Empirical Significance, Predictive Power, and Explication

[Penultimate draft, forthcoming in *Synthese*]

Jonathan Surovell

jonathansurovell@gmail.com
https://txstate.academia.edu/JonathanSurovell

Abstract
Criteria of empirical significance are supposed to state conditions under which (putative) reference to an unobservable object or property is "empirically meaningful." The intended kind of empirical meaningfulness should be necessary for admissibility into the selective contexts of scientific inquiry. I defend Justus's recent argument that the reasons generally given for rejecting the project of defining a significance criterion are unpersuasive. However, as I show, this project remains wedded to an overly narrow conception of its subject matter. Even the most cutting edge significance criteria identify empirical significance with predictive power, and thereby rule out vocabulary with legitimate scientific functions. In a nutshell, the problem is that there are ("shortcut") terms that reduce the computational burden of extracting predictions from theory, and that may therefore be scientifically useful, but that do not produce any *novel* predictions, and so are ruled scientifically inadmissibility by existing significance criteria. I spell out this objection by specifying shortcut terms that are ruled inadmissible by Creath's and Schurz's criteria. Having objected in this way to extant criteria, and to the equation of empirical significance with predictive power in general, I discuss an approach to defining empirical significance that is capable of avoiding my objection and, more ambitiously, that may break the cycle of "punctures and patches" that has plagued the project from the beginning. I gloss Goldfarb and Ricketts's idea of "case-by-case" delineations of empirically significant terms as the provision of special rather than general explications of the informal concept of empirical significance.

1. Introduction

Criteria of empirical significance are supposed to state conditions under which reference to an unobservable object or property ('electron', e.g.) is "empirically meaningful". The intended kind of empirical meaningfulness is supposed to be necessary for admissibility

into the highly selective contexts of scientific inquiry. In their heyday, the logical empiricists—the original champions of significance criteria—believed that these criteria would illuminate scientific methodology and also expose metaphysics as a kind of defective attempt to talk about unobservables. Of course, these two rationales are mutually independent and the project of defining a significance criterion—call this the *significance project*—does not depend on either one individually.

According to standard historical accounts, the logical empiricists failed decisively in their attempts to state significance criteria. This failure is said to have been exposed by the objections of Church (1949), Hempel (1951), Achinstein (1968), and Kaplan (1975). Thus the significance project has been almost completely supplanted by the view that empirical meaningfulness cannot belong to individual sentences, but only to larger networks of theories.

There is good reason to doubt this received view. Justus (2014) points out that it focuses disproportionately on the shortcomings of Ayer's principles of verification, and that "scrutiny of the criticisms actually made of later significance criteria, some of which differ radically from Ayer's, provides little support for" (Justus 2014, 416) the received view's general pessimism about the significance project. More importantly, proponents of significance criteria have successfully responded to the major objections that have been raised. Thus Carnap (1956a) addresses Church's (1949) and Hempel's (1951) objection, Creath (1976) addresses Kaplan's (1975), and Schurz (1991; 2014b) Achinstein's (1968); the latter two responses remain unscathed decades after their publications and, as I argue below, can be combined into a single criterion that retains the advantages of each.

So Justus is correct when he says that the significance project has been written off prematurely and for unpersuasive reasons. However, as I will show, this project remains wedded to an overly narrow conception of its subject matter. Even the most cutting edge and formally sophisticated significance criteria identify empirical significance with predictive power, and thereby rule out vocabulary with legitimate scientific functions. In a nutshell, the problem is that there are terms—I call them 'shortcut terms'—that reduce the computational burden of extracting predictions from theory, and that may therefore be scientifically useful, but that cannot be used to augment the containing theory's observational consequence class, and so are ruled scientifically inadmissible by existing significance criteria. In what follows, I will spell out this objection by specifying shortcut terms that are ruled inadmissible by Creath's and Schurz's criteria.

Having objected in this way to extant criteria, and to the equation of empirical significance with predictive power in general, I will discuss an approach to defining empirical significance that is capable of avoiding my objection and, more ambitiously, that may break the cycle of "punctures and patches" (Lewis 1988) that has plagued the significance project since its inception. The approach is inspired by Goldfarb and Ricketts's (1992) idea of "case-by-case" delineations of empirically significant terms; I gloss this idea as the provision of, in Carnap's (1937) terminology, (relatively) "special" rather than (fully) "general" explications of the informal concept of empirical significance.

2. Creath's Response to Kaplan's Objection

The project of defining empirical significance begins with the informal idea that theoretical science is distinguished from other kinds of discourse at least in part by the nature of its connection to observation. Carnap (1956a) suggests that the appropriate connection to observation occurs, in the first instance, at the level of sub-sentential terms: a term, he suggests, is empirically significant just in case it occurs in a sentence that makes a "difference for the prediction of an observable event" (Carnap 1956a, 49). This characterization of the connection is informal and awaits precise formulation; it is an *explicandum* waiting to be explicated. (I say more about the methodology of explication in section 8.) Many explicata have been proposed, and many of these have been "punctured" by counter-examples. I will not here rehearse the complete history of these punctures and the subsequent patches, but will instead begin with the current state of the art.

I begin, then, with Creath's (1976) criterion, and with a kind of linguistic apparatus that it presupposes, *viz.*, what Carnap called a "language for science". This is an axiomatization of some body of scientific theory in an artificial language whose grammatical, proof theoretic, and semantic rules have been explicitly stated. A Carnapian language for science is divided into an observation language L_O and a theoretical language L_T, where all the descriptive terms of L_O—the set of which is V_O—refer to directly observable properties, processes, or individuals. L_T's sentences contain "theoretical terms"—the set of which is V_T—whose ties to observation are more indirect.[1]

---

[1] Carnap provides a "rough explanation" of the notion of an observation term, which grounds the distinction between V_O and V_T:

> A predicate 'P' of a language L is called *observable* for an organism (e.g. a person) N, if, for suitable arguments, e.g. 'b', N is able under suitable circumstances to come to a decision with the help of few observations about a full

Carnap gives as examples of theoretical terms terms that refer "to micro-particles like electrons or atoms, to the electromagnetic field or the gravitational field in physics, to drives and potentials of various kinds in psychology" (Carnap 1956a, 38). A language for science is used to state *theoretical postulates* T and *correspondence rules* C. T is the conjunction of the posited natural laws containing only terms of V_T; C is the conjunction of postulates that contain both observation and theoretical terms. One of the purposes of a significance criterion is to identify conditions of admissibility into V_T (Carnap 1956a, 38-39).

Creath proposed his criterion partially in response to Kaplan's (1975) contention that significance criteria should be insensitive to *deoccamization*. A language is deoccamized when its theoretical postulates T′ and correspondence rules C′ are formed from some other theoretical postulates T and correspondence rules C "either by replacing all occurrences of certain elements of V_T by the conjunction of two new primitive constants of the same type, or by replacing all occurrences of certain elements of V_T by the disjunction of two new primitive constants of the same type" in T and C (Kaplan 1975, 91). Kaplan assumes that significance criteria should be *insensitive* to deoccamization, i.e., that if a term M of L is reckoned significant by a given criterion, then this criterion should reckon M_1 and M_2 significant when M_1&M_2 replaces M in a

---

sentence, say 'P(b)', i.e. to a confirmation of either 'P(b)' or '~P(b)' of such a high degree that he will either accept or reject 'P(b)'. (Carnap 1936-7, 454-455)
Degrees of observationality, so understood, lie on a continuum. Carnap proposes that the sharp line between V_O and V_T be drawn at the most convenient point on the continuum (Carnap 1966, 226).

The distinction between observational and theoretical language is contested, but I will presuppose it for the sake of argument. For an argument that Carnap only requires an innocuous version of this distinction, see Justus (2014, 420).

deoccamization of L. I will later argue (following Justus (2014)) that this assumption is in need of argument. But well motivated or not, Kaplan's criticism contributed to the view that Carnap's (1956a) criterion, which was sensitive to deoccamization, was inadequate, and to increased skepticism of the significance project.

Creath's (1976) criterion, which results from a small amendment to Carnap's (1956a), was the first to be insensitive to deoccamization. His criterion is given in a sequence of three clauses. The first of which, D1′, defines empirical significance for primitive terms[2] relative to a large theoretical context—a context that includes, in particular, other assumptions in L_T in addition to T and C. D1′ requires of terms that inserting them into this large theoretical context would, through the formulation of a sentence containing them, allow us to non-vacuously derive[3] an observational prediction that we could not otherwise derive in this context.

> D1′. A [primitive] term 'M' is *significant* relative to the class *K* of terms, with respect to L_T, L_O, T, and $C =_{df}$ the terms of *K* belong to V_T, there is a class *J* of terms to which 'M' belongs and such that each term of *J* belongs to V_T but not to *K*, and there are three sentences S_J and S_K in L_T and S_O in L_O, such that the following conditions are fulfilled:
>
> (a) S_J contains only terms of *J* as its descriptive terms.

---

[2] Creath specifies that the criterion is meant to apply to primitive terms and suggests that "[f]or defined terms we need only require in addition that every descriptive term in the definiens thereof must be *antecedently* shown to be significant" (1976, 395).

[3] Carnap's (1956a) criterion has an analogous requirement that the observation sentence be logically derivable within the theoretical context. He seems to treat this requirement as a simplifying assumption. A more complete formulation of his criterion would encompass also probabilistic relations between theory and observational prediction (Carnap 1956a, 49). This point can be applied *mutatis mutandis* to subsequent criteria.

(b) The descriptive terms of S_K belong to K.

(c) The conjunction $S_K \bullet S_J \bullet T \bullet C$ is consistent (i.e., not logically false).

(d) S_O is logically implied by the conjunction $S_J \bullet S_K \bullet T \bullet C$.

(e) S_O is not logically implied by $S_K \bullet T \bullet C$.

(f) There is no set J', where J' is a proper subset of J, such that there are three sentences $S_{J'}$ and $S_{K'}$ in L_T and $S_{O'}$ in L_O, such that the following conditions are fulfilled:

(fa) $S_{J'}$ contains only terms of J' as its descriptive terms.

(fb) The descriptive terms in $S_{K'}$ belong to K.

(fc) The conjunction $S_{J'} \bullet S_{K'} \bullet T \bullet C$ is consistent (i.e., not logically false).

(fd) $S_{O'}$ is logically implied by the conjunction $S_{J'} \bullet S_{K'} \bullet T \bullet C$.

(fe) $S_{O'}$ is not logically implied by $S_{K'} \bullet T \bullet C$. (Creath 1976, 398)

Creath's response to Kaplan is contained in this first clause of the definition. The role of the class J in D1′ "secure[s] in a single step whole collections (sets) of terms" (Creath 1976, 397), including the pairs of predicates that replace single predicates through deoccamization, within the sphere of the empirically significant.

Building on D1′, D2′ removes the latter's relativization to a class K of terms antecedently certified empirically significant:

D2′. A [primitive] term 'M' is *significant* with respect to L_T, L_O, T, and $C =_{df}$ there is a sequence of sets $J_1, ..., J_n$ of terms of V_T such that 'M' belongs to J_n, and every member of every set J_i ($i = 1, ..., n$) is significant relative to the union of J_1 through $J_{(i-1)}$, with respect to L_T, L_O, T, and C. (Creath 1976, 398)

The idea is to start with a set J of terms that occur in a sentence p that non-vacuously implies an observation sentence with the help of only the relatively small theoretical context of T and C; sentences like p "directly" entail observation sentences. D2′ certifies the members of J empirically significant relative to this small context. Now p can be used to help certify the terms composing set K as empirically significant relative to the same smaller context: the members of K then just need to be certified significant, in the sense of D1′, relative to J, L_T, L_O, T, and C.

The third and final clause of Creath's definition defines sentence significance in terms of the component expressions' significance (in the sense of D2′).

D3[′]. An expression A of L_T is a *significant sentence* of L_T = $_{Df}$

a. A satisfies the rules of formation of L_T,

b. every descriptive constant in A is a significant term (in the sense of

D2[′]).[4] (Carnap 1956a, 60; quoted in Creath 1976, 394)

---

[4] It is worth noting that by beginning with term significance and working up to sentence significance, the definition avoids Church's (1949) famous counter-example to Ayer's (1946) criterion. Let O_1, O_2, and O_3 be primitive observation predicates and A (for 'Absolute') a primitive "metaphysical" predicate, i.e., a predicate every attribution of which the logical empiricists are determined to reject as nonsense. Translating Church's argument from sentential to predicate logic, we get the sentences

    1. $[(\sim(x)O_1x)\&(x)O_2x]\lor[(x)O_3x\&(\sim(x)Ax)]$

and

    2. $(x)O_1(x)$

which together imply the observation sentence $(x)O_3x$. According to Ayer's criterion, a sentence is "directly verifiable" if, with other observation sentences, it entails an observation sentence that these other observation sentences alone do not entail (Ayer 1946, 16). So sentence (1) is directly verifiable. And a sentence is empirically significant if, together with other directly verifiable sentences, it entails an observation sentence that these other directly verifiable sentences alone do not entail. Therefore, since $(x)Ax$ and (1) together entail $(x)O_2x$, and since $(x)Ax$ is essential to this entailment, $(x)Ax$ is

3. Schurz's Response to Achinstein's Objection

So one of the putative refutations of the significance project has been addressed. I turn

now to a separate solution to a separate problem for significance criteria: Schurz's (1991;

2014b) solution to Achinstein's (1968) problem of variance over logical equivalents.

> Achinstein assumes

> Principle *B*: If a sentence (or conjunction of sentences) *S* is such that the

> occurrence of term *M* in *S* suffices to guarantee that *M* is significant, then the

> occurrence of *M* in any sentence logically equivalent to *S* also suffices to

> guarantee that *M* is significant. (1968, 78)

He then shows that Carnap's (1956a) criterion violates this principle; his argument

applies *mutatis mutandis* to Creath's (1976) criterion. Consider a language with

correspondence rule $S_m \rightarrow S_o$, where S_o is an observation sentence and S_m is a sentence

whose only descriptive term is the theoretical term M. Relative to this language, M is

significant according to Creath's criterion—it is the only theoretical term in a sentence

that, in conjunction with the postulates, allows us to non-vacuously derive S_o (which we

would not otherwise be able to derive). But now consider a language with theoretical

postulate is S_m and observational postulate S_o. Creath's criterion reckons M non-

---

empirically significant. Of course, any sentence can be put in the truth-functional position
of (x)Ax in (1) and thereby rendered empirically significant by Ayer's (1946) criterion.

      Creath's (1976) criterion avoids this problem. The set of primitive descriptive
predicates occurring in (1) has a proper subset—$\{O_1, O_2, O_3\}$—whose members occur in
sentences that non-vacuously imply (x)O_2x. So A's role in the derivation of (x)O_2x does
not make it empirically significant, according to this criterion. And since the criterion
builds sentence significance out of term significance, (1) is not empirically significant
according to Creath's criterion. (The basic strategy here originates in Carnap's (1956a)
criterion.)

significant relative to this language. However, $S_m \& (S_m \rightarrow S_o)$ is logically equivalent to $S_m \& S_o$. So Creath's criterion contradicts Principle B.

Achinstein's example poses a trilemma for significance criteria: deny, counter-intuitively, Principle B; reckon M significant with respect to both conjunctions and accept, counter-intuitively, that $S_m \& S_o$ is significant; or reckon M non-significant with respect to both conjunctions and accept, counter-intuitively, that $S_m \& (S_m \rightarrow S_o)$ is non-significant (call this the *non-significance horn* of the trilemma).

Schurz argues that the solution to this trilemma lies in recognizing that $S_m \rightarrow S_o$ is an "irrelevant consequence of" S_o. He defines a novel notion of relevance for this purpose for an arbitrary theory T, with 'theory' here understood in a broad sense, encompassing both Carnapian theoretical postulates, correspondence rules, and observation statements:

> [Definition]   Sentence p is a *relevant consequence* of theory T if and only if no predicate in p "is replaceable on some of its occurrences by any other predicate of the same arity, salva validitate of" the argument from T to p (1991, 409).[5]

> [Definition]   The set of T's *relevant consequence elements*, $(T)_r$, $= \{$p: (i) p is a relevant consequence of T and (ii) there is no conjunction $q_1 \& \ldots \& q_n$ such that (a) p is equivalent to $q_1 \& \ldots \& q_n$, (b) $q_1 \& \ldots \& q_n$ is a relevant consequence of T, and (c) each q_i is shorter than p$\}$.

Schurz uses the notion of a theory's relevant consequence elements to explicate empirical significance through the following two definitions:

> [Definition]   A *minimal empirically equivalent axiomatization of* $(T)_r$ is any subset Δ of $(T)_r$ such that Δ and T imply exactly the same observation sentences

---

[5] Schurz (1991) gives a more formally rigorous inductive definition of 'relevant consequence' on pp. 410-412.

and no subset of $(T)_r$ of smaller cardinality than Δ has exactly the same observational consequences as T.

[Definition]   A theoretical term t is *empirically significant* relative to T if and only if there is no minimal empirically equivalent axiomatization of $(T)_r$ which does not contain t (Schurz 1991, 425).

Applying these definitions to Achinstein's argument, we find that $\{S_o\}$ is a minimal empirically equivalent axiomatization of both $(\{S_m \& S_o\})_r$ and $(\{S_m \& (S_m \rightarrow S_o)\})_r$. And since S_o does not contain M, the term is not empirically significant relative to either conjunction.

Although Schurz's criterion reckons M non-significant relative to $S_m \& (S_m \rightarrow S_o)$, the criterion is not impaled on the non-significance horn of Achinstein's trilemma—rather, it grasps the horn. The scientific point of theoretical jargon is to allow us to in some way systematize or unify a relatively large body of observational knowledge on a manageably small basis (Schurz 1991, 424-425; Schurz and Lambert 1994). Theoretical postulates with meager observational consequences could just as well be replaced by a list of these consequences themselves. So for a postulate of theoretical science to earn its keep, it would, roughly speaking, have to increase the theory's observational consequences by at least as much as it increases the size or complexity of the theory; as Schurz puts it,

[i]t is the main achievement of a scientific theory to *unify* its empirical consequences, and a theoretical term *t* is empirically significant within T iff the empirical *unification* provided by T were not possible in a 'subtheory' of T which does not contain *t*. (1991, 425; emphasis added)

This expectation of theoretical postulates is reflected in Schurz's requirement that for theory T's terms to be significant, there must be no set of sentences from which the terms are absent that achieves empirical equivalence with T by means of fewer sentences than occur in $(T)_r$. This is where M falls short with respect to $S_m \& (S_m \to S_o)$.[6]

So the two biggest thorns in the side of the significance project—the problems of deoccamization and variance over logically equivalent theories—have been removed. And, happily, the two solutions are compatible. Schurz's insight fits neatly into Creath's D1′ as the further sub-clause:

> (g) There is no minimal empirically equivalent axiomatization of
>
> $(S_K \bullet S_j \bullet T \bullet C)_r$ which does not contain 'M',

thus immunizing Creath's criterion against Achinstein's objection. Call this amended version of the first clause of the criterion D1″. D1″ is, of course, more demanding than D1′; the former but not the latter excludes, for example, M relative to $S_m \& (S_m \to S_o)$. The criterion due to Schurz that I just discussed already preserves empirical significance over deoccamization. If we uniformly replace occurrences in T of an empirically significant predicate with a disjunction of predicates, then the disjunction of predicates will occur in all of T's minimal empirically equivalent axiomatizations.

---

[6] Schurz, writing with Lambert, indicates one way for theories to meet this requirement, *viz.*, inclusion of "many T_i's [i.e., laws of the form $\forall x(F_i x \supset Tx)$ for various *i*] plus many T'_k's [i.e., laws of the form $\forall x(Tx \supset G_k x)$ for various *k*]" (Schurz and Lambert 1994, 88), where each F_i and each G_k is an observational predicate. They comment that

> [t]his is the typical situation where one conjectures in science that there is a certain 'intrinsic' and not directly observable property of objects, say $T(x)$, common to all [observational] properties $F_1, \ldots, F_m$, which has as its empirical effects the [observational] properties $G_1, \ldots, G_n$. $T(x)$ may e.g. mean that '*x* has metallic structure'; then the F_i are different kinds of metals, and the G_k are typical properties of metals. (Schurz and Lambert 1994, 87)

But if, like Kaplan, one is convinced that deoccamizations of empirically significant theories must be empirically significant, then one might think Schurz's criterion supersedes Creath's, and that there is no point in fusing the two into D1″. The purpose of Creath's criterion is insensitivity to deoccamization, but Schurz's criterion has this property. On the other hand, Schurz's criterion addresses Achinstein's objection, whereas Creath's does not.

However, there are at least two properties of Creath's criterion that are preserved in D1″ and that are absent from Schurz's criterion. Whether D1″ is worth considering will depend on the appeal of these properties. First, Creath's definition, like Carnap's, accords the postulates T and C a special role; it assumes that there is (are)

> some part(s) of theory that observation bears on directly, without the mediation of any other theoretical terms or statements. [D2′] reflects this requirement by assuming at least one theoretical term is empirically significant through a direct connection with observation, that is, unfacilitated by other theoretical terms (when $[\{J_1, \ldots, J_{i-1}\}]$ is null). (Justus 2014, 423)

As I discussed above, the "direct connection" intended here is occurrence in a sentence that, together with T and C alone, non-vacuously implies an observation sentence that T and C do not imply. Schurz's criterion does not make any such requirement of theories. It therefore does not require any machinery analogous to the classes of terms J and K in Creath's criterion. Second, D1′(a) requires that S_K, the auxiliary sentence antecedently certified empirically significant, contain only theoretical terms. There is no such requirement in Schurz's criterion.

Schurz (1991, 425) proposes the significance criterion I have ascribed to him as necessary *and* sufficient. But in his (2014b, 312), he treats it as one of two independent, merely necessary conditions, the second condition being:

"[a] theoretical concept τ… is *empirically significant* in [theory] T *only if* T entails a law of correspondence for τ" (2014b, 312),

where a law of correspondence for a term τ is "a lawlike (synthetic) sentence of the form $\forall x \forall t(A_i xt \rightarrow (\tau(x) \leftrightarrow R_i xt))$, and where $A_i xt$ and $R_i xt$ are [observational] or pre-theoretical expressions of T in [theory net] N" (2014b, 307), are descriptive and distinct from τ, and where i = (i,…, n) for some whole number n. A descriptive term τ is *pre-T-theoretical* just in case there is a theory T* that entails a law of correspondence for τ and whose observational vocabulary is a subset of T's. A theory T's *net* is an intricate set- and model-theoretic object—its make-up need not concern us—that represents different ways of specifying T's "intended applications" and restricting T's posits (Schurz 2014a, 1534). Note that this notion of a *law* of correspondence is more restrictive that Carnap's notion of a correspondence *rule*.

I have overlooked Schurz's correspondence law condition up until now because it is not among the targets of my objection. However, its eligibility to be combined with the other conditions further increases the range of options for the proponent of significance criteria. In particular, those who do not share Kaplan's intuition that deoccamizations of significant terms are themselves significant could include Schurz's correspondence law condition to rule out deoccamized terms (Schurz 2014b, 312-313). Also, my account of shortcut terms, below, will make use of the second of Schurz's conditions.

Schurz's first, minimal empirical equivalence criterion, as well as Creath's criterion and D1″, adhere to what I call *the predictive power conception* of empirical significance: all of these criteria certify as empirically significant only terms that allow us to augment our total theory's observational consequence class (where the "total theory" includes general theoretical postulates, auxiliary hypotheses about the experimental setting, etc.). They are not the only criteria to do so—their predecessors, Carnap (1936-7) and Ayer (1946), also adhere to this conception.[7] But as I argue in the following section, this conception does not do justice to the diversity of contributions to scientific inquiry that language can make.

4. The Problem of Shortcut Terms

A problem with Creath's and Schurz's criteria, and with the predictive power conception more generally, is that they give negative verdicts on terms that, though they do not augment the theory's predictive power, make it easier to determine which predictions are already implicit in a body of knowledge. I call such terms *shortcut* terms, as they allow us to chart a "shortcut" from theory to prediction. The informal gloss just stated suggests three conditions as collectively sufficient for a term's being of this kind:

[SC]   M is a shortcut term with respect to language L, theory T, and

correspondence rules C containing M if

(a) Any observation sentence implied by T&C in L is implied by T in L,

(b) It is easier, given certain possible states of knowledge, to derive some

observational prediction from T&C in L than from T in L, and

_____

[7] As does Wright's (1986) criterion. I pass over this criterion because, unlike those of Creath and Schurz, it currently suffers from unpatched punctures (Yi 2001).

(c) C indicates how to observationally detect the presence or magnitude of M in certain contexts.

Condition SC(a) will select terms that predictive power conceptions reckon non-significant. However, terms that meet the other two conditions will have a strong claim to scientific legitimacy.

By a derivation's being *easier* for a given agent than another I mean, roughly, that fewer inferences or less searching (for, e.g., the solution to an equation) would be, in all likelihood, required for the agent to know the premises or prove the conclusion of the derivation. For present purposes, knowing a proposition may be understood to be a matter of having derived it from the postulates of one's language or achieved a sufficiently high credence in it by means of empirical evidence. Computational complexity theory's notions of computational time and space may be taken as precisifications of the notion of ease of derivation, so my objection will not hinge on a derivation's being easier for some particular agent who idiosyncratically works better with one set of postulates than another.[8]

Condition SC(c) ensures that scientists' applications of shortcut terms will be empirically constrained. It should thereby forestall the objection to a previous definition that lacked this condition, which was put to me by Gerhard Schurz: 'God' might be a shortcut term given the "theory" 'everything obeys God's will and if Newtonian mechanics is true, then God wills that the planets orbit the sun elliptically'. This "theory" would simplify the task of deriving the ellipticality of the planets' orbits from Newtonian mechanics, as the derivation would require no mathematical calculation. But clearly this

---

[8] Christian Feldbacher-Escamilla expressed to me the worry that, without a condition like this, something could count as a shortcut term on the basis of such an idiosyncracy.

does not make 'God's will' a legitimate scientific term. To derive the predictions in this way would be, to borrow Russell's famous wording, to steal the results that Newtonian astronomers earned through honest toil. The requirement that God's will be observationally accessible would pre-empt such useless postulation.

We might interpret condition SC(c) in a variety of ways. For the sake of concreteness I interpret it as the requirement that the theory entail a law of correspondence, in terms of Schurz's (2014a; 2014b) sense, for the term.

The notion of a shortcut term is not a purely formal or semantic one; its reference in SC(b) to the agent's states of knowledge makes it partially pragmatic. This sets it apart from the concepts figuring in previous objections to significance criteria. But on reflection, it should come as no surprise that significance criteria should have to contend with the pragmatic aspect of language: significance criteria are supposed to capture a dimension of scientific legitimacy, and some of the theoretical virtues guiding science are pragmatic.

5. The Problem Applied to Creath's and Schurz's Criteria

I will now introduce the predicate Q_1 as shortcut term that Creath's criterion would eliminate from its containing language L_1. L_1's primitive vocabulary includes

- theoretical predicates Q_1 and R, each of which applies to object-number pairs, where each object has at most one such Q_1-magnitude and at most one R-magnitude;[9]

- the terms of basic arithmetic;

---

[9] $Q_1(x,y)$ and $R(x,y)$ might each be defined as the functions $q_1(x)=y$ and $r(x)=y$, respectively. Thanks to an anonymous reviewer for pointing this out.

- observation predicates O_1 and O_2, which, like Q_1 and R, apply to object-number pairs;

- an infinite set of individual constants a, b, c, d, e,…;

- the classical truth-functional connectives; and

- first-order quantifiers and variables w, x, y, and z.

L_1 also contains the defined predicates E (even), P (prime), and < (strictly less than). Its theoretical postulates and laws of correspondence (in Schurz's (2014a) sense) are

[T] $(w)(x)[((R(x,w)\&Ew\&3<w)\&\exists y\exists z(Py\&Pz\&w=y+z))\leftrightarrow O_1(x,w)]$

[CR] $(x)(y)[O_1x\rightarrow(R(x,y)\leftrightarrow O_2(x,y))]$

[CQ] $(x)(z)(w)[R(x,w)\rightarrow((Q_1(x,y)\&Py\&Ew\&3<w)\leftrightarrow O_1(x,w))]$

L_1's axioms also include Peano arithmetic and its inference rules the standard introduction and elimination rules for the classical truth-functional connectives, first-order quantifiers, and identity.

Q_1 meets clause SC(a) with respect to L_1, which is to say that it is reckoned non-significant by Creath's criterion, as it yields no additional observational predictions. T, which does not feature Q_1, states that if something's R-magnitude (the number to which it stands in the R-relation) satisfies Goldbach's conjecture,[10] then it also stands in the O_1 relation to this magnitude. So if Goldbach's conjecture is true, as it seems to be, then T entails that anything whose R-magnitude is an even integer x greater than three has O_1-magnitude x. CQ, which does feature Q_1, says that if something's R-magnitude is an even integer greater than three, *and* if its Q_1-magnitude is prime, then it is O_1-related to its R-

---

[10] Goldbach's conjecture says that every even integer greater than three is the sum of two primes.

magnitude. And these are all the observational consequences of the two postulates.
Therefore, for any x, y, and z, if O is an observational consequence in L_1 of
$CQ\&Q_1(x,z)\&R(x,y)$ then it is also a consequence of $T\&R(x,y)$; but the converse does
not hold. This lack of predictive power is responsible for Q_1's failure to meet $D1'(e)$.

That Q_1 satisfies SC(b) can be illustratedby an agent who

[A1] Does not justifiably believe Goldbach's conjecture, and

[A2] Justifiably believes that the set of observation statements entailed by CQ,
together with the relevant beliefs about Q_1 and R, is a subset of the set of
observation statements entailed by T and these same beliefs about R.

While belief states satisfying A1 and A2 may be unlikely, they are perfectly coherent.
Such a state could arise through the agent's

[A3] Having checked T and CQ for enough of the same values for A2 to be the
case, but

[A4] Not having checked T for enough values to justifiably believe Goldbach's
conjecture.

Such an agent knows that Goldbach's conjecture holds for even integers greater than
three that are something's R-magnitude and that, when added to a prime number, yield
this same thing's Q_1-magnitude; but she does not know that it holds for all integers. She
justifiably believes that what she has found to hold of even integers greater than three that
stand in the right Q_1- and R-relations to things cannot be assumed to hold of all even
integers greater than three. If the coherence of this last stipulation should require further
argument, we can assume also that the R-magnitudes which she has found to satisfy
Goldbach's conjecture all have some conspicuous number-theoretic property—e.g., being

the sum of a prime and a semiprime (i.e., the product of two primes).

Now suppose that this same agent knows that R(a,c) and that Q_1(a,b)&R(a,c), where b is prime and c is a very large even number greater than three. She suspects that O_1(a,c) might be an observational consequence of her theoretical postulates and knowledge, but has not rigorously derived it. She can see that it will be easier to do so if she relies on Q_1(a,b)&R(a,c)&CQ rather than R(a,c)&T. On the latter approach, she would have to rely on T. Given that, by hypothesis, she cannot get the antecedent of the appropriate instantiation of T through a direct appeal to Goldbach's conjecture, in order to become entitled to use it as a premise, she would need to find a pair of primes that add up to c (a *Goldbach partition* of c) or, in the case of the agent discussed above, a prime and a semiprime that add up to c (a *Goldbach semi-partition* of c). But Goldbach (semi-)partitions can be difficult to find and verify for large values. By contrast, if she tries to derive O_1(a,c) from Q_1(a,b)&R(a,c)&T&CQ, then she could rely on CQ instead of T, and would only need to show that the b is prime. And primeness is generally easier—less computationally complex—to establish than possession of a Goldbach (semi-)partition: to find a Goldbach partition of x, one must verify that two numbers are each prime and add up to x.[11]

Finally, CQ ensures that Q1 satisfies SC(c) with respect to L_1 and Creath's criterion.

---

[11] To be concrete, we might suppose that the agent uses the algorithm developed by e Silva, Herzog, and Pardi (2013), which has computed Goldbach partitions for every even number between six and 4×10^{18} (which entails the validity of the conjecture up to 8.37×10^{26}). This algorithm contains a sub-algorithm (a sieve of Eratosthenes) that generates the prime numbers. And such a sub-algorithm is of course less complex, in regards to computational time or space, than its containing super-algorithm. So since using Q_1 would require a sub-algorithm and eschewing it the super-algorithm, using Q_1 eases the computational burden of deriving the prediction in question.

Schurz's criterion also deems Q_1 non-significant, but in part by design: T's observational consequences are too *few* to bring the predicate up to the criterion's standards. From Schurz's perspective, to include in the language a purely computational aid to R would be to compound needless complications, as R itself unifies too little observational knowledge to earn its keep. So Q_1 does not ground an objection to Schurz's criterion or to D1''.

But Q_2 of L_2 and theory T stands in an analogous relationship to Schurz's criterion and D1''. L_2 is the same as L_1 except that it replaces Q_1 with Q_2, and contains the additional observation predicates O_3, O_4, O_5, O_6, O_1', and O_2' and theoretical predicate S. T is composed of the following postulates and laws of correspondence:

[T1-6] $(x)(y)(S(x,y)\leftrightarrow O_i(x,y))$, i = (1,..., 6)

[T7] $(w)(x)[((R(x,w)\&Ew\&3<w)\&\exists y\exists z(Py\&Pz\&w=y+z))\rightarrow S(x,w)]$

[T8] $(x)(y)(O_1'(x,y)\rightarrow R(x,y))$

[T9] $O_1'(a,b)\&Eb\&3<b$

[T10] $(x)(y)(O_2'(x,y)\rightarrow R(x,y))$

[T11] $O_2'(c,d)\&Ed\&3<d$

[T12] $O_2'(a,e)\&Pe$

[T13] $(x)(y)(O_2'(x,y)\rightarrow Q_2(x,y))$

[T14] $(w)(x)(z)[((R(x,w)\&Ew\&3<w)\&(Q_2(x,z)\&Pz))\rightarrow O_1(x,y)]$

T1&…&T12 does not contain Q_2 and is a minimal empirically equivalent axiomatization of the set of relevant consequence elements of T. So Q_2 meets SC(a) and by Schurz's criterion is non-significant relative to T.

Q_2 also meets SC(b) with respect to this criterion and language; the reasons why

are analogous to the reasons, given above, for Q_1's satisfying these conditions with respect to Creath's criterion. Consider an agent who

[A5]   Does not justifiably believe Goldbach's conjecture, and

[A6]   Justifiably believes that the set of observation statements entailed by T9&…&T14 is a subset of the set of observation statements entailed by T1&…&T12.

For this agent, it will be easier to predict $O_1(a,b)$ by deriving it from the Q_2-containing T9&…&T14 than from the Q_2-deficient T1&…&T9; the latter but not the former derivation would require a difficult to find Goldbach partition.

Finally, T13 satisfies SC(c) for Q_2 in T with respect to Schurz's criterion.

Creath's criterion instructs us to remove Q_1 from L_1 and Schurz's criterion and D1″ to remove Q_2 and from L_2 and T, respectively. However, Q_1 and Q_2's satisfaction of SC(b) and SC(c) with respect to the relevant criteria, languages, and theories calls both criteria's verdicts into question. The agents discussed above—the ones that meet conditions A1 and A2, and A5 and A6 respectively—could determine their postulates' observational consequences by means of the postulates containing the shortcut terms and thereby spare themselves the trouble of finding Goldbach partitions. This is a good reason for these agents to keep the shortcut terms in their languages or theories. It is widely recognized by philosophers of science that this kind of ease of derivation is a theoretical virtue. Some consider it to be a kind of simplicity (Baker 2013; Barrios 2015), while others take simplicity to consist in nothing more than it (Peirce 1935; Lindsay 1937; Ludlow 2011). Furthermore, shortcut terms are scientifically useful from the standpoint of what I will call 'Carnap's pragmatism', which is an important rationale for the

significance project, and which I discuss in section 7.

Are there examples of shortcut terms that have contributed to actual scientific inquiry? Psillos claims that the concept of asymptotic freedom in quantum chromodynamics helps to "establish connections between other theoretical terms" without "contributing to the derivation of fresh observational consequences" (Psillos 2008, 137). While the theoretical role ascribed to asymptotic freedom here is similar to that of a shortcut term, there is a difference: shortcut terms not only establish connections between theoretical terms, but also facilitate derivations of observational predictions. Of course, theoretical connections might have this kind of collateral impact on prediction, but to show as much would require further discussion, which I am not in a position to give. Asymptotic freedom might be a counter-example to predictive power criteria on the basis of its inter-theoretical connections alone, depending on whether and how it makes physical theories easier to work with.

An anonymous reviewer suggested a method for constructing shortcut terms by means of the concept of gamma rays. Let T_1 and T_2 be theories and M_1 and M_2 theoretical predicates, such that $T_1\&M_1(a)$ and $T_2\&M_2(a)$ are empirically equivalent, and such that empirically establishing $M_1(a)$ involves detecting a microparticle, which is relatively difficult to do, whereas empirically establishing $M_2(a)$ involves detecting a gamma ray, which is relatively easy to do. M_2 is therefore a shortcut term.

The reviewer also pointed out that this example is highly general, in that it puts few restrictions on the theories or predicates involved, requiring only on the inclusion of basic particle and wave physics.

I conjecture—though I will not argue—that the shortcut term objection to

predictive power conceptions can be generalized in another way: for any reasonable explicatum E of the notion of essential contribution to an observational prediction, there will be some possible shortcut term that does not satisfy E but that would allow its users to circumvent searches for Goldbach partitions in their derivations of their theories' observable consequences.

6. Responses and Counter-Responses

One might respond on behalf of the predictive power conception that the advantages of including Q_1 and Q_2 in their respective languages are relatively minor and are outweighed by the increased syntactic complexity that the terms and their associated correspondence rules would engender. The premise of this response may be true, but the conclusion does not help the predictive power conception. The objection points out that charting a shortcut to an observation statement is not sufficient for admissibility into the theoretical language. But shortcut terms will pose a problem for the predictive power conception so long as there is *pro tanto* reason to admit them. Significance criteria are supposed to state standards for what counts as a scientific theory; they should not provide any guidance beyond this concerning whether to use a minimally acceptable term. And if there can be a *pro tanto* reason to admit a term, then the term satisfies the minimal standards that significance criteria aim to capture.

Justus (2014) contrasts Church's objection to Ayer's criterion with Kaplan's (1975) objection to Carnap's criterion: whereas the former objection identifies a decisive technical flaw, the latter relies on the controversial philosophical intuition that deoccamization preserves significance. Justus then contends that there is no "plausible

basis for claiming deoccamization should preserve empirical significance" (Justus 2014, 426). He cites Glymour's contention that deoccamized theories are

> just the sort of theories that theorists abjure; physicists say they have 'redundant quantities' or 'unobservable quantities' and regard them with suspicion and worse… without appropriate restrictions, the hypothetico-deductive view [which is reflected in Carnap's informal characterization of empirical significance] is committed to the legitimacy of deoccamized theories, and that commitment may not accord with either intuitive judgment or scientific practice. (Glymour 1980, 32; quoted in Justus 2014, 426)

It is true that the problem of shortcut terms, like Kaplan's and Achinstein's and unlike Church's counter-examples, does not expose a decisive formal flaw in the target significance criteria. However, unlike Kaplan's objection, it does not rely on a philosophical intuition that the proponent of significance criteria can easily abandon. As noted above, the legitimacy of shortcut terms follows from the fact that ease of use is a theoretical virtue. I as I discuss next, it also follows from a leading rationale for significance criteria, *viz*. Carnap's.

7. Carnap's Pragmatic Rationale for Significance Criteria

I want to now consider the reasons for seeking a significance criterion in the first place. I focus on the reasons given by Carnap, which have been highly influential. As I will show, Carnap's conception of the significance project entails the admissibility of shortcut terms. And it is also conducive to the proposal, which I make in the next section, to give various "special" explications empirical of significance, for various languages and theories.

I will being relying on an interpretation I have developed at greater length elsewhere (Surovell 2017), according to which Carnap's conception of language, which encompasses his empiricism, derives from his *pragmatism*. The latter involves two stances. The first treats language as a tool that can be used for a chosen purpose:

> the choice of a certain language structure... is a practical decision like the choice of an instrument; it depends chiefly upon the purposes for which the instrument—here the language—is intended to be used and upon the properties of the instrument. (1956b, 43)

The proposal to use a particular language is thus practically justified to the extent that it achieves this purpose *efficiently*: "[a]s a hammer helps a man do better and more efficiently what he did before with his unaided hand, so a logical tool helps a man do better and more efficiently what he did before with his unaided brain" (Carnap 1943, viii).

On this view, we are practically justified in using a significance criterion in so far as it helps us select the vocabulary that would be minimally acceptable for achieving the purpose that we have chosen for language; it should admit all and only the terms that advance our purpose in some way.

The second component of Carnap's pragmatism specifies which linguistic aim is most amenable to the scientific enterprise; it says that the language of science is an instrument to help with inferences to and from observation reports. In his (1953a), Carnap conceives of the "formal" sciences—i.e. logic and mathematics—as auxiliaries to the "factual"—i.e. theoretical or empirical—sciences; "[*f*]*ormal science* has no independent significance, but is an auxiliary component introduced for technical reasons

in order to facilitate linguistic transformations in the *factual sciences*" (1953a, 127). He expresses this view again in his (1939, 2). Logic and mathematics thus serve an auxiliary, purely inferential function on behalf of the empirical part of the language for science, according to Carnap.

As we have seen, Carnap further sub-divides empirical science into theoretical and observational sub-languages. And he conceives of the postulates of the theoretical language as instruments for inferring observational predictions:

> [f]or an observer *X* to "accept" the postulates of *T* means here not simply to take *T* as an uninterpreted calculus, but to use *T* together with specified rules of correspondence *C* for guiding his expectations by deriving predictions about future observable events from observed events with the help of *T* and *C*. (1956a, 45)

In short, Carnap's pragmatism says that languages for science are "prediction machines" whose practical purpose is to facilitate inferences to and from observational predictions and reports.

I claimed that Carnap's pragmatism provides his rationale for a criterion of empirical significance. Here's how. As I discussed above, significance criteria are explicata for the informal notion of a term's making a "difference for the prediction of an observable event" (Carnap 1956a, 49). But from the point of view of Carnap's pragmatism, all terms that make such a difference make *some* contribution to a language's fulfillment of its inferential function in science. And only such terms make such a contribution; terms that make no difference for prediction would increase the language's complexity without any compensating pragmatic advantage. So we should not admit these terms into our language, and we should remove any that we might have

already admitted. Significance criteria expedite the Carnapian pragmatist's choice of a language for science by screening out terms that are useless given her aims.

In section 6, I considered the argument that shortcut terms' veneer of scientific legitimacy, like that of Kaplan's deoccamized predicates, is a mere philosophical intuition. We can now see how Carnap's pragmatism entails the scientific legitimacy of shortcut terms: the Carnapian pragmatist's purpose in using scientific language is to more efficiently perform inferences to and from observation sentences; but this is exactly what shortcut terms do.

8. Special Explication as an Alternative

Perhaps the hole that shortcut terms puncture in recent criteria can be patched. Still, given the persistence of the cycle of punctures and patches, it would be prudent for the advocate of significance criteria to consider alternatives to the assumptions that have so far guided her project. Specifically, I want to suggest, the significance project should not limit itself to the predictive power conception or to the kinds of highly general criteria that have so far been the norm. In this section, I will sketch an alternative to these prevailing assumptions.

My proposal is inspired by the following remarks by Goldfarb and Ricketts:

[t]he lack of such a general criterion [of empirical significance] is not as

damaging to Carnap's program as is commonly thought. Carnap can still criticize

what he calls metaphysics by first demanding that the linguistic framework in

which the metaphysical claims are made be laid out clearly; and then invoking

pragmatic criticisms, of the form that in the given framework most interesting-

sounding claims turn out to be analytic, or that certain vocabulary doesn't add to

explanatory scope, or whatnot. The point is that the criticisms can be made case-

by-case, without general criteria. (1992, 74-75)

I will recast the approach described in this passage as a method of explicating the

informal concept of empirical significance. This recasting may help to clarify the

proposal and highlight its advantages.

As I discussed above, the various criteria of empirical significance are attempts to

explicate the informal explicandum of an expression's making a difference for

observational prediction. And they are attempts to do so in a highly general way; extant

significance criteria are supposed to apply to any language or theory whatsoever. I want

to consider this aspect of significance criteria through the lens of Carnap's notion of call

general and *special* concepts. A general concept is defined so as to be "applicable to any

language whatsoever" (Carnap 1937, 167), or to "all languages of a certain kind" (Carnap

1937, 153), e.g., first-order languages, languages with classical connectives, etc. Though

Carnap does not explicitly elucidate the notion of application in play here, he clearly

intends a concept C-in-L to "apply" to a language L* only if L ranges over a class of

languages to which L* belongs. When I say that a sentence is logically true-in-L, in the

sense of being true-in-L on all re-interpretations of L's descriptive vocabulary, I mean for

L to range over all languages. By contrast, the 'English' in the concept of truth-in-English

ranges only over the various dialects of English.

In place of Carnap's sharp distinction between general and special, I will use

comparative relations of *more or less general* and *more or less special than*; thus the

concept of logical truth-in-L is more general and less special than the concept of truth-in-

English.[12] For present purposes, these comparative concepts may be loosely understood.

In his discussions of general vs. special concepts, Carnap focuses on restrictions on the scope of applicability having to do with languages' logical notions e.g., restrictions to intuitionistic or classical languages. But given that a language for science, in his sense, is also partially constituted by its theoretical and observational postulates, we can also restrict concepts to languages with or without certain such postulates. The relatively special explicata that I will propose for empirical significance will involve restrictions of this kind.

So far, all attempts to explicate empirical significance have been highly general: they are supposed to apply to any axiomatization of any body of scientific theory whatsoever. I believe that Goldfarb and Ricketts's proposal of "case-by-case", pragmatic selection of the scientifically admissible languages is best understood as the proposal to give relatively special explications of the informal notion of empirical significance.

What would a relatively special explication of empirical significance look like? There are various possibilities. For certain languages one of the extant significance criteria might work. For example, Creath's or Schurz's criteria could work specifically for many languages that have no shortcut terms. For others, such as L_1, discussed above, we could resort to enumeration: we could ask for each candidate set of descriptive terms of L_1 whether it accords with the informal notion of making a difference for the prediction of an observable event and certify its members empirically significant if and only if it does so.

---

[12] I do so to address an argument, put to me by an anonymous reviewer, that contends that there are no meta-linguistic concepts that apply to all languages; rather, such concepts must be relativized to a logical type.

The obvious advantage of this case-by-case approach is that relatively special explicata will be less susceptible than general explicata to counter-example. Since we can see, through consideration of the informal concept of that which makes a difference for the prediction of an observable event, that Q_1 falls under the concept, we include it in L_1's theoretical vocabulary. Moreover, if we overlook a deserving term or mistakenly include an undeserving one in our enumeration, adding or removing it will not be an "unintuitive and ineffective [patch]" (1988, 4). There is nothing untoward about improving on a definition by enumeration by adding to or removing from the list.

Is the case-by-case approach to explication, so understood, methodologically sound? The desiderata governing explication do not rule it out. Explication is, fundamentally, an attempt to state a precise definition for, and thereby improve upon, an informal notion. Thus, one desideratum of explication is *exactness* of the precisely defined *explicatum*. A second is *similarity* of the explicatum to the explicandum. But unlike, say, dictionary definition, explication does not require complete or even significant similarity to the explicandum; large deviations are permitted. But such deviations should improve upon the concept by making it more *fruitful* or *simpler*, which are the third and fourth desiderata for explication (Carnap 1953b, section 3). Relatively special explications can do well according to all four desiderata.

Carnap explicitly argues that one may legitimately give special explications of a notion without giving it a fully general explication. In response to Quine's objection that 'analytic' has not been defined for an arbitrary language, Carnap writes:

> [i]n case Quine's remarks are meant as a demand to be given one definition
> applicable to all systems [i.e. languages], then such a demand is manifestly

unreasonable; it is certainly neither fulfilled nor fulfillable for semantic or syntactic concepts, as Quine knows. (Carnap 1991, 430)

Carnap goes on to state that it is therefore unlikely that Quine meant to make such a demand of explications of the informal notion of analyticity. And Carnap is right that Quine did not, in general, require that concepts possess fully general characterizations; Quine argues that various concepts that are indispensible to linguistics and logic—those of grammatical category, construction, and morpheme—resist fully general definitions, but that "anyway there is no need to force transcendence" (Quine 1986, 20) (transcendence being Quine's analogue for generality), as the relatively special definitions of the notions in question suffice for the relevant research programs.

Why, then, have the proponents of significance criteria devoted so much effort to constructing fully general criteria? Generality does seem *prima facie* desirable. A more general explication can provide a more complete or more exact understanding: with its help, one can avoid relying on the inexact explicandum to determine whether a given special explicatum is adequate. This might be reason enough to continue to pursue the project of explicating generally empirical significance. But it is no reason to think that it is full generality or bust for empirical significance.

Relatively special explications of empirical significance would serve the same purpose as more general ones. As the passage quoted earlier from Goldfarb and Ricketts indicates, it would ground the same criticisms of metaphysics and pseudo-science. The criticisms made on a case-by-case basis

have no different status from those that would be generated by a general criterion of testability; for even a general criterion could have only pragmatic sanction.

That is, a general criterion would simply be a proposal for rating linguistic frameworks, whose merits could be urged only on pragmatic grounds. It would avoid the drudgery of case-by-case criticism, but would not escape its noncognitive nature. (Goldfarb and Ricketts 1992, 75)

In section 7, I explained how Carnap's pragmatism motivates the use of significance criteria. We can make the same point in Goldfarb and Ricketts's terms: Carnap's pragmatism provides one interpretation of the "pragmatic grounds" of either a "general criterion" or the "case-by-case" criticism mentioned in the passage above. Ricketts might disagree with this proposal *qua* Carnap interpretation. Elsewhere, he maintains that any systematic statement of the standards for scientific explication (or in Ricketts's terminology, "clarification") would amount to the kind of philosophy that Carnap "eschews in [*The Logical Syntax of Language*]" (Ricketts 1994, 196-197). Against this interpretation, I would offer the textual evidence presented in section 7. But in any case, my present concern is not exegetical. My point is that what I have called 'Carnap's pragmatism' provides a serviceable basis for the pragmatic case-by-case criticism of candidate languages for science.

9. Conclusion

Optimists about a precise demarcation, at the level of individual sentences, between the empirically meaningful and the empirically meaningless, have so far assumed that the defining characteristic of the empirical is some kind of predictive power, and that this concept of predictive power must be explicated through a highly general definition. I have argued against both of these assumptions. Against the predictive power conception,

and in particular, Creath's (1976) and Schurz's (1991; 2014b) significance criteria, I

raised the problem of shortcut terms, i.e., terms that add no new observational predictions

but that make it easier to determine latent predictions. Against the assumption that the

demarcation between the empirical and non-empirical needs to be general, I argued that

Goldfarb and Ricketts's (1992) case-by-case procedure for identifying the scientifically

admissible terms may be understood as a series of relatively special explications, without

any fully general explication, of the informal notion of empirical significance. I claimed

that there is nothing methodologically unsound in such a procedure. Furthermore, from a

Carnapian perspective at least, a demarcation arrived at through this case-by-case process

would have the same force as one that relies on a fully general significance criterion; both

are ultimately motivated by the same pragmatic considerations.

Bibliography

Achinstein, P. (1968). *Concepts of Science*. Baltimore, MD: Johns Hopkins Press.

Ayer, A. J. (1946). *Language, Truth, and Logic* (2nd ed.). London: Gollancz.

Baker, A. (2013, Fall). *Simplicity*. (E. N. Zalta, Editor) Retrieved Jan. 13, 2014, from The Stanford Encyclopedia of Philosophy: http://plato.stanford.edu/archives/fall2013/entries/simplicity/

Barrios, E. (forthcoming). Simple is not Easy. *Synthese* .

Carnap, R. (1936-7). Testability and Meaning. *Philosophy of Science , 3-4,* 410-71 and 1-40.

Carnap, R. (1937). *The Logical Syntax of Language*. (A. Smeaton, Trans.) New York: Open Court.

Carnap, R. (1939). *Foundations of Logic and Mathematics* (Vol. I). Chicago: University of Chicago Press.

Carnap, R. (1943). *Formalization of Logic*. Cambridge, MA: Harvard University Press.

Carnap, R. (1953a). Formal and Factual Science. In H. Feigl, & M. Brodbeck, *Readings in the Philosophy of Science* (H. Feigl, & M. Brodbeck, Trans., pp. 123-8). New York: Appleton-Century-Crofts.

Carnap, R. (1953b). *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Carnap, R. (1956a). The Methodological Character of Theoretical Concepts. (H. Feigl, & M. Scriven, Eds.) *Minnesota Studies in the Philosophy of Science, The Foundations of Science and the Concepts of Psychology and Psychoanalysis , I*, 38-76.

Carnap, R. (1956b). *Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago: University of Chicago Press.

Carnap, R. (1966). *Philosophical Foundations of Physics: An Introduction to the Philosophy of Science*. New York: Dover Publications.

Carnap, R. (1991). Quine on Analyticity. In R. Creath, *Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Work* (pp. 427-432). Berkeley, CA: The University of California Press.

Church, A. (1949). Review of the Second Edition of Language, Truth and Logic. *Journal of Symbolic Logic , 14*, 52-53.

Creath, R. (1976). On Kaplan on Carnap on Significance. *Philosophical Studies , 30* (6), 393-400.

e Silva, O. T., Herzog, S., & Pardi, S. (2013). Empirical Verification of the Even Goldbach Conjecture and Computation of Prime Gaps Up to $4{*}{*}{*}10^{18}$. *Mathematics of Computation , 83* (288), 2033-2060.

Glymour, C. (1980). *Theory and Evidence*. Princeton, NJ: Princeton University Press.

Goldfarb, W., & Ricketts, T. (1992). Carnap and the Philosophy of Mathematics. In D. Bell, & W. Vossenkuhl, *Science and Subjectivity: The Vienna Circle and Twentieth Century Philosophy* (pp. 61-78). Berlin: Akademie-Verlag.

Hempel, C. (1951). The Concept of Cognitive Significance. *Proceedings of the American Academy of Arts and Sciences , 80*, 61-77.

Justus, J. (2014). Carnap's Forgotten Criterion of Empirical Significance. *Mind , 123* (490), 415-36.

Kaplan, D. (1975). Significance and Analyticity: A comment on Some Recent Proposals of Carnap. In J. Hintikka (Ed.), *Rudolf Carnap, Logical Empiricist* (pp. 87-94). Dordrecht, Holland: D. Reidl.

Lewis, D. (1988). Statements Partly About Observation. *Philosophical Papers , XVII* (1), 31.

Lindsay, R. (1937). The Meaning of Simplicity in Physics. *Philosophy of Science , 4*, 151-167.

Ludlow, P. (2011). *The Philosophy of Generative Linguistics*. New York: Oxford University Press.

Peirce, C. S. (1935). *Collected Papers of Charles Sanders Peirce, Volume VI: Pragmatism and Pragmaticism and Scientific Metaphysics*. In C. Hartshorne, & P. Weiss (Eds.), Cambridge, MA: Belknap.

Psillos, S. (2008). Carnap and Incommensurability. *Philosophical Inquiry , XXX* (1-2), 135-156.

Quine, W. (1986). *Philosophy of Logic* (Second ed.). Cambridge, Mass: Harvard University Press.

Ricketts, T. (1994). Carnap's Principle of Tolerance, Empiricism, and Conventionalism. In B. Hale, & P. Clark, *Reading Putnam* (pp. 176-200). Oxford: Blackwell.

Schurz, G. (1991). Relevant Deduction: From Solving Paradoxes Towards a General Theory. *Erkenntnis , 35* (1/3), 391-437.

Schurz, G. (2014a). Criteria of Theoreticity: Bridging Statement and Non-Statement View. *Erkenntnis , 79*, 1521-1545.

Schurz, G. (2014b). *Philosophy of Science: A Unified Approach*. New York: Routledge.

Schurz, G., & Lambert, K. (1994). Outline of a Theory of Scientific Understanding. *Synthese , 101* (1), 65-120.

Surovell, J. (2017, June). Language, Ontology, and the Carnap-Quine Debate. *Philosophia, 45*(2), 811-833.

Wright, C. (1986). Scientific Realism, Observation and the Verification Principle. In G. Macdonald, & C. Wright, *Fact, Science, and Morality*. Oxford: Basil Blackwell.

Yi, B.-U. (2001). Compact Entailment and Wright's Verification Principle. *Mind , 110* (438), 413-421.