

Article

Symmetry and Evidential Support

Michael G. Titelbaum

Department of Philosophy, University of Wisconsin–Madison, 600 North Park Street, Madison, WI 53706, USA; E-Mail: titelbaum@gmail.com; Tel. +1-608-263-6708; Fax: +1-608-265-3701

Received: 27 April 2011; in revised form: 25 August 2011 / Accepted: 30 August 2011 /

Published: 16 September 2011

Abstract: This article proves that formal theories of evidential favoring must fail because they are inevitably language dependent. I begin by describing Carnap’s early confirmation theories to show how language dependence problems (like Goodman’s grue problem) arise. I then generalize to show that any formal favoring theory satisfying minimal plausible conditions will yield different judgments about the same evidence and hypothesis when they are expressed in alternate languages. This does not just indict formal theories of favoring; it also shows that something beyond our evidence must be invoked to substantively favor one hypothesis over another.

Keywords: evidence; evidential support; Carnap; grue; language dependence

1. Introduction

We are constantly judging hypotheses in light of evidence, to see whether the evidence supports one hypothesis over another, is better explained by one hypothesis, makes one hypothesis more likely than the other, *etc.* In doing so we invoke a general *evidential favoring relation*, according to which an evidential proposition may favor one hypothesis proposition over another. When that relation holds among such relata, it reveals an important asymmetry between the hypotheses relative to the evidence—an asymmetry we rely on in making predictions, designing experiments, and deciding among courses of action. Little wonder, then, that philosophers since Hempel ([1] and [2]) have tried to formalize the relation of evidential favoring. I begin this article by explaining and evaluating Carnap’s [3] probabilistic theory of evidential favoring. Among other problems, that theory has trouble with language dependence: it yields different favoring judgments when the same hypotheses and evidence are re-expressed in a different language. In another article [4], I have proven that these language dependence problems generalize

beyond the particulars of Carnap's theory, and even beyond probabilistic approaches to evidential favoring. Here I will explain how the proof proceeds, what assumptions it makes about favoring, and how it shows that an accurate formal theory of evidential favoring is impossible. Finally, I will sketch some philosophical consequences of this result, in particular that hypotheses are more symmetrically related to evidence than we think and that choosing among hypotheses almost always requires a contribution from agents that goes beyond their evidence.

2. Carnap's Early Theories of Favoring

Carnap's central idea is that a body of evidence favors one hypothesis over another just in case it renders the former more probable than the latter. To make this precise, he first represents evidence and hypothesis propositions in a first-order formal language $\mathcal{L}$. Such a language has *predicates* ($F, G, H, \dots$) that can be applied to *constants* ($a, b, c, \dots$) to form *atomic sentences* ($Fa, Gb, Hc, \dots$). By combining atomic sentences with sentential connectives, we build more complex sentences. Carnap's languages are *interpreted*: each constant represents a particular object in the world; each predicate represents a property such objects may display; each sentence represents a particular proposition.

With his interpreted language in place, Carnap defines a real-valued function c on ordered pairs of sentences in $\mathcal{L}$. $c(h, e)$ measures how strongly the evidence represented by e confirms the hypothesis represented by h . The proposition represented by e favors the proposition represented by h_1 over the proposition represented by h_2 just in case $c(h_1, e) > c(h_2, e)$.¹ The substance of Carnap's theory comes from the particular values of c . c is defined in terms of a real-valued one-place function m on $\mathcal{L}$. For any $x, y \in \mathcal{L}$ (with y non-contradictory),

$$c(x, y) = \frac{m(x \& y)}{m(y)} \quad (1)$$

Carnap requires that m be a *probability function* in the sense of Kolmogorov [5]: non-negative, normal, and finitely additive. He also requires that m be *regular*, in the sense that it assigns positive values to all non-contradictions. Combined with Equation (1), these requirements make $c(\cdot, e)$ a probability function for any non-contradictory $e \in \mathcal{L}$. This squares with Carnap's central idea that $c(h, e)$ —the degree to which e confirms h —is just the probability of h on e .

Carnap then needs to specify values for m . He proves that a probability distribution over $\mathcal{L}$ can be fully specified by assigning values to $\mathcal{L}$'s state descriptions. A *state description* is a maximal consistent conjunction of literals of $\mathcal{L}$. (A *literal* is either an atomic sentence or its negation.) Any non-contradictory sentence $x \in \mathcal{L}$ is logically equivalent to a disjunction of state descriptions; that disjunction is called x 's *disjunctive normal form*.

For example, take a simple language $\mathcal{L}^G$ with two constants a and b and one predicate G . Table 1 lays out a set of state descriptions for $\mathcal{L}^G$. (The table's leftmost column provides a name for each state description.)² Suppose we assign m -values to each of the four state descriptions in Table 1. Carnap

¹This is Carnap's "firmness" explication of confirmation, as opposed to the "increase in firmness" explication he distinguishes from it in the preface to the second edition of [3]. We will stick with the firmness explication because it is easier to work with, but my criticisms below apply equally well to the increase in firmness explication.

²To prevent logical redundancy in a language's set of state descriptions, we assume that no literal appears more than once in a state description and the literals appear in alphabetical order. So $Ga \& Gb$ is a state description of $\mathcal{L}^G$, but $Ga \& Ga \& Gb$ and $Gb \& Ga$ are not.

proves that the state descriptions' m-values generate a probabilistic distribution over $\mathcal{L}^G$ just in case they are non-negative and sum to 1. If that condition is met, the m-value of any non-contradictory $x \in \mathcal{L}^G$ can be calculated by summing the m-values of the state descriptions in its disjunctive normal form. (A contradictory x receives an m-value of 0.)

Table 1. $m^\dagger$ - and m^* -values for language $\mathcal{L}^G$.

Name	State description	$m^\dagger$	m^*
s_1	$Ga \ \& \ Gb$	1/4	1/3
s_2	$Ga \ \& \ \sim Gb$	1/4	1/6
s_3	$\sim Ga \ \& \ Gb$	1/4	1/6
s_4	$\sim Ga \ \& \ \sim Gb$	1/4	1/3

Take the $m^\dagger$ column of Table 1, for instance. Setting aside for the moment where $m^\dagger$ comes from, the $m^\dagger$ -values specified for the state descriptions of $\mathcal{L}^G$ are clearly non-negative and sum to 1. This induces a probability function $m^\dagger$ over the entire $\mathcal{L}^G$ language. For example, since the disjunctive normal form of sentence Ga is a disjunction of s_1 and s_2 , $m^\dagger(Ga)$ is the sum of the $m^\dagger$ -values in Table 1's first two rows. In other words, $m^\dagger(Ga) = 1/2$.

Carnap has now moved from the problem of specifying c-values to the problem of specifying m-values, and from the problem of specifying an entire m-distribution to the problem of specifying m over state descriptions. He notes that for any x , $m(x) = c(x, T)$, where T is an arbitrary tautology in $\mathcal{L}^G$. So he thinks of $m(x)$ as the probability of the proposition represented by x relative to a tautology, or relative to an evidence set containing no empirical information.³ The state descriptions of $\mathcal{L}^G$ describe the basic states of the world $\mathcal{L}^G$ is capable of discriminating among, and it is natural to think that lacking any empirical information, each of these states should be treated symmetrically. That is, we should assign each state description the same m-value. This gives us distribution $m^\dagger$ in Table 1, which (applying Equation (1)) yields a confirmation function Carnap calls $c^\dagger$.

$c^\dagger$ captures some intuitive favoring relations. For example, a bit of calculation with Table 1 and Equation (1) reveals that

$$1 = c^\dagger(Ga, Ga \ \& \ Gb) > c^\dagger(\sim Ga, Ga \ \& \ Gb) = 0 \quad (2)$$

Recall that evidence favors one hypothesis over another just in case the former receives a higher c-value on the evidence than the latter. So $c^\dagger$ is suggesting a sensible favoring relation: Since the evidence represented by $Ga \ \& \ Gb$ entails the hypothesis represented by Ga while refuting the hypothesis represented by $\sim Ga$, that evidence favors the former over the latter.

But broadening our view reveals clear flaws in $c^\dagger$ as an explication of evidential favoring. Suppose, for example, that a and b represent two emeralds to be sampled, and G represents the property of being

³On a probabilistic approach like Carnap's, an evidence set is represented by a conjunction each of whose conjuncts represents one of the set's members. It makes no difference to the probability calculations if an extra, tautologous conjunct is added to this evidential conjunction. So we can imagine that if we removed empirical facts from an evidence set one at a time, in the end our "evidential conjunction" would be just a tautology.

green. We might think that the evidence represented by Ga favors the hypothesis represented by Gb over the hypothesis represented by $\sim Gb$. Yet from Table 1 we can calculate

$$c^\dagger(Gb, Ga) = c^\dagger(\sim Gb, Ga) = 1/2 \quad (3)$$

The problem generalizes. When we construct $c^\dagger$ for larger languages, we find that even when many objects are represented the evidence that all but one have the property G does not confirm the hypothesis that the last object has G over the hypothesis that it does not. As Carnap puts it, $c^\dagger$ does not allow learning from experience.

The trouble is that $m^\dagger$ is too symmetrical. To get learning from experience, we need a state in which all the emeralds are green to be more probable than a state in which the first run of emeralds are green but the last one is not. By treating all state descriptions symmetrically, $m^\dagger$ renders our evidence incapable of discriminating among full states of the world consistent with that evidence.

So Carnap constructs a new confirmation function c^* from the probability function m^* . Instead of treating state descriptions symmetrically, m^* treats *structure descriptions* symmetrically. Intuitively, a structure description describes the distribution of properties over objects at an abstract level, without specifying which objects occupy which places in the property distribution. A structure description might say “one sampled emerald is green while another is not,” without telling us which of the emeralds is the green one.

At the technical level, a structure description is a disjunction of state descriptions obtainable from each other by permuting constants. Our simple language $\mathcal{L}^G$ has three structure descriptions: s_1 by itself (“both emeralds are green”), the disjunction of s_2 and s_3 (“exactly one of the emeralds is green”), and s_4 by itself (“neither emerald is green”). m^* assigns each structure description the same value; within a given structure description, m^* -value is divided up equally among the state descriptions. The resulting m^* -values are displayed in Table 1.⁴

c^* solves the learning from experience problem, because

$$2/3 = c^*(Gb, Ga) > c^*(\sim Gb, Ga) = 1/3 \quad (4)$$

This evidential favoring is possible because s_1 has a higher m^* -value than s_2 . There’s only one way to arrange the world such that both emeralds are green, so there’s only one state description in s_1 ’s structure description and that state description gets the full $1/3$ probability. Yet there are two possible arrangements in which exactly one emerald is green (first emerald green, second emerald green), so s_2 has to share a structure description with s_3 and gets an m^* -value of only $1/6$. Symmetric treatment of structure descriptions yields asymmetric treatment of state descriptions, making the sought-after favoring relations possible.⁵

But c^* has a new problem: language dependence. Even with a language as simple as $\mathcal{L}^G$, we can construct a version of Goodman’s [6] “grue” problem. Suppose we define a language $\mathcal{L}^H$ with the same constants as $\mathcal{L}^G$ representing the same objects. But $\mathcal{L}^H$ has a predicate H related to G as follows: for any

⁴Marlos Viana points out to me (in correspondence) that the transition from distributions over state descriptions to distributions over structure descriptions is familiar from Bose–Einstein, Fermi–Dirac, and Boltzmann–Maxwell statistics.

⁵ c^* can also retrieve the favoring judgment represented in Equation (2)—in fact, any c -function derived from a regular, probabilistic m -function will retrieve that judgment.

constant x , Hx has the same truth-value as $Gx \equiv (x = a)$.⁶ Table 2 displays a set of state descriptions of $\mathcal{L}^H$ and their m^* -values, calculated just as before. But it also reveals a further fact: each state description of $\mathcal{L}^H$ expresses the same proposition as a state description of $\mathcal{L}^G$. That means $\mathcal{L}^H$ can express every proposition expressible in $\mathcal{L}^G$: given any $\mathcal{L}^G$ -sentence, we find its disjunctive normal form in $\mathcal{L}^G$, replace its $\mathcal{L}^G$ state descriptions with the corresponding $\mathcal{L}^H$ state descriptions from Table 2, and are left with an $\mathcal{L}^H$ synonym for the $\mathcal{L}^G$ original. Each $\mathcal{L}^G$ sentence has a *synonym* in $\mathcal{L}^H$: an $\mathcal{L}^H$ sentence that expresses the same proposition as the $\mathcal{L}^G$ original.

Table 2. m^* -values for language $\mathcal{L}^H$.

State description	m^*	Expresses same proposition as
$Ha \ \& \ Hb$	$1/3$	s_2
$Ha \ \& \ \sim Hb$	$1/6$	s_1
$\sim Ha \ \& \ Hb$	$1/6$	s_4
$\sim Ha \ \& \ \sim Hb$	$1/3$	s_3

Consulting Table 2, we find that

$$1/3 = c^*(\sim Hb, Ha) < c^*(Hb, Ha) = 2/3 \quad (5)$$

But Ha is a synonym for Ga and Hb is a synonym for $\sim Gb$. Equation (5) indicates exactly the opposite favoring relation—relative to the same evidence—as Equation (4)!

Goodman reverses the confirmation relations among propositions indicated by a theory of evidential favoring by re-expressing the same propositions in a new language. But we can also make those favoring relations disappear entirely. Consider a new language $\mathcal{L}^S$. $\mathcal{L}^S$ has one constant, o , which names the ordered pair consisting of first the emerald named by a and then the emerald named by b in $\mathcal{L}^G$. $\mathcal{L}^S$ has two predicates. G_1x obtains when the first element of the ordered pair named by x is green. Sx obtains when the two objects in the ordered pair are the same with respect to greenness—they are either both green or they are both not. Since there is only one constant in this new language, permuting constants does not turn any state description into another; each state description has its own structure description. Table 3 displays the resulting m^* -values for $\mathcal{L}^S$.

⁶This is a metalinguistic statement about how the truth-values of propositions represented by atomic $\mathcal{L}^H$ sentences relate to the truth-values of propositions represented by atomic $\mathcal{L}^G$ sentences. None of our object languages contain both G and H as predicates, nor do any of our languages represent the identity relation.

Table 3. m^* -values for language $\mathcal{L}^S$.

State description	m^*	Expresses same proposition as
$G_1o \ \& \ So$	1/4	s_1
$G_1o \ \& \ \sim So$	1/4	s_2
$\sim G_1o \ \& \ So$	1/4	s_4
$\sim G_1o \ \& \ \sim So$	1/4	s_3

Again, each state description of $\mathcal{L}^G$ has a synonym in $\mathcal{L}^S$, so every proposition expressible in $\mathcal{L}^G$ is expressible in $\mathcal{L}^S$. And now we have

$$c^*(G_1o \equiv So, G_1o) = c^*(\sim[G_1o \equiv So], G_1o) = 1/2 \quad (6)$$

G_1o expresses the same proposition as Ga did in language $\mathcal{L}^G$, while $G_1o \equiv So$ is a synonym for Gb . In $\mathcal{L}^S$, the favoring relations we found earlier disappear. $\mathcal{L}^H$ and $\mathcal{L}^S$ reveal that the facts c^* relies upon—facts about which propositions are expressed in sentences that share structure descriptions with others—are artifacts of language choice.

3. Alternative Approaches to Favoring

Evidential favoring is a relation among propositions—evidence and hypotheses. With only a few exceptions (which we'll discuss later), favoring relations among propositions should turn out the same regardless of which language the propositions are expressed in. Yet Carnap's formal analyses of confirmation yield different favoring judgments for the same propositions when those propositions are expressed in different languages.

Once the possibility of a language like $\mathcal{L}^S$ —a language in which all the objects are referred to via a single n -tuple name—has come up, it may seem like any formal Carnap-style theory of favoring is doomed. But some features of evidence and hypotheses are invariant among all the languages we have considered, even $\mathcal{L}^S$. Consider again the situation in which our evidence is that emerald a is green, our first hypothesis is that emerald b is green, and our second hypothesis is that emerald b is not. In each of the three languages we have seen, that evidence and those hypotheses are each expressed by a disjunction of two state descriptions. For example, in $\mathcal{L}^G$ their disjunctive normal forms are:⁷

$$\begin{array}{ll} h_1 : s_1 \vee s_3 & h_2 : s_2 \vee s_4 \\ e : s_1 \vee s_2 \end{array}$$

Notice also that each hypothesis shares one state description with the evidence but no state descriptions with the other hypothesis. That remains true across $\mathcal{L}^G$, $\mathcal{L}^H$, and $\mathcal{L}^S$.

We could imagine a formal confirmation theory that works with these facts about state-description counts and state-description sharing. For instance, we can invent a “Proportional Theory” that counts what

⁷Since s_1 through s_4 are names we have given to state descriptions of $\mathcal{L}^G$ —and not sentences appearing in $\mathcal{L}^G$ —the disjunctive normal forms written here for h_1 , h_2 , and e are metalinguistic indications of what those disjunctive normal forms actually look like in $\mathcal{L}^G$.

proportion of its state descriptions a hypothesis shares with the evidence, then favors a hypothesis that has a higher proportion of shared state descriptions over one that has a lower proportion. In the example at hand each hypothesis would have a shared proportion of $1/2$, so neither h_1 nor h_2 would be favored by e over the other. So the Proportional Theory fails to yield intuitively plausible favoring judgments about learning from experience—but it can still serve as our toy example of a theory that gives consistent favoring results across the transition from $\mathcal{L}^G$ to $\mathcal{L}^H$ and $\mathcal{L}^S$. It appears at least possible that with a bit of work something in the Proportional Theory's neighborhood could yield plausible favoring results.

But this appearance is illusory. In Sections 5 and 6 I will describe a general proof that rules out all formal theories of evidential favoring. Starting with very general conditions on what any formal favoring theory would have to achieve, I will show that even if a theory yields consistent indications of favoring for a set of evidence and hypotheses across $\mathcal{L}^G$ through $\mathcal{L}^S$, those indications disappear when the relata are expressed in yet another language. After describing the proof, I explain in Section 7 what we can learn from it about the underlying nature of evidential favoring.

Possessing such a general proof is also important for historical reasons. Despite the positive results c^* yields for simple cases of enumerative induction, Carnap was dissatisfied with that confirmation function's inability to properly model what he called "arguments by analogy". So Carnap produced and refined a number of successors to c^* . Meanwhile, Jaynes ([7] and [8]) developed a rival confirmation approach based on maximizing entropy in probabilistic distributions. Other authors, such as Maher [9] have since tried to develop further formal theories of evidential favoring.⁸

While presenting Carnap's first steps is useful for illustrative purposes, there's no need to work through these further proposals because they all run into the same problem Carnap did: language dependence.⁹ The result I present in the next few sections reveals that this is not a coincidence—any formal theory of favoring meeting very general conditions will have language dependence problems.

4. General Conditions on Evidential Favoring

What, in general, do we know about the evidential favoring relation? It is a relation among propositions, but we must express those propositions as sentences in a language to work with them. So if $h_1, h_2, e \in \mathcal{L}$ represent two hypotheses and a body of evidence in interpreted first-order language $\mathcal{L}$ (with a finite number of constants and predicates, and no quantifiers or identity symbol), we will write $f(h_1, h_2, e)$ when the evidence represented by e favors the hypothesis represented by h_1 over the hypothesis represented by h_2 . Unlike Carnap, we will not assume that f has anything to do with probabilities or numerical functions, though what we do assume about f will be compatible with that possibility. For example, we will assume that f is antisymmetric relative to a given e —that is, we cannot have both $f(h_1, h_2, e)$ and $f(h_2, h_1, e)$.

We might also think that f should obtain in some clear cases involving entailment relations. We saw one such case in Equation (2): the evidential proposition represented in $\mathcal{L}^G$ by $Ga \ \& \ Gb$ favors the hypothesis represented by Ga over the hypothesis represented by $\sim Ga$. But if we want a theory of f to detect these entailment-based favorings, we cannot expect the theory to yield correct results for

⁸See Maher's [10] for discussion of arguments by analogy and Carnap's later views, as well as favoring theories by Hesse and others. Maher freely admits the language-dependence of his own systems.

⁹For example, [11] develops language dependence problems for Jaynes's maximum entropy approach.

every conceivable language, because some languages will hide the relevant entailments. For example, a language with a single constant (like the ordered-pair constant in $\mathcal{L}^S$) might express the evidence and pair of hypotheses from Equation (2) with the atomic sentences To , Uo , and Vo . In that language there would be no way to formally recover the entailment facts in virtue of which the favoring holds.

This example shows that we cannot demand invariance of a confirmation theory's results over every language capable of expressing the evidence and hypotheses of interest. As we go along, we will consider which kinds of language-independence we want and which are inessential. We will judge this by asking what it would reveal about the underlying evidential favoring relation if a correct theory of that relation were invariant across a particular language type. Since it seems plausible that some evidential favorings arise from entailments, we will require a formal theory to detect evidential favoring only when propositions are represented in *faithful* languages. While faithfulness is defined precisely in [4], the key condition is that the state descriptions of a faithful language express a set of mutually exclusive, exhaustive propositions. Because of this, two sentences x and y in a faithful language will have $x \vdash y$ just in case the proposition represented by x entails the proposition represented by y .¹⁰ A faithful language captures in its syntax all the entailment relations among the propositions it represents.

Unfaithful languages go wrong by failing to capture entailment relations among the propositions they represent. But a language may also go wrong by failing to represent some propositions entirely. When we work with scientific hypotheses, a body of evidence may favor one hypothesis over another because it reveals the truth of a prediction made by the former but not the latter.¹¹ A language that faithfully represents the entailment relations among evidence and hypotheses but lacks a sentence representing that prediction may leave formal theories incapable of detecting the favoring relation that obtains.

We will call a language *adequate* for a particular set of evidence and hypotheses if it contains sentences representing not only those three propositions but also any other propositions necessary for detecting favoring relations among the three. I have no precise characterization of the conditions a language must meet to be adequate for a particular set of evidence and hypotheses. Luckily, all we need for our result is that the concern for adequacy is a concern about representational paucity. We will suppose that for any evidence and two hypotheses, there is a set of languages adequate for those three relata. We will require formal theories of evidential favoring to yield correct results when applied to adequate languages. And we will assume that if language $\mathcal{L}$ is adequate for a set of evidence and hypotheses, and language $\mathcal{L}'$ contains synonyms for every sentence in $\mathcal{L}$, then $\mathcal{L}'$ is adequate for those relata as well.

To this point we have required that the f relation be anti-symmetric for a given e and that formal favoring theories detect the relation's presence when its relata are represented in a language that is adequate for them and faithful. Although I believe various favorings based on entailments (such as the one represented in Equation (2)) hold, we will be able to prove our result without assuming there are any such favorings in the extension of f . We will, however, assume that the extension of f contains something besides entailment-based favorings. This follows Hume's point in [13, Book I] that if favoring

¹⁰The full definition of faithfulness also ensures that if we think of propositions as sets of worlds, a faithful language will have $\sim x$ represent the complement of the proposition represented by x , $x \& y$ represent the intersection of the propositions represented by x and y , and $x \vee y$ represent the union of those propositions.

¹¹This idea is familiar from the hypothetico-deductivist theory of confirmation; see [12] for discussion.

underwrites the inductive inferences that get us through our days, it cannot be restricted to cases in which the evidence either entails or refutes one of the hypotheses.

To be precise: We will say that h_1 , h_2 , and e in faithful language $\mathcal{L}$ are *logically independent* just in case any conjunction obtainable by inserting a non-negative number of negation symbols into “ h_1 & h_2 & e ” represents a non-empty (non-contradictory) proposition. If there exists at least one faithful $\mathcal{L}$ with logically independent $h_1, h_2, e \in \mathcal{L}$ such that $f(h_1, h_2, e)$, we will say that f is *substantive*. I take it Hume showed us evidential favoring is substantive.¹²

Finally, we need a condition capturing what it is to say that a theory of the evidential favoring relation is “formal”. Typically, formality means that a theory operates on the structure of sentences without noticing which particular items play which roles in that structure. For example, suppose a theory of evidential favoring said that in $\mathcal{L}^G$, Ga favored Gb over $\sim Gb$, but Gb did not favor Ga over $\sim Ga$. These two triples are structurally identical; such a theory would be differentiating between them strictly on the grounds that a appeared in the evidence in the first case while b appeared in the evidence in the second. A formal theory treats constants in a language as interchangeable. It also treats predicates as interchangeable, which is the condition that will play a central role in our proof. We will require f to *treat predicate permutations identically*, by which we mean that for any language $\mathcal{L}$ that is faithful and adequate for $h_1, h_2, e \in \mathcal{L}$ and any permutation π of the predicates of $\mathcal{L}$, $f(h_1, h_2, e)$ entails $f(\pi(h_1), \pi(h_2), \pi(e))$. (Where $\pi(x)$ is the sentence that results from replacing each predicate occurrence in x with its image under the permutation π .)

5. First Stage of the Proof

With these notions in place, our general result can be stated simply:

General Result: If the evidential favoring relation is antisymmetric and substantive, it does not treat predicate permutations identically.

For the full details of the proof, I refer the reader to [4]. Here I will simply explain how it works, starting with an overview of the proof strategy: Suppose for *reductio* that evidential favoring is substantive and antisymmetric and treats predicate permutations identically. By f ’s substantivity there exist an h_1 , h_2 , and e in faithful, adequate language $\mathcal{L}$ such that these three relata are logically independent and $f(h_1, h_2, e)$. We will construct another faithful, adequate language $\mathcal{L}^*$ with h_1^* , h_2^* , and e^* representing the same propositions as h_1 , h_2 , and e respectively. Since f concerns a relation among the propositions expressed by sentences, we will have $f(h_1^*, h_2^*, e^*)$. Moreover, $\mathcal{L}^*$ will be constructed so as to make available a predicate permutation π such that $\pi(h_1^*) = h_2^*$, $\pi(h_2^*) = h_1^*$, and $\pi(e^*) = e^*$. Since f treats predicate permutations identically, $f(h_2^*, h_1^*, e^*)$. But that violates f ’s antisymmetry, yielding a contradiction.

The difficult part of this proof is demonstrating that the required $\mathcal{L}^*$, h_1^* , h_2^* , e^* , and π can be constructed in the general case. That generalization proceeds in two stages. The first stage begins by noting that while $\mathcal{L}$ might contain multi-place predicates and any (finite) number of constants, given any

¹²Evidential favoring will not be substantive if evidence can only discriminate among hypotheses by logically ruling some out, as suggested by the falsificationism of [14]. Miller [15] is moved to falsificationism in part by language dependence issues related to those discussed in this article. Nevertheless, falsificationism remains a minority view of evidential favoring.

faithful adequate $\mathcal{L}$ we can always find another language that is faithful and adequate for the propositions expressed by h_1 , h_2 , and e but contains a single constant and only single-place predicates. We do this in much the same way we moved from language $\mathcal{L}^G$ to language $\mathcal{L}^S$ earlier: We make the new language's constant represent a tuple of the objects represented by the constants of $\mathcal{L}$, then make the new language's predicates represent properties of objects at particular places in the tuple ("the first object in the tuple is green") or relations among those objects ("the objects in the tuple match with respect to greenness"). Since whenever there is a faithful, adequate language representing h_1 , h_2 , and e there is also a faithful, adequate language representing the same propositions with only one constant and single-place predicates, we will assume without loss of generality that our original language $\mathcal{L}$ has only one constant, a , and only single-place predicates.

Having made this assumption, let's fix a case in our minds by imagining that besides its one constant $\mathcal{L}$ has five predicates—it does not matter which particular predicates they are. We can then refer to the 32 state descriptions of $\mathcal{L}$ as s_1 through s_{32} . And to further fix a particular case, we can imagine that the h_1 , h_2 , and e in question have these disjunctive normal form equivalents:¹³

$$\begin{aligned} h_1 &: s_1 \vee s_4 \vee s_5 \vee s_7 & h_2 &: s_2 \vee s_4 \vee s_6 \vee s_7 \\ e &: s_3 \vee s_5 \vee s_6 \vee s_7 \end{aligned}$$

The reader may establish that this particular h_1 , h_2 , and e are logically independent as required.

We will now construct language $\mathcal{L}^*$ with one constant, a , representing the same tuple as it represents in $\mathcal{L}$. Like $\mathcal{L}$, $\mathcal{L}^*$ will have five predicates—they will be F^* , G^* , B_1^* , B_2^* , and B_3^* . Like $\mathcal{L}$, $\mathcal{L}^*$ will be an interpreted language; the key to our proof will be how we assign meanings to the sentences of $\mathcal{L}^*$. $\mathcal{L}^*$ will be faithful, and will be capable of expressing all the propositions expressed by sentences of $\mathcal{L}$. So, for instance, $\mathcal{L}^*$ will contain an h_1^* , h_2^* , and e^* that are synonyms for h_1 , h_2 , and e respectively. Because $\mathcal{L}^*$ expresses every proposition expressible in $\mathcal{L}$, and because we have assumed $\mathcal{L}$ is adequate for h_1 , h_2 , and e , $\mathcal{L}^*$ will be adequate for h_1^* , h_2^* , and e^* . But $\mathcal{L}^*$ is designed so that a predicate permutation π that maps each B^* -predicate to itself and interchanges F^* with G^* will map h_1^* to h_2^* , h_2^* to h_1^* , and e^* to itself.

To construct $\mathcal{L}^*$, we begin with the disjunctive normal form equivalents of h_1 , h_2 , and e in $\mathcal{L}$. These sentences have some state descriptions in common and some distinct. For instance, s_4 appears in the disjunctive normal forms of h_1 and h_2 but not of e . So we will refer to s_4 as an " h_1h_2 -sd". (A state description that does not appear in any of our three disjunctive normal forms will be a " ϕ -sd".) The basic strategy for setting up $\mathcal{L}^*$ will be to make each state description of $\mathcal{L}^*$ the synonym of a different state description of $\mathcal{L}$. (Once meanings are assigned to the state descriptions of $\mathcal{L}^*$, the meanings of other $\mathcal{L}^*$ sentences will be built up by interpreting logical connectives in the usual way.)¹⁴ This means that $s_4 \in \mathcal{L}$, for instance, will have a synonym state description in $\mathcal{L}^*$ occurring in the disjunctive normal forms of h_1^* and h_2^* but not e^* . We will label that state description in $\mathcal{L}^*$ an " $h_1^*h_2^*$ -sd".

We will achieve the permutations we want by carefully selecting which propositions are expressed by which state descriptions in $\mathcal{L}^*$. For example, a proposition expressed by an h_1 -sd in $\mathcal{L}$ will be expressed by an h_1^* -sd in $\mathcal{L}^*$. But we get to select which $\mathcal{L}^*$ state description plays that role, so we will select one

¹³Again, these are metalinguistic indications of what the disjunctive normal forms look like, not actual object-language expressions of the evidence and hypotheses.

¹⁴Because $\mathcal{L}$ is faithful, its state descriptions express a set of mutually exclusive, exhaustive propositions. Since the state descriptions of $\mathcal{L}^*$ express these same propositions, $\mathcal{L}^*$ will wind up faithful as well.

that is mapped by π to an h_2^* -sd. More generally, we will assign state descriptions of $\mathcal{L}^*$ to propositions so that π maps state description types to each other as described in Table 4. (The rows of Table 4 have been numbered for reference later.)

Table 4. In $\mathcal{L}^*$, π maps...

(i)	each h_1^* -sd to	an h_2^* -sd and <i>vice versa</i> ,
(ii)	each e^* -sd to itself,	
(iii)	each $h_1^*h_2^*$ -sd to itself,	
(iv)	each $e^*h_1^*$ -sd to	a $e^*h_2^*$ -sd and <i>vice versa</i> ,
(v)	each $e^*h_1^*h_2^*$ -sd to itself, and	
(vi)	each ϕ^* -sd to	a ϕ^* -sd.

Notice how these mappings have been set up. Once we know which state descriptions express which propositions, we can determine the disjunctive normal form equivalent of h_1^* . When π is applied to that disjunctive normal form equivalent, it will replace each h_1^* -sd with an h_2^* -sd and each $e^*h_1^*$ -sd with an $e^*h_2^*$ -sd. In other words, it will replace each state description that appears in h_1^* but not h_2^* with a state description that appears in h_2^* but not h_1^* . As a result, applying π converts h_1^* into h_2^* . Similarly, if the mappings in Table 4 hold, π will convert h_2^* to h_1^* while leaving e^* unchanged.¹⁵

How do we assign state descriptions to propositions so as to achieve the mappings described in Table 4? For the specific example we have been working with, the state description assignments in Table 5 will do the job. Here I have indicated state descriptions schematically, leaving out the “ a ”s and replacing negations with overbars. The table indicates, for example, that the proposition expressed by s_4 in $\mathcal{L}$ is expressed in $\mathcal{L}^*$ by the state description $F^*a \& G^*a \& B_1^*a \& \sim B_2^*a \& B_3^*a$. Notice that in our example s_4 is an h_1h_2 -sd, and our permutation π maps the assigned $\mathcal{L}^*$ state description to itself. More generally, the positions in Table 5 match the positions in Table 4: s_1 is an h_1 -sd, s_2 is an h_2 -sd, s_3 is an e -sd, *etc.* (Table 5 does not assign state descriptions for ϕ -sds because those assignments turn out not to matter, as long as each ϕ -sd receives an $\mathcal{L}^*$ equivalent that has not been assigned to any other state description of $\mathcal{L}$.) Assignments in the rows of Table 5 ensure that the mappings in the same-numbered rows of Table 4 hold. In row (i), for example, when π swaps F^* and G^* the h_1^* -sd synonym of s_1 is exchanged with the h_2^* -sd synonym of s_2 .¹⁶

The assignments in Table 5 employ a system that can be extended to the general case. First, using guidance from Table 4 we have divided the $\mathcal{L}$ state descriptions appearing in h_1 , h_2 , and/or e into pairs and singletons, depending on whether π will map the relevant state description to another state

¹⁵Technically, the order of some conjuncts in the disjunctive normal form of e may be changed. But here and elsewhere in the argument that follows, I treat a sentence of a faithful language as interchangeable with other sentences in that language logically equivalent to it. If we think of propositions as sets of possible worlds, then any two logically equivalent sentences in

Table 5. $\mathcal{L}^*$ synonyms for some $\mathcal{L}$ state descriptions.

(i)	$s_1 : F^* \overline{G^*} B_1^* B_2^* B_3^*$	$s_2 : \overline{F^*} G^* B_1^* B_2^* B_3^*$
(ii)	$s_3 : F^* G^* B_1^* B_2^* \overline{B_3^*}$	
(iii)	$s_4 : F^* G^* B_1^* \overline{B_2^*} B_3^*$	
(iv)	$s_5 : F^* \overline{G^*} B_1^* \overline{B_2^*} B_3^*$	$s_6 : \overline{F^*} G^* B_1^* \overline{B_2^*} B_3^*$
(v)	$s_7 : F^* G^* \overline{B_1^*} B_2^* B_3^*$	

description or to itself. The pairs and singletons have been aligned on the individual rows. Each row is assigned a unique binary code using the B^* -predicates. Row (i) is assigned the binary code $B_1^* B_2^* B_3^*$; row (ii) gets $B_1^* B_2^* \overline{B_3^*}$, etc. We then determine the F^* s and G^* s. If one state description is to be mapped onto another (and *vice versa*) by π , the state description belonging to h_1^* affirms F^* and denies G^* , while the description it'll be mapped onto denies F^* and affirms G^* . The singletons, meanwhile, affirm both F^* and G^* so that π will map them to themselves.¹⁷

With all this work done, π maps h_1^* to h_2^* and *vice versa* while leaving e^* intact. And this was our goal: We started with $f(h_1, h_2, e)$, so we have $f(h_1^*, h_2^*, e^*)$. We also have $\pi(h_1^*) = h_2^*$, $\pi(h_2^*) = h_1^*$, and $\pi(e^*) = e^*$ in faithful, adequate language $\mathcal{L}^*$. So by the identical treatment of predicate permutations, $f(h_2^*, h_1^*, e^*)$. Together with f 's antisymmetry this yields our contradiction.

6. Second Stage of the Proof

Hopefully I have described the strategy for this single case so that the reader can generalize to similar cases. Unfortunately, though, there is a class of cases to which the strategy (as so far described) will not generalize. Suppose our h_1 , h_2 , and e in faithful, adequate $\mathcal{L}$ (with five predicates) have the following disjunctive normal form equivalents:

$$\begin{aligned} h_1 &: s_1 \vee s_4 \vee s_5 \vee s_7 & h_2 &: s_0 \vee s_2 \vee s_4 \vee s_6 \vee s_7 \\ e &: s_3 \vee s_5 \vee s_6 \vee s_7 \end{aligned}$$

These relata are exactly as before, except that I have added an extra disjunct s_0 to h_2 .

a faithful language express the same proposition. Since f relates the propositions expressed by sentences, any two equivalent sentences in a faithful language will enter into f -relations in exactly the same ways.

¹⁶One might wonder what the predicates of $\mathcal{L}^*$ mean—for instance, what property of the a -tuple is represented by the predicate G^* ? We can construct the meaning of G^* from the meanings of the (presumably well-understood) predicates of $\mathcal{L}$. Each state description of $\mathcal{L}$ says something about a , and there are tuples in the world that make that state description true when referred to by a . Each state description of $\mathcal{L}^*$ is a synonym of an $\mathcal{L}$ state description; an $\mathcal{L}^*$ state description says the same thing as its $\mathcal{L}$ counterpart. G^*a is equivalent to a disjunction of state descriptions of $\mathcal{L}^*$, so we can determine which tuples make G^*a true when referred to by a . G^* expresses the property of belonging to the set containing just those tuples.

¹⁷Depending on the size of $\mathcal{L}$ and the particular constitution of h_1 , h_2 , and e , this assignment strategy may require $\mathcal{L}^*$ to have more predicates, and thus more state descriptions, than $\mathcal{L}$. The relevant calculations and a general recipe for making this work are described in [4, Appendix A].

If we pursue the strategy of the previous section, each of $\mathcal{L}$'s state descriptions will have a state-description synonym in $\mathcal{L}^*$. But even if we are clever and set things up so that the $\mathcal{L}^*$ state description synonym for s_2 is mapped by π to the $\mathcal{L}^*$ state description synonym for s_1 , there will be nothing left in h_1^* for π to map the synonym of s_0 onto. So π will not map h_2^* onto h_1^* as desired.

The problem is that in our new case the numbers of each type of state description do not line up in the right way. Looking at Table 4, we can see that in order for our mapping scheme in $\mathcal{L}^*$ to work, we need the following equalities met:

$$\begin{aligned}\text{number of } h_1^* \text{-sds} &= \text{number of } h_2^* \text{-sds} \\ \text{number of } e^* h_1^* \text{-sds} &= \text{number of } e^* h_2^* \text{-sds}\end{aligned}$$

(Notice from Table 4 that the number of e^* -, $h_1^* h_2^*$ -, $e^* h_1^* h_2^*$ -, and ϕ^* -sds is irrelevant, as they are mapped by π to themselves.) As our strategy stands, the state descriptions of $\mathcal{L}^*$ match up one-to-one with state descriptions of $\mathcal{L}$, so unstarred versions of those equalities must be met as well. But the first (unstarred) equality is not met by the h_1 and h_2 under consideration; in this case there are more h_2 -sds than h_1 .

And this case is important, because it represents a favoring instance under the Proportional Theory discussed in Section 3. Recall that under the Proportional Theory, one hypothesis is favored over another if a greater proportion of its state descriptions is shared with the evidence. In the case at hand, h_1 shares half the state descriptions in its disjunctive normal form with e , while h_2 shares only $2/5$. So according to the Proportional Theory $f(h_1, h_2, e)$. Moreover, every case in which the Proportional Theory indicates that one hypothesis is favored over another by evidence will violate the unstarred version of at least one of the equalities above.¹⁸ So for our proof to rule out the Proportional Theory as a viable theory of evidential favoring, it must apply to cases in which the unstarred version of at least one of these equalities fails.

In the present case there is one more h_2 -sd than there are h_1 -sds, but for our permutation mapping in $\mathcal{L}^*$ to work out we need the number of h_2^* -sds to match the number of h_1^* -sds. To make this happen, we must break the one-to-one mapping between propositions expressed by state descriptions of $\mathcal{L}^*$ and propositions expressed by state descriptions of $\mathcal{L}$. (Breaking the one-to-one mapping allows the starred equalities to hold when their unstarred analogs do not.) In particular, if the proposition expressed by s_1 —a member of h_1 's disjunctive normal form but not h_2 's—was expressed by the disjunction of two state descriptions in $\mathcal{L}^*$, the first of these state descriptions could be mapped by π to the synonym of s_2 while the other could be mapped to the synonym of s_0 . “Splitting” s_1 —an h_1 -sd—into two state descriptions in $\mathcal{L}^*$ would increase the number of h_1^* -sds by one while leaving the number of h_2^* -sds unchanged. This would bring the number of h_1^* -sds and h_2^* -sds into the required alignment.

The trick is to take the proposition expressed by s_1 in $\mathcal{L}$ and express it as the disjunction of two state descriptions in $\mathcal{L}^*$, while leaving $\mathcal{L}^*$ adequate and faithful and matching all the other $\mathcal{L}$ state descriptions that participate in h_1 , h_2 , or e to state descriptions of $\mathcal{L}^*$ one-to-one (so as not to alter the count of any non- ϕ -sd-types besides the h_1 -sds). We will achieve this using a technique I call “explode and gather”. This technique involves introducing two intermediary languages, $\mathcal{L}^\dagger$ and $\mathcal{L}'$, through which we

¹⁸Quick proof: If we let $\#h_1$ represent the number of h_1 -sds and so forth, the proportion of its state descriptions that h_1 shares with e is $(\#h_1 e + \#h_1 h_2 e)/(\#h_1 + \#h_1 h_2 + \#h_1 e + \#h_1 h_2 e)$ and the proportion of its state descriptions that h_2 shares with e is $(\#h_2 e + \#h_1 h_2 e)/(\#h_2 + \#h_1 h_2 + \#h_2 e + \#h_1 h_2 e)$. If the unstarred versions of the two equalities are met, these proportions are equal. So if the unstarred equalities are both met, the Proportional Theory indicates no favoring. Contraposing, if the Proportional Theory indicates a favoring relation at least one of the unstarred equalities is violated.

move from $\mathcal{L}$ to $\mathcal{L}^*$. Instead of assigning the state descriptions of $\mathcal{L}$ synonyms directly in $\mathcal{L}^*$, we will assign them synonyms in $\mathcal{L}^\dagger$. Sentences in $\mathcal{L}^\dagger$ will in turn receive synonyms in $\mathcal{L}'$, which will finally receive synonyms in $\mathcal{L}^*$. The net result will be the representation of every proposition expressed by a state description of $\mathcal{L}$ in $\mathcal{L}^*$, but the number of h_1^* -sds in $\mathcal{L}^*$ will be one greater than the number of h_1 -sds in $\mathcal{L}$.

To get a rough idea of how “explode and gather” works, imagine we had a language $\mathcal{B}$ whose one constant p referred to a pig and whose one predicate N indicated that the pig was north of the barn. Our h_1 might be Np and our h_2 might be $\sim Np$. We would then introduce a second language $\mathcal{B}^\dagger$ that contained the constant p , the predicate $N^\dagger$, and the predicate $W^\dagger$ indicating that the pig was west of the barn. This is the “explosion” phase; while each of our hypotheses was represented by a single state description in $\mathcal{B}$, each is now represented by a disjunction of two state descriptions in $\mathcal{B}^\dagger$. (Np , for instance, has become $(N^\dagger p \ \& \ W^\dagger p) \vee (N^\dagger p \ \& \ \sim W^\dagger p)$.) Notice that just like the state descriptions of $\mathcal{B}$, the state descriptions of $\mathcal{B}^\dagger$ express a set of mutually exclusive, exhaustive propositions; $\mathcal{B}^\dagger$ is faithful just like $\mathcal{B}$.

Now we “gather” with a new language $\mathcal{B}'$. $\mathcal{B}'$ has three state descriptions. One is a synonym for $\sim N^\dagger p$, one is a synonym for $N^\dagger p \ \& \ W^\dagger p$, and the last is a synonym for $N^\dagger p \ \& \ \sim W^\dagger p$. $\mathcal{B}'$ cannot express every proposition expressible in $\mathcal{B}^\dagger$, but it can express every proposition expressible in $\mathcal{B}$, so it inherits its adequacy from that language. Also, the state descriptions of $\mathcal{B}'$ express a set of mutually exclusive, exhaustive propositions, so $\mathcal{B}'$ will be faithful as well. And in $\mathcal{B}'$, the synonym of h_2 is a single state description while the synonym of h_1 is a disjunction of two. There is one more h'_1 -sd than there are h_1 -sds, but the number of h'_2 -sds equals the number of h_2 -sds.

This example, while hopefully suggestive, cannot actually be carried through in the details—for one thing, a standard first-order language cannot contain exactly three state descriptions. But now that we have got the broad idea down, we can implement the technical details of “explode and gather” for the $\mathcal{L}$ -example we have been considering.

We begin by constructing $\mathcal{L}^\dagger$. $\mathcal{L}^\dagger$ has the same single constant, a , as $\mathcal{L}$, representing the same tuple as in $\mathcal{L}$. $\mathcal{L}^\dagger$ contains every predicate in $\mathcal{L}$, representing the same properties as they represent in $\mathcal{L}$. But $\mathcal{L}^\dagger$ has two additional predicates, which (for the sake of definiteness) we will say are $D^\dagger$ and $E^\dagger$. $D^\dagger$ and $E^\dagger$ represent properties not already represented by any of the predicates of $\mathcal{L}$. (Their introduction is analogous to the introduction of $W^\dagger$ in the pig example.)¹⁹ Notice that under this construction each state description of $\mathcal{L}$ will have a synonym in $\mathcal{L}^\dagger$. But the $\mathcal{L}^\dagger$ synonym of an $\mathcal{L}$ state description s_i will not be a state description of $\mathcal{L}^\dagger$. Instead, it will be a disjunction of four state descriptions: One that looks just like s_i but also affirms $D^\dagger a \ \& \ E^\dagger a$, another that looks just like s_i but appends $D^\dagger a \ \& \ \sim E^\dagger a$, etc. Thus we have “exploded” each state description of $\mathcal{L}$ into multiple state descriptions of $\mathcal{L}^\dagger$.

Now we “gather” the state descriptions of $\mathcal{L}^\dagger$ into the language $\mathcal{L}'$. $\mathcal{L}'$ has the same constant as $\mathcal{L}$ and $\mathcal{L}^\dagger$ representing the same tuple. But $\mathcal{L}'$ has only one more predicate than $\mathcal{L}$. Instead of saying what the predicates of $\mathcal{L}'$ represent, I will describe how we assign $\mathcal{L}'$ state descriptions as synonyms of sentences

¹⁹In fact, $D^\dagger$ and $E^\dagger$ must be chosen so that it is logically possible for the tuple represented by a to have any combination of the properties represented by $D^\dagger$, $E^\dagger$, and the predicates of $\mathcal{L}$. This is required so that the state descriptions of $\mathcal{L}^\dagger$ will express a set of mutually exclusive, exhaustive propositions and $\mathcal{L}^\dagger$ can be faithful. The possibility of finding properties of the tuple expressible by $D^\dagger$ and $E^\dagger$ that are independent in the relevant sense of each other and of the properties represented in $\mathcal{L}$ is guaranteed by an assumption I call the Availability of Independent Properties. For a defense of that assumption see [4, p. 502].

in $\mathcal{L}^\dagger$. First, assign distinct $\mathcal{L}'$ state descriptions as synonyms for the $\mathcal{L}^\dagger$ sentences expressing state descriptions of $\mathcal{L}$ —except for the $\mathcal{L}^\dagger$ sentences expressing s_1 and the ϕ -sds of $\mathcal{L}$. (Recall that s_1 is the state description we’re trying to “split” in two.) s_1 , like all the state descriptions of $\mathcal{L}$, has been exploded into four “parts” in $\mathcal{L}^\dagger$. Take one of those parts and give it a synonym that is a state description in $\mathcal{L}'$. Then take the other three parts and give their disjunction another state description synonym in $\mathcal{L}'$.²⁰

The resulting $\mathcal{L}'$ will not be able to express every proposition expressible in $\mathcal{L}^\dagger$. But it will be able to express every proposition expressible in $\mathcal{L}$, because each $\mathcal{L}$ state description has a synonym in $\mathcal{L}'$. And since $\mathcal{L}$ is adequate for h_1 , h_2 , and e , $\mathcal{L}'$ will be adequate for the propositions expressed by those relata as well.²¹ Moreover, each state description of $\mathcal{L}$ appearing in h_1 , h_2 , or e has a state description synonym in $\mathcal{L}'$ —except for s_1 . The synonym of s_1 is a disjunction of two state descriptions of $\mathcal{L}'$. Since they compose the synonym of s_1 and s_1 is a disjunct of h_1 , both of these state descriptions will be disjuncts of h'_1 . So while there was only one h_1 -sd in $\mathcal{L}$, there will be two h'_1 -sds in $\mathcal{L}'$. Thus the disjunct-type counts in $\mathcal{L}'$ will satisfy our earlier equalities (with primes taking the place of stars). That means we can construct $\mathcal{L}^*$ from $\mathcal{L}'$ with a one-to-one mapping between state descriptions, cleverly assigning meanings to state descriptions of $\mathcal{L}^*$ as described in Section 5. $\mathcal{L}^*$ will be adequate and faithful, and the permutation π that exchanges F^* with G^* while leaving B^* -predicates intact will effect exactly the mappings described in Table 4. So our proof will go through as before.

Let’s take a step back and assess. The Proportional Theory and theories like it work by counting the number of disjuncts in hypotheses’ disjunctive normal forms and the number of disjuncts those hypotheses share with the evidence. Yet these count facts are artifacts of the language in which we choose to express our evidence and hypotheses. In this section I have worked through a particular example in which “explode and gather” takes us from one faithful, adequate language to another faithful, adequate language in which the disjunct counts of sentences expressing the same propositions have changed. No matter what logically independent evidence and hypotheses we are presented with, the explode and gather technique can be applied (multiple times if necessary) to bring disjunct counts into the balance we need to apply the proof maneuver from Section 5. Accurate theories of evidential favoring yield the same favoring judgments across all faithful, adequate languages, and if they are formal they do so while treating predicate permutations identically. But given logically independent relata we will always be able to construct an $\mathcal{L}^*$ in which the hypotheses are predicate permutation variants, so that a formal theory will detect no favoring between them. This means that a formal theory will never find that a body of evidence differentially favors logically independent hypotheses. So if it can be captured by a formal theory, evidential favoring is not substantive.

²⁰What about the sentences in $\mathcal{L}^\dagger$ that are synonyms of ϕ -sds? For our particular h_1 , h_2 , and e , $\mathcal{L}$ has 32 state descriptions 24 of which are ϕ -sds. Each ϕ -sd has a synonym that is a disjunction of 4 $\mathcal{L}^\dagger$ state descriptions. $\mathcal{L}'$ has one more predicate than $\mathcal{L}$, so it has 64 state descriptions. We have already assigned synonyms to 9 of them, leaving 55. Of the 24 ϕ -sds in $\mathcal{L}$, use $\mathcal{L}^\dagger$ to split 17 of them into disjunctions of 2 state descriptions in $\mathcal{L}'$, and split the other 7 into disjunctions of 3 state descriptions in $\mathcal{L}'$. This will assign meanings to the remaining 55 state descriptions of $\mathcal{L}'$ and ensure that each ϕ -sd has a synonym in $\mathcal{L}'$. (For a generalization of the math involved here, including how big $\mathcal{L}^\dagger$ and $\mathcal{L}'$ will typically have to be, see [4, Appendix A], especially notes 68 through 70.)

²¹A bit of thought will also reveal that because the state descriptions of $\mathcal{L}^\dagger$ express a mutually exclusive, exhaustive set of propositions the state descriptions of $\mathcal{L}'$ do so as well. This makes $\mathcal{L}'$ faithful.

One final note about the proof: One might wonder why h_1 , h_2 , and e have to be logically independent for our argument to go through. The “explode and gather” technique increases the number of h_1 -sds (say) by “splitting” an existing h_1 -sd into two h'_1 -sds. This works only if there is an h_1 -sd to begin with. If h_1 , h_2 , and e were arranged so that there were some h_2 -sds but no h_1 -sds, we would not be able to split h_1 -sds to make their number equal the h_2 -sds'. Logical independence of h_1 , h_2 , and e guarantees that there is at least one state description in $\mathcal{L}$ of each type, avoiding situations that would derail the proof. But once we have seen that proof, it is clear that the logical independence requirement could be relaxed. For example, we could run the argument for an h_1 , h_2 , and e with no h_1 -sds as long as there were no h_2 -sds either. More generally, the relata we work with have to make some ϕ -sds available, and if they make one sd-type on a row of Table 4 available they have to make any other types on that row available as well. This means that we could, for instance, accommodate positions on which evidential favoring only ever occurs between mutually exclusive hypotheses.²² In a faithful language two mutually exclusive hypotheses h_1 and h_2 have no h_1h_2 -sds and no eh_1h_2 -sds. But eliminating those rows from Table 4 would not disrupt our mapping scheme. So while I will continue to assume logical independence among the relata for simplicity's sake, it is significant that this requirement could be relaxed.

7. The Favoring Relation Itself

Even if the logical independence requirement were relaxed, our general result would still have shortcomings. For example, it applies only to evidence and hypotheses adequately expressible in a first-order language with no quantifiers or identity symbol. In [4, Appendix B] I argue that this is not as serious a limitation as it might seem. Certainly most of the evidential favoring on which we rely in daily life, scientific research, *etc.* involves more complicated logical structures than this. But notice that all we need to get our proof going is one instance of logically independent evidence and hypotheses that are first-order expressible. Moreover, these syntactical limitations apply only to the evidence and hypotheses, not to the processes used by theories that detect evidential favoring. Even if an entropy calculation or something more mathematically complex is used to determine the extension of f , our result will apply as long as at least one triple in that extension consists of sentences as simple as “This emerald is green” and “The next one will be too.”

Our result may also seem aimed at a project most philosophers have abandoned; very few people still maintain that evidential favoring can be ferreted out by formal means. So let's stop thinking about formal theories of favoring, and start thinking about the evidential favoring relation itself. Most of the conditions we imposed in Section 4 concerned that relation directly; the only one motivated by the prospect of formal theorizing was identical treatment of predicate permutations. But the identical treatment of predicate permutations is significant even when we set formal theories aside. If f fails to treat predicate permutations identically, there exist hypotheses and evidence in adequate, faithful languages whose favoring relations disappear when their predicates are swapped around. Since predicates represent properties, a favoring relation that treats predicate permutations non-identically behaves differently towards propositions containing some properties than it does towards otherwise-identical propositions containing different properties. A favoring relation that does not treat predicate permutations identically

²²For one such position, see [16].

plays favorites among properties. Our general result says that if evidential favoring is antisymmetric and substantive, it must play favorites.

The standard lesson drawn from language dependence results such as Goodman's "grue" problem is that evidential favoring privileges some properties over others. Goodman himself [6, Lecture IV] tried to distinguish "projectible" properties from nonprojectible; later philosophers such as Quine [17] attempted to identify "natural properties" that play a special role in evidential favoring. If there are such special properties, favoring theories need not yield consistent results across faithful, adequate languages; such theories (formal or not) may restrict their attention to languages whose predicates represent the privileged properties. If greenness, for example, is a natural property, then Carnap's c^* function should be applied only to language $\mathcal{L}^G$ —not $\mathcal{L}^H$ or $\mathcal{L}^S$ —and the intuitive confirmation result of Equation (4) stands.

This standard response leaves an epistemic problem: How can agents determine which properties are natural? The obvious response is that natural properties are revealed by evidence from the natural world. But here the generality of our result kicks in. Imagine some multi-stage process agents could apply: first, the process would use an agent's empirical evidence to determine a list of natural properties. Next, the process would sort languages whose predicates represented natural properties from those whose predicates did not. Finally, the process would work within the set of preferred languages to determine which hypotheses were favored over others by the agent's evidence. Now consider a relation np that captures the net effect of this entire process: $np(h_1, h_2, e)$ holds just in case the natural property list generated by the evidence represented by e yields a favoring relation on which e favors the hypothesis represented by h_1 over the hypothesis represented by h_2 .

Now np is just a relation, so the process that takes agents from the inputted e to an ultimate favoring judgment between h_1 and h_2 need not proceed in the sequence just described. The key point is that however it is generated, np should be antisymmetric and had better be substantive, so that favoring relations may obtain among hypotheses neither entailed nor refuted by the evidence. And so (substituting np for f) our general result applies: np will fail to treat predicate permutations identically. That is, antecedent to the introduction of a particular body of evidence e , np will already prefer some properties over others. Anyone who tries to work out the natural properties from a body of empirical evidence will need a preferred property list before that evidence is even consulted.

Apparently substantive favoring among hypotheses arises from something more than just evidence; it requires an extra-evidential element to select among properties. As far as the evidence and hypotheses themselves are concerned—considered alone as propositions, without any further information or influences—any two logically independent hypotheses are related to a given body of evidence in the same way. The appearance of asymmetry among these propositions, the sense that the evidence has more in common with one hypothesis or favors it over the other, is an artifact of the language in which the propositions are expressed. And the propositions themselves cannot tell you which language is best.²³

²³Notice how much stronger this result is than standard "underdetermination of theory by evidence" arguments. Underdetermination of theory by evidence typically argues that while some e may favor h_1 over h_2 , we can always manufacture an h_3 that is just as well supported by e as h_1 . The present result shows that even if e , h_1 , and h_2 are selected for us, there will be no favoring of h_1 over h_2 by e unless some logical entailment holds among them. It is not just that an outlandish theory can be made up that accounts for the data; it is that no theory under serious consideration accounts for the data better than any of the others.

This leaves a few options in the theory of evidential favoring. First, one can be an externalist: One can insist that there are facts in the world about which properties are natural, or projectible—it is just that those facts cannot be discerned by agents from their evidence. While there are favoring relations among hypotheses and evidence, agents cannot ever know that they have got those relations right. [4] presents arguments against this option; I will simply say here that I find favoring facts inaccessible by agents highly unattractive. Second, one can hold that the preferred property list need not be discerned from empirical evidence because it can be determined *a priori*. Besides requiring a very strong view of our *a priori* faculties, this response contradicts the positive theories philosophers have offered of what makes projectible properties projectible. The standard theory (from [18], among others) is that projectible properties recur regularly because they play a particular role in the natural laws of the universe, but such laws must be gleaned from empirical data.²⁴

The remaining option maintains that the element privileging some properties over others comes from agents, and so is accessible to them and allows them to determine when favoring holds. This “subjectivist” option denies that there is a three-place evidential favoring relation (among two hypotheses and a body of evidence) at all; it relativizes favoring to a fourth relatum that is a feature of subjects. Perhaps agents grow up speaking a language whose predicates express certain properties; perhaps agents evolved to think using certain categories; perhaps for some reason agents have a prior disposition to project some properties more readily than others. Wherever their preferred property lists come from, subjects with different lists may not be able to adjudicate disputes between those lists (and the favoring relations that attend them) by citing evidence or *a priori* considerations. While this approach avoids the problems of the other options—it permits substantive evidential favoring, posits no unrealistic *a priori* faculties, and allows agents access to favoring facts—it may require us to radically rethink what we are doing when we ask which of two hypotheses is favored by our evidence.

Acknowledgements

I am grateful to Hayley Clatterbuck, Casey Helgeson, Marlos Viana, Paul Bartha, and Brandon Rdzak for discussion of earlier drafts of this article.

²⁴In correspondence Paul Bartha asks me to consider whether the set of languages to which a favoring theory should apply can be restricted further than I have suggested on an *a priori* basis. His example is a language whose single constant represents a tuple of 10 emeralds and whose three predicates represent the properties “tuple elements 1 through 6 are green,” “tuple elements 7 through 9 are green,” and “tuple element 10 is green.” Bartha suggests that this language can be ruled out (and in particular that a formal favoring theory need not treat permutations of its predicates identically) on the grounds that some predicates talk about more objects than others, or alternatively that some atomic sentences convey more information than others.

My response is that the proper individuation of objects for purposes of projection should be determined by our empirical evidence. For example if we are making predictions about emeralds, is evidence more significant if it is about more emeralds or if the emeralds it talks about have a greater total mass? (What if emeralds 7 through 9 in Bartha’s example together weigh more than emeralds 1 through 6?) Put another way, should a scientific language for predicting features of emeralds have constants that pick out individual emeralds or individual 1-gram chunks of emerald mass? I do not see how this kind of question could be answered *a priori*, but if we try to answer it using empirical evidence and then develop our evidential favoring relation from languages that individuate objects properly, we will wind up in the same kind of circle as we saw with *np*. Once more something beyond our evidence will have to play a decisive role.

References

1. Hempel, C.G. Studies in the logic of confirmation (I). *Mind* **1945**, *54*, 1–26.
2. Hempel, C.G. Studies in the logic of confirmation (II). *Mind* **1945**, *54*, 97–121.
3. Carnap, R. *Logical Foundations of Probability*; University of Chicago Press: Chicago, IL, USA, 1950.
4. Titelbaum, M.G. Not enough there there: Evidence, reasons, and language independence. *Philos. Perspect.* **2010**, *24*, 477–528.
5. Kolmogorov, A.N. *Foundations of the Theory of Probability*; Chelsea Publishing Company: New York, NY, USA, 1950; Translated by Nathan Morrison.
6. Goodman, N. *Fact, Fiction, and Forecast*; Harvard University Press: Cambridge, MA, USA, 1979.
7. Jaynes, E.T. Information theory and statistical mechanics I. *Phys. Rev.* **1957**, *106*, 620–630.
8. Jaynes, E.T. Information theory and statistical mechanics II. *Phys. Rev.* **1957**, *108*, 171–190.
9. Maher, P. Probabilities for two properties. *Erkenntnis* **2000**, *52*, 63–91.
10. Maher, P. Probabilities for multiple properties: The models of Hesse and Carnap and Kemeny. *Erkenntnis* **2001**, *55*, 183–216.
11. Seidenfeld, T. Entropy and uncertainty. *Philos. Sci.* **1986**, *53*, 467–491.
12. Glymour, C. *Theory and Evidence*; Princeton University Press: Princeton, NJ, USA, 1980.
13. Hume, D. *A Treatise of Human Nature*, 2nd ed.; Oxford University Press: Oxford, UK, 1978.
14. Popper, K.R. *The Logic of Scientific Discovery*; Science Editions, Inc.: New York, NY, USA, 1961.
15. Miller, D. *Out of Error: Further Essays on Critical Rationalism*; Ashgate: Burlington, VT, USA, 2005.
16. Chandler, J. Contrastive Confirmation: Some Competing Accounts. Forthcoming in *Synthese*; DOI: 10.1007/s11229-010-9845-9.
17. Quine, W. Natural kinds. In *Ontological Relativity, and Other Essays*; Columbia University Press: New York, NY, USA, 1969; pp. 114–138.
18. Lewis, D. New work for a theory of universals. *Australas. J. Philos.* **1983**, *61*, 343–377.

© 2011 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>.)