from $\mathcal{L}'$ by taking each state description that appears as a disjunct of e', h_1', or h_2' and giving it a synonym that is a 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, for example, 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 state description of $\mathcal{L}'$, but we give one of the h_1-only disjuncts a synonym that is the disjunction of *three* 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 outlines an algorithm for constructing a language $\mathcal{L}^*$ that witnesses the incompatibility of substantivity and the identical treatment of predicate permutations.

Proof of Result 9. Suppose for *reductio* that substantivity and the identical treatment of predicate permuations hold. Suppose also that faithful language $\mathcal{L}$ is adequate for the propositions expressed by $e, h_1, h_2 \in \mathcal{L}$, that $f(h_1, h_2, e)$, and that the propositions represented by e, h_1, and h_2 are logically independent. If $C(\mathcal{L}) > 1$, Result 7 allows us to construct a faithful, adequate single-constant 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 $f(h_1', h_2', e')$, 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 constant. The equivalence condition also allows us to assume without loss of generality that e, h_1, and h_2 are in dnf. (Since they are logically independent, none of these three sentences can be contradictions.)

It will help to introduce some notation at this point. We will be dividing the sds of $\mathcal{L}$ up by type, according to which of our three sentences of interest they are disjuncts of. For example, an sd of $\mathcal{L}$ will be described as an eh_2-sd if it is a disjunct of e and of h_2 but not of h_1. (Put another way, an eh_2-sd is an sd that entails e and h_2 but does not entail h_1.) $\#(eh_2)$ will be the number of distinct eh_2-sds. An sd that entails none of e, h_1, and h_2 will be described as a ϕ-sd. The logical independence of e, h_1, and h_2 guarantees that there is at least one sd of each of the eight possible types. (Once we construct $\mathcal{L}'$, an $e'h_2'$-sd will be a disjunct of both

e' and h'_2, $\#(e'h'_2)$ will be the number of such sds in $\mathcal{L}'$, and a ϕ'-sd will be an sd of $\mathcal{L}'$ that does not appear in e', h'_1, or h'_2; our notation will work in a similar way for $\mathcal{L}^*$.)

Our ultimate goal will be to construct a language $\mathcal{L}^*$; an $h^*_1, h^*_2, e^* \in \mathcal{L}^*$; and a permutation π that maps the sds of $\mathcal{L}^*$ to each other in the following fashion:

h^*_1-sd	to	h^*_2-sd	and *vice versa*
$e^*h^*_1$-sd	to	$e^*h^*_2$-sd	and *vice versa*
$h^*_1h^*_2$-sd	to itself		
$e^*h^*_1h^*_2$-sd	to itself		
e^*-sd	to itself		
ϕ^*-sd	to any ϕ^* sd		

Step 1: Construct language $\mathcal{L}'$. Ultimately the sds of $\mathcal{L}^*$ will be interpreted by matching them to the sds of $\mathcal{L}'$ one-to-one; our goal in constructing $\mathcal{L}'$ is to put the numerical relationships in place that will allow the π mappings described above. So we want to construct $\mathcal{L}'$ and $h'_1, h'_2, e' \in \mathcal{L}'$ such that

$$\#(h'_1) = \#(h'_2)$$
$$\#(e'h'_1) = \#(e'h'_2)$$

If the numbers of sds of the relevant types already line up in this way in $\mathcal{L}$, $\mathcal{L}$ can serve as our $\mathcal{L}'$ and we can skip to Step 2. But this may not be the case. If it is not, we will enforce the equalities in $\mathcal{L}'$ by inflating the number of sds of particular types. (Keep in mind that because h_1, h_2, and e are logically independent there will necessarily be at least one sd of each of the eight types. Our goal is just to increase the numbers of certain types so as to ensure the equalities above.)

Here we will work out one example of such inflation in detail, leaving it to the reader to generalize the process for other examples. Suppose we have a faithful, adequate $\mathcal{L}$ with $C(\mathcal{L}) = 1$ and $P(\mathcal{L}) = 4$. We will also imagine that $\#(h_1) = \#(h_2) - 2$ (as in the proof overview above). We will construct a faithful, adequate $\mathcal{L}'$ and $h'_1, h'_2, e' \in \mathcal{L}'$ such that:

$$\#(h'_1) = \#(h_1) + 2 \qquad \#(h'_2) = \#(h_2)$$
$$\#(h'_1h'_2) = \#(h_1h_2) \qquad \#(e') = \#(e)$$
$$\#(e'h'_1) = \#(eh_1) \qquad \#(e'h'_2) = \#(eh_2)$$
$$\#(e') = \#(e)$$

Notice that this construction will make $\#(h'_1) = \#(h'_2)$ while leaving the numbers of other sd-types unchanged. If necessary, we may then apply similar constructions to inflate the numbers of other sd-types. For example, we might need to increase the number of eh_1-sds to equal the number of eh_2-sds. Since our construction alters the number of sds of one type while leaving all the other type-counts intact, additional constructions will not undo the work we do in this example construction. (The construction may increase the number of ϕ-sds, but that number is immaterial to satisfying our target equalities.)

We will construct $\mathcal{L}'$ from $\mathcal{L}$ by moving through an intermediary $\mathcal{L}^\dagger$. We construct $\mathcal{L}^\dagger$ from $\mathcal{L}$ directly by applying the construction described in the Availability of Independent Properties four times. Each application adds one new predicate to $\mathcal{L}$ that represents an independent property. The result of four applications is an $\mathcal{L}^\dagger$

that is faithful, has one constant, and has twice as many predicates as $\mathcal{L}$. Moreover, since $\mathcal{L}^\dagger$ has a synonym for each atomic sentence of $\mathcal{L}$, $\mathcal{L}^\dagger$ expresses all of $\mathcal{L}$.

Now take one h_1-sd in $\mathcal{L}$ and one ϕ-sd in $\mathcal{L}$ and designate them s_1 and s_ϕ, respectively. We are going to arrange $\mathcal{L}'$ so that while most sds in $\mathcal{L}$ have synonyms that are sds of $\mathcal{L}'$, s_1 has an $\mathcal{L}'$-synonym that is a disjunction of three $\mathcal{L}'$-sds. This will increase the total number of h_1'-sds to two more than the number of h_1-sds.

$\mathcal{L}'$ will have one constant and 5 predicates. Having one more predicate than $\mathcal{L}$ will ensure that $\mathcal{L}'$ has enough sds to split s_1 into three sds (so to speak).[69] We now construct the interpretation for $\mathcal{L}'$ as follows:

Step 1.1: For each sd of $\mathcal{L}$ that is not s_1 or s_ϕ, pick a distinct sd in $\mathcal{L}'$ and assign it as the synonym of the sd in $\mathcal{L}$.

Step 1.2: Now find the dnf sentence of $\mathcal{L}^\dagger$ that is a synonym of s_1. Call it $s_1^\dagger$. It will be a disjunction of 2^4 sds of $\mathcal{L}^\dagger$, which we will call the "parts" of $s_1^\dagger$. Make one of these parts a synonym of one $\mathcal{L}'$ sd that does not yet have an interpretation—call it s'_{1a}. Take another of the $s_1^\dagger$ parts and assign it as the synonym of another $\mathcal{L}'$ sd that does not yet have an interpretation—call it s'_{1b}. Finally, take the disjunction of all the remaining parts of $s_1^\dagger$ and make it the synonym of another unused sd of $\mathcal{L}'$—call it s'_{1c}.

Step 1.3: We now have a number of remaining $\mathcal{L}'$ sds that lack an interpretation. Take all of them but one and assign them synonyms in $\mathcal{L}^\dagger$ that are "parts" of $s_\phi^\dagger$—that is, that are $\mathcal{L}^\dagger$ sds that are disjuncts of the dnf $\mathcal{L}^\dagger$ synonym for s_ϕ. Finally, take the last remaining $\mathcal{L}'$ sd and assign it an $\mathcal{L}^\dagger$ synonym that is the disjunction of all the remaining unassigned $s_\phi^\dagger$ parts.[70]

Step 1.4: We have now assigned an interpretation to every sd of $\mathcal{L}'$. For any other non-contradictory sentence of $\mathcal{L}'$, assign it the proposition that is the union of the propositions represented by the sds in the sentence's dnf equivalent. Assign any contradictory sentence of $\mathcal{L}'$ the empty proposition.

We have now given an interpretation to every sentence in $\mathcal{L}'$. Notice that every sd of $\mathcal{L}$ has a synonym in $\mathcal{L}'$. The synonyms of most of the $\mathcal{L}$ sds were generated in Step 1.1. Step 1.2 makes $s'_{1a} \vee s'_{1b} \vee s'_{1c}$ the synonym of s_1. Step 1.3 makes the disjunction of a huge number of $\mathcal{L}'$ sds the synonym of s_ϕ. Since every sd of $\mathcal{L}$ has a synonym in $\mathcal{L}'$, $\mathcal{L}'$ expresses all of $\mathcal{L}$. So $\mathcal{L}'$ is adequate for the propositions expressed by h_1, h_2, and e in $\mathcal{L}$. These propositions have synonyms in $\mathcal{L}'$ which we will refer to as h_1', h_2', and e'.

Furthermore, $\mathcal{L}^\dagger$ and $\mathcal{L}'$ are related according to the conditions for $\mathcal{L}^1$ and $\mathcal{L}^2$ in Result 4. An $\mathcal{L}'$ sd that is the synonym of an $\mathcal{L}$ sd other than s_1 or s_ϕ (call the $\mathcal{L}$ sd s_n) represents the same proposition as the disjunction of the set of $\mathcal{L}^\dagger$ "parts" of $s_n^\dagger$. The sets of $\mathcal{L}^\dagger$ sds represented by the other $\mathcal{L}'$ sds are spelled out in the second

---

[69] In general, we will want $\mathcal{L}'$ to have one sd for each sd of $\mathcal{L}$ plus an sd for each number by which we need to "inflate" s_1. (In the case we're currently working through that inflation number is 2.) Since the inflation number will never be greater than the number of sds in $\mathcal{L}$, making $\mathcal{L}'$ have twice as many sds as $\mathcal{L}$ will always do the trick. Since both languages have exactly one constant, this can always be accomplished by making $\mathcal{L}'$ have one more predicate than $\mathcal{L}$.

[70] If you've been keeping track, you'll notice that prior to Step 1.3 we assigned interpretations to $2^4 + 2$ sds of $\mathcal{L}'$. That leaves $2^5 - (2^4 + 2) = 2^4 - 2$ sds of $\mathcal{L}'$ in need of interpretations. The reason we needed $\mathcal{L}^\dagger$ to have twice the number of predicates of $\mathcal{L}$ (that is, 4 more predicates than $\mathcal{L}$) was so that there would be at least $2^4 - 2$ parts of $s_\phi^\dagger$ available.

and third steps above. Notice that each $\mathcal{L}^\dagger$ sd appears in exactly one of these sets. Since $\mathcal{L}^\dagger$ is faithful, Result 4 guarantees that $\mathcal{L}'$ is faithful as well.

So $\mathcal{L}'$ is a faithful language that is adequate for h_1', h_2', and e'. Except for s_1 and s_ϕ, each sd of $\mathcal{L}$ has a synonym that is a single sd of $\mathcal{L}'$. s_1 has a synonym that is a disjunction of three sds of $\mathcal{L}'$, so $\#(h_1') = \#(h_1) + 2$. (Also, $\#(\phi') \geqslant \#(\phi)$.) So we now have $\#(h_1') = \#(h_1)$, as desired. Hopefully it is clear how to repeat this process a second time (if needed) so as to generate a further language in which the number of eh_1-sds equals the number of eh_2-sds, lining up the numbers perfectly for the next step of our proof.

Step 2: Starting with $\mathcal{L}$, we may have had to apply the process described in Step 1 zero, one, or two times to get our sd numbers lined up correctly. However many times we applied that process, let's call the language we finally wound up with $\mathcal{L}'$. Recall that $\mathcal{L}'$ is faithful and adequate for $h_1', h_2', e' \in \mathcal{L}'$ and that $\#(h_1') = \#(h_2')$ and $\#(e'h_1') = \#(e'h_2')$.

Now construct a language $\mathcal{L}^*$. $\mathcal{L}^*$ will have the same single constant as $\mathcal{L}$ and $\mathcal{L}'$, which we will suppose is a. It will have the same number of predicates as $\mathcal{L}'$. Those predicates will be $F^*, G^*, B_1^*, B_2^*, \ldots, B^*_{P(\mathcal{L}')-2}$. We will take a state description of $\mathcal{L}^*$ to be a conjunction that 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 in numerical order by subscript. Notice that $\mathcal{L}^*$ has the same number of sds as $\mathcal{L}'$. We will then 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 $\mathcal{L}^*$ that maps the sds of $\mathcal{L}'$ one-to-one to sd synonyms in $\mathcal{L}^*$, as follows:

Step 2.1: Assign each $e'h_1'h_2'$-, $h_1'h_2'$-, and e'-sd 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 which pattern of B^*-predicates is assigned, as long as no sd of $\mathcal{L}^*$ is assigned more than once.) Note that π will map each of these assigned sds of $\mathcal{L}^*$ to an equivalent of itself.

Step 2.2: Pair off each h_2'-sd with an h_1'-sd. (Our equalities above guarantee that they will pair off without remainder.) Each h_2'-sd 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 which pattern of B^*-predicates is assigned, as long as no two h_2'-sds receive the same pattern.) The h_1'-sd paired with that h_2'-sd 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 h_2'-sd will map it to a synonym of the h_1'-sd, and *vice versa*.

Step 2.3: Pair off each eh_2'-sd with an eh_1'-sd. (The equalities above guarantee that they will pair off without remainder.) Each eh_2'-sd 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 which pattern of B^*-predicates is assigned, as long as that pattern hasn't already been used for an h_2'-sd or another eh_2'-sd.) The eh_1'-sd paired with that eh_2'-sd 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 eh_2'-sd will map it to a synonym of the eh_1'-sd, and *vice versa*.

Step 2.4: For each ϕ'-sd of $\mathcal{L}'$, assign a synonym that is an sd of $\mathcal{L}^*$ that has not been assigned yet. Since $\mathcal{L}'$ and $\mathcal{L}^*$ have the same number of sds, this will exhaust the remaining supply of both $\mathcal{L}'$ sds and $\mathcal{L}^*$ sds.[71]

Step 2.5: We have now assigned an interpretation to every sd of $\mathcal{L}^*$. For any other non-contradictory sentence of $\mathcal{L}^*$, assign it the proposition that is the union of the propositions represented by the sds in the sentence's dnf equivalent. Assign any contradictory sentence of $\mathcal{L}^*$ the empty proposition.

Notice that because the interpretive mapping between $\mathcal{L}'$ sds and $\mathcal{L}^*$ sds is one-to-one, the numbers of sds of each type in $\mathcal{L}^*$ will correspond to the numbers of sds of each type in $\mathcal{L}'$.

Since every sd of $\mathcal{L}'$ has a synonym in $\mathcal{L}^*$, all of $\mathcal{L}'$ is expressed in $\mathcal{L}^*$. So there will be synonyms $h_1^*, h_2^*, e^* \in \mathcal{L}^*$ of h_1', h_2', and e' respectively. These will also be synonyms of h_1, h_2, and e by the transitivity of synonymy. Because $\mathcal{L}^*$ expresses all of $\mathcal{L}'$ and $\mathcal{L}'$ is adequate for h_1', h_2', and e', $\mathcal{L}^*$ will be adequate for h_1^*, h_2^*, and e^*. Moreover, the one-to-one mapping between their sds means that $\mathcal{L}'$ (playing the role of $\mathcal{L}^1$) and $\mathcal{L}^*$ (playing the role of $\mathcal{L}^2$) meet the conditions for Result 4. So since $\mathcal{L}'$ is faithful, $\mathcal{L}^*$ is too.

By language invariance $f(h_1^*, h_2^*, e^*)$. By our construction, π maps $e^*h_1^*h_2^*$-, $h_1^*h_2^*$-, and e^*-sds to equivalents of themselves. π maps distinct h_2^*-sds to equivalents of distinct h_1^*-sds and *vice versa*, while $\#(h_2^*) = \#(h_1^*)$.[72] Also, π maps distinct $e^*h_2^*$-sds to equivalents of distinct $e^*h_1^*$-sds and *vice versa*, while $\#(e^*h_2^*) = \#(e^*h_1^*)$. Putting these facts together, we can see that $\pi(e^*) \dashv\vdash e^*$, $\pi(h_1^*) \dashv\vdash h_2^*$, and $\pi(h_2^*) \dashv\vdash h_1^*$. By Result 8, it is not the case that $f(h_1^*, h_2^*, e^*)$. But now we have a contradiction.

□

We will now prove a corollary of our main result concerning mutual-adjustment equilibrium processes of the type described at the end of Section 3. The idea of such a process is that it takes a list of properties and a proposed favoring relation (and also perhaps a body of evidence) as inputs, mutually adjusts them until they

[71] In order for steps 2.1 through 2.3 to work, the number of B^*-predicates in $\mathcal{L}^*$ (that is, $P(\mathcal{L}')-2$) has to exceed certain minima. In particular, since the number of available B^*-predicate patterns is $2^{P(\mathcal{L}')-2}$, we need the following inequalities to hold:

$$2^{P(\mathcal{L}')-2} \geqslant \#(e'h_1'h_2') + \#(h_1'h_2') + \#(e') \qquad \text{(for Step 2.1)}$$
$$2^{P(\mathcal{L}')-2} \geqslant \#(e'h_2') + \#(h_2') \qquad \text{(for Steps 2.2 and 2.3)}$$

If $P(\mathcal{L}')$ is not great enough for these inequalities to be met, we can repeat the process in Step 1 to increase the number of predicates in $\mathcal{L}'$ while leaving all the sd-type-numbers except $\#(\phi')$ unchanged. To do this, start with $\mathcal{L}'$ playing the role of $\mathcal{L}$, introduce an $\mathcal{L}^\dagger$ with twice as many predicates, then go through Steps 1.1 through 1.4 without designating an s_1 sd in $\mathcal{L}$. That is, don't inflate any of the sds that appear in e, h_1, or h_2. The result will be a language (call it $\mathcal{L}'$) with more predicates than the number you started with, an inflated $\#(\phi')$, but all the other sd-type-numbers unchanged. This new language $\mathcal{L}'$ will meet the inequalities above (if it doesn't, just repeat the process again) and allow you to proceed with Step 2 as described.

[72] Why does π map an h_2^*-sd to an *equivalent* of an h_1^*-sd, and not to the h_1^*-sd itself? Because applying π to an h_2^*-sd will generate a conjunction in which Ga appears before Fa, which technically does not count as an sd of $\mathcal{L}^*$. However, there will be an h_1^*-sd that is equivalent to the result of applying π to the h_2^*-sd; it can be found simply by reversing the order of the Fa and Ga conjuncts. Given the equivalence condition, this small technical detail is of no consequence.

match in some way, and then outputs a three-place favoring relation. This favoring relation then determines evidential favoring among evidence and hypotheses. So let's suppose we have such a process. Given the process, we will introduce a five-place relation $ep(\cdot,\cdot,\cdot,\cdot,\cdot)$. The first three relata of this relation are sentences in a formal language of the type we have been considering; the fourth relatum is a set of predicates from that language; the fifth relatum is a three-place relation whose relata are sentences in the language. $ep(h_1, h_2, e, l, g)$ holds just in case if we start with the property list represented by l and the relation represented by g (and perhaps also the evidence represented by e) and apply our mutual-adjustment process, it will output a relation according to which the proposition represented by h_1 is favored over the proposition represented by h_2 relative to the proposition represented by e.

We now define two conditions that ep may or may not possess. First:

Accuracy: Given any h_1, h_2, e, a faithful language $\mathcal{L}$ adequate for those three, set l of predicates from $\mathcal{L}$, and three-place relation g over sentences in $\mathcal{L}$, $ep(h_1, h_2, e, l, g)$ just in case $f(h_1, h_2, e)$ (where f is the true, objective three-place evidential favoring relation).

We would not want to employ our equilibrium process unless at the end of the process we had a relation that reflected the actual evidential favoring facts as they stand; hence we want ep to be accurate in the sense just defined.[73] Second:

Property Neutrality: Suppose we have:

(1) Any h_1, h_2, and e in adequate, faithful language $\mathcal{L}$;
(2) any two equal-sized sets l and l' of predicates from $\mathcal{L}$;
(3) any one-to-one mapping τ from l to l';
(4) any three-place relations g and g' over sentences in $\mathcal{L}$ meeting the condition that for any $x, y, z \in \mathcal{L}$, $g(x, y, z)$ just in case $g'(\tau(x), \tau(y), \tau(z))$.[74]

Then $ep(h_1, h_2, e, l, g)$ just in case $ep(\tau(h_1), \tau(h_2), \tau(e), l', g')$.

Property neutrality says that our equilibrium process does not have any biases towards certain properties prior to being given a property list, proposed favoring relation, and (perhaps) body of evidence to mutually adjust. If ep is property neutral, then it will give the same result for a set of hypotheses, evidence, property list, and proposed relation as it gives if, for instance, one property is replaced in all of those things by another that doesn't already appear.

Result 10. If f is substantive then ep cannot be both accurate and property neutral.

Proof. Suppose for *reductio* that f is substantive and ep is both accurate and property neutral. Now take logically independent h_1, h_2, and e in faithful, adequate language $\mathcal{L}$ such that $f(h_1, h_2, e)$. Perform the construction of $\mathcal{L}^*$, h_1^*, h_2^*, and e^* described in Result 9. Now let l be a set of predicates in $\mathcal{L}^*$ that includes F^* and all the B^*-predicates; l' be a set that includes G^* and all the B^*-predicates; and let τ map F^* to G^* and each B^*-predicate to itself. Let g be any three-place relation over the sentences of $\mathcal{L}^*$, and construct g' so that for any $x, y, z \in \mathcal{L}^*$, $g(x, y, z)$ just in case $g'(\tau(x), \tau(y), \tau(z))$.

---

[73]Notice that if ep is accurate, both of the possibilities listed with small Roman numerals on page 14 are ruled out.

[74]Here "$\tau(x)$" denotes the sentence generated by taking each predicate symbol appearing in x that belongs to l and replacing it with that predicate's image under τ.

By assumption $f(h_1, h_2, e)$ and by language invariance $f(h_1^*, h_2^*, e^*)$. Since $\mathcal{L}^*$ is faithful and adequate for h_1^*, h_2^*, and e^*, the accuracy of ep gives us $ep(h_1^*, h_2^*, e^*, l, g)$. By the property neutrality of ep, $ep(\tau(h_1^*), \tau(h_2^*), \tau(e^*), l', g')$. By the accuracy of ep, $f(\tau(h_1^*), \tau(h_2^*), \tau(e^*))$. By our construction of $\mathcal{L}^*$ and our selection of τ, we have $\tau(h_1^*) \dashv\vdash h_2^*$, $\tau(h_2^*) \dashv\vdash h_1^*$, and $\tau(e^*) \dashv\vdash e^*$. By the equivalence condition, we then have $f(h_2^*, h_1^*, e^*)$. But that violates the antisymmetry of f, so we have a contradiction. □

Appendix B. Responses to Technical Objections

Objection 1. We should not analyze evidential favoring by representing evidence and hypotheses in faithful languages.

In Appendix A I suggested that the set of formal languages to which we should expect a formal favoring-detection algorithm to apply should be restricted not just to languages that are adequate for the evidence and hypotheses in question but also to languages that are faithful (meaning, among other things, that their atomic sentences represent logically independent propositions). This strikes me as fair because there are important cases of evidential favoring that supervene on entailments. For example, if a body of evidence deductively entails one hypothesis while refuting another, I would say that the evidence favors the former over the latter. The easiest way to reflect this in a favoring-detection algorithm that works on formal languages is to have the algorithm rule that $f(h_1, h_2, e)$ whenever $e \vdash h_1$ and $e \vdash {\sim}h_2$. Yet if the algorithm is applied to languages that lack the entailment property, there may be cases in which $\mu e \vDash \mu h_1$ even though $e \nvdash h_1$, so the algorithm may miss imporant favorings. The general idea that the formal languages to which an algorithm is expected to apply must reflect important semantic features in their syntax is what led me to restrict our results in Appendix A to faithful languages.

Notice that the faithfulness requirement restricts the languages under consideration to those that would be most amenable to formal analysis if evidential favoring were detectable by purely formal property-neutral means. So restricting our attention to faithful languages only makes our job more difficult and in the end makes our general result stronger. (It would be much easier to construct adequate languages meeting various criteria if those languages didn't have to be faithful as well.) Also, Goodman reports that something like a faithfulness requirement was commonly assumed among early formal evidential-favoring theorists.[75]

Carnap, however, repudiated this requirement in the preface to the second edition of his (1950). His stated reason was that some predicates by their very meaning make certain state descriptions impossible. Carnap's example was a language with a two-place predicate W for the warmer-than relation; by the inherent antisymmetry in that relation a state description that affirmed both Wab and Wba would be impossible. (We could generate a similar problem with one-place predicates if, for example, we had predicates for the properties of being red-all-over, blue-all-over, and yellow-all-over in the same language.) Carnap proposed adopting Kemeny's solution of accompanying each (unfaithful) language with a set of "meaning postulates" or "A-postulates" that would specify the impossible state descriptions.

[75]Goodman (1946, n. 2) reports that "although this requirement is not explicitly stated in the articles cited" (articles by Hempel, Oppenheim, Helmer, and Carnap), according to Hempel a requirement that primitive predicates represent logically independent properties "was recognized by all the authors concerned."

How can we acommodate Carnap's maneuver within our formal framework? The key is to note that whenever we have an unfaithful language adequate for a given set of hypotheses and evidence, we can always construct a faithful language that is adequate as well. Suppose we are working with adequate, unfaithful language $\mathcal{L}$. As Carnap points out, the effect of the A-postulates is to pick out certain sds of $\mathcal{L}$ as expressing impossible propositions. All we need to do, then, is construct a language $\mathcal{L}'$ whose sds are synonyms for the possible propositions expressed by sds of $\mathcal{L}$. $\mathcal{L}'$ will then be faithful, and it will express all of $\mathcal{L}$. (Each possible sd of $\mathcal{L}$ will have an $\mathcal{L}'$ synonym that's an sd, while the impossible $\mathcal{L}$ sds will have a contradiction in $\mathcal{L}'$ as their synonym.) So $\mathcal{L}'$ will be adequate as well.[76] We can then follow the steps in the proof of our general result to generate a faithful, adequate language in which the hypotheses and evidence are expressed by sentences that are predicate permutations.

In short, I believe that faithfulness is a good requirement on languages to which a favoring-detection algorithm is to be applied, but for those who reject this requirement we can loosen the stipulations of our general result so that it applies to unfaithful languages as well.

Objection 2. Evidential favoring is contrastive and applies only to hypotheses that are mutually exclusive.[77] Substantivity requires there to be favoring between hypotheses that are logically independent. So this objection would urge that substantivity is a bad requirement on an evidential favoring relation.

I happen to think that there can be evidential favoring even between hypotheses that are not mutually exclusive. (See, for instance, the Cadillac and number-drawing examples in Section 1 of this article.) But let's see how the logical independence of h_1 and h_2 factors into our general result. We use the logical independence of e, h_1, and h_2 to guarantee that sds of all eight types are available in $\mathcal{L}$. But the proof doesn't actually require the availability of *all* sd types. For instance, if there were no h_1h_2-sds and no eh_1h_2-sds the proof would proceed just fine (because the synonyms of those sds in $\mathcal{L}^*$ just get mapped by π to equivalents of themselves). Allowing h_1 and h_2 to be mutually exclusive would cause only sds of these two types to disappear. So once more we can loosen the stipulations of our general result: the logical independence requirement in the substantivity condition can be relaxed to allow mutually exclusive hypotheses without undermining the outcome.

(To be maximally precise about the weakening of the general result being considered here: Our proof will show that the evidential favoring relation does not treat predicate permutations identically as long as there exists one set of h_1, h_2, e in an adequate, faithful language $\mathcal{L}$ such that $f(h_1, h_2, e)$ and the following conditions are met:

[76] The languages with which we are working always have a number of sds that is a power of 2. What if the number of sds that represent possible propositions in $\mathcal{L}$ is not a power of 2? How can we construct $\mathcal{L}'$?

The answer is to apply something like the maneuver described in Step 1 of our proof of the general result so as to generate a language $\mathcal{L}^\dagger$ in which some of the propositions represented as single sds in $\mathcal{L}$ are represented as disjunctions of multiple sds. Applied carefully, this procedure will ensure that the number of sds in $\mathcal{L}^\dagger$ representing possible propositions is a power of 2, thereby allowing us to construct $\mathcal{L}'$ by making each of its sds a synonym for an sd of $\mathcal{L}^\dagger$ that represents a possible proposition.

[77] Chandler (ms), for instance, takes this position.

(i) $\mu({\sim}h_1 \,\&\, {\sim}h_2 \,\&\, {\sim}e)$ is nonempty.
(ii) If $\mu(h_1 \,\&\, {\sim}h_2 \,\&\, e)$ is nonempty then $\mu({\sim}h_1 \,\&\, h_2 \,\&\, e)$ is as well, and *vice versa*.
(iii) If $\mu(h_1 \,\&\, {\sim}h_2 \,\&\, {\sim}e)$ is nonempty then $\mu({\sim}h_1 \,\&\, h_2 \,\&\, {\sim}e)$ is, and *vice versa*.

Put in more recognizable terms, these become:

(i) ${\sim}h_1 \,\&\, {\sim}h_2 \nvdash e$. That is, h_1 and h_2 are not logically exhaustive, and it is not the case that every rival hypothesis to h_1 and h_2 entails e.
(ii) $e \,\&\, h_1 \vdash h_2$ just in case $e \,\&\, h_2 \vdash h_1$. That is, we are not in a situation in which one of the hypotheses is strictly stronger than the other relative to a background of e.
(iii) ${\sim}e \,\&\, h_1 \vdash h_2$ just in case ${\sim}e \,\&\, h_2 \vdash h_1$. That is, we are not in a situation in which one of the hypotheses is strictly stronger than the other relative to a background of ${\sim}e$.)

Objection 3. Evidence and hypotheses should not be thought of as propositions. The objection is that instead of thinking of evidence and hypotheses as propositions (understood as sets of possible worlds), we should treat them as propositions under modes of presentation, or structured propositions, etc. For instance, even if Hesperus is Phosphorous a body of evidence about Hesperus might not be properly treated as identical with a body of evidence that says the same things about Phosphorous.

The simplest response to this objection is to adjust the μ function so that it maps sentences to objects other than propositions with a distinct entailment relation. For instance, we might create a space of *epistemically* possible worlds in which some worlds assign Hesperus different properties than Phosphorous. The μ function would then interpret sentences in a formal language as sets of epistemically possible worlds. The problem revealed by our general result isn't really about an agent's inability to identify the same objects (or same propositions, if you like) across modes of presentation. In the end that information is just entailment information, and can do no better than the agent's total evidence in helping him figure out which properties are treated preferentially by a substantive favoring relation.

Objection 4. Our formal framework does not allow for favoring relations among logical truths. Since all logical truths express the same proposition, any logical truths expressed by a faithful language will be syntactically equivalent, and so by the equivalence condition cannot be differentially favored by any evidence. Yet we might think that some mathematical truths, for instance, are better evidence for certain mathematical theorems than for others.

One solution here is to work with epistemically possible worlds again, but now allow epistemically possible worlds that are logically impossible. A more conservative response would be to say that if there is a substantive evidential favoring relation that ranges over both empirical and logical truths, it will have a subpart that concerns favoring only among propositions that are not logically necessary. That subpart is the relation being modeled by our $f(\cdot,\cdot,\cdot)$, and is the relation to which our results apply. Since defenders of an objective three-place evidential favoring relation presumably want it to yield substantive results even when the hypotheses and evidence are empirical propositions, I trust this will still be very bad news for them.

Objection 5. Evidential favoring cannot be adequately represented in first-order languages.

Goodman's grue discussion fails to be perfectly general in two ways. First, it applies only to the formal theories of evidential favoring that were available in his day. I take it our general result has addressed that lack of generality. Second, however, Goodman limits his discussion to examples of evidence and hypotheses that are adequately representable in first-order languages, and our general result is limited to such evidence and hypotheses as well. We are examining efforts to model a subtle, complex evidential favoring relation, yet the formal languages in which we represent evidence and hypotheses contain no quantifiers, relations, higher-order predicates, modal relations, etc.

The first response to this objection is that some additional formal apparati can be incorporated into our languages as they stand. Result 7 describes how the propositions represented in a language with multiple constants can be re-represented in a language with a single constant representing a tuple of the objects represented by the constants in the first language. Once we have made this transition, multi-place relations among the objects represented in the first language can be represented by one-place predicates in the second language (applied to the constant representing the tuple). So multi-place relations can be represented within our simple languages as well.

However, there remain many types of logical structures that can't be represented within our languages. Our second response is to pin down exactly where this matters. Our simple formal languages are used to represent bodies of evidence and the hypotheses among which they establish favoring relations. Whatever algorithms, reasoning processes, etc. are used to determine when those favoring relations hold, such processes do *not* need to be representable within our formal languages. So, for instance, we might have evidence that contains no modal expressions favoring hypotheses with no modal expressions, yet the process by which we detect that favoring could involve using the evidence to derive modal claims and then determining the implications of those modal claims for our (nonmodal) hypotheses. Our general result treats the process by which agents work out evidential favoring relations as a "black box"—regardless of *how* agents do it, the result shows that if their process determines a substantive favoring relation it must treat some properties differently from others.

Moreover, even if hypotheses involving modals, quantifiers, etc. can be favored by bodies of evidence, anyone who believes in an objective, three-place evidential favoring relation will also believe that evidence can sometimes establish favoring relations among hypotheses expressible in our simple first-order languages. So we can think of our systems as modeling the favoring relations established among such simple hypotheses by various bodies of evidence, regardless of what favoring relations among other hypotheses may also be established by that evidence. And then our conclusions concerning substantivity and preferential property treatment will apply over this limited span of the favoring relation's range.

In other words, the simplicity of the languages in our general result is not a limitation as far as representing complex favoring-detection *processes* goes and is not a crippling limitation as far as representing potentially-favored *hypotheses* goes. The complaint must focus on the limited ability of our languages to represent bodies of *evidence*.

Our third response is that there are probably important evidential situations (situations to which the defender of an objective, three-place evidential relation would think that relation applies) in which the evidence does not involve the sorts of logical structures that our languages lack the ability to represent. For instance, we can meaningfully inquire which hypotheses are favored by our perceptual evidence, and I think it's dubious to suggest that our direct perceptual evidence contains quantified propositions. (There's also an important argument to be had in the philosophy of perception as to whether it includes any modals.) If perceptual evidence does not include quantified propositions (even if quantified propositions such as existentials can be immediately *inferred* from what that evidence contains), then the lack of quantifiers in our formal languages will not present an obstacle to representing the favoring of hypotheses by perceptual evidence, and the conclusions we draw from our general result will apply to that sphere.

One might reply that there are still important features unmistakably present in evidence that our simple languages are unable to represent. For example, perceptual evidence almost certainly includes second-order predications (as when we observe that certain color properties are similar to each other). If this is to create an objection to the conclusions we derive from our general result, the objection must be that while those conclusions (about the inconsistency of substantivity and the identical treatment of predicate permutations, for instance) are true for evidence and hypotheses in simple first-order languages, they do not apply to evidence and hypotheses that can be adequately expressed only with more complex logical structures. If the argument of this article's main text is correct, this would be to concede that there is no objective favoring among simple evidence and hypotheses and to maintain that evidential favoring comes into play only once the evidence achieves a certain level of logical complexity.

To argue that evidential favoring depends crucially on evidence's having a certain degree of logical complexity is to abandon a significant strand of the history of formal inquiry into such favoring. Carnap, Hempel, and their interlocutors began their investigations with favoring among simple sentences and then tried to work their way up to more complex situations. These authors (and Goodman as well) would have been surprised indeed to learn that there just are no substantive favoring relations among simple predications. But there are other important strands in the analysis of evidential support on which the relevant features only kick in when, say, equations involving adjustable parameters become involved. And those who believe in the deep subtlety of evidential favoring—and the impossibility of modeling it by formal means—may take the failure of the bottom-up approach as more grist for their mill. Someone who believes that objective evidential relations may be detected purely *a priori*, for instance, can maintain that there is a pure, property-neutral *a priori* method for detecting such substantive relations and brush off our negative results by claiming that these relations only emerge once the evidence has more logical complexity than the formal languages entertained here can represent.

Still, anyone who took this approach would have to concede that for beings who were incapable of gaining evidence with the requisite logical complexity, there simply would be no substantive evidential favoring (that is, favoring beyond direct entailment and refutation) among the evidence and hypotheses those beings entertained. It seems to me a very strong position to claim that only once we achieved the ability to gain evidence with, say, modal structure were we able to make evidentially-based inductive predictions about our world.

Despite these responses, I think this objection is the best objection to our general result. The obvious next step for the technical approach pursued here is to see if something like it can be applied to higher-order languages. For the time being I will reiterate that a main goal of this article is to emphasize the importance of language-dependence results for broad issues in philosophy. If the defenders of popular positions in epistemology and normativity theory concede that the viability of those positions hangs on whether certain technical results can be extended to higher-order languages, I will consider that goal to have been achieved.