Carnap on Empirical Significance

Sebastian Lutz*

Preprint: 2014–04–30

Abstract

Carnap's search for a criterion of empirical significance is usually considered a failure. I argue that the results from two out of his three different approaches are at the very least problematic, but that one approach led to success. Carnap's criterion of translatability into logical syntax is too vague to allow definite results. His criteria for terms—introducibility by reduction sentences and his criterion from "The Methodological Character of Theoretical Concepts"—are almost trivial and have no clear relation to the empirical significance of sentences. However, his criteria for sentences—translatability, verifiability, falsifiability, confirmability—are usable, and under assumption of the Carnap sentence, verifiability, falsifiability, and translatability become equivalent. The price for the Carnap sentence approach is that metaphysics cannot always be shown to be non-significant.

Keywords: empirical significance; cognitive significance; meaningfulness; Carnap; logical empiricism; Ramsey sentence; Carnap sentence; verifiability; falsifiability; testability; translatability

Contents

*Munich Center for Mathematical Philosophy, Ludwig-Maximilians-Universität München. sebastian.lutz@gmx.net. A previous version was presented in 2013 at the workshop *Carnap on Logic* at the Ludwig-Maximilians-Universität München, some aspects were presented in 2013 at the workshop *Formal Epistemology and the Legacy of Logical Empiricism* at the University of Texas at Austin and in 2012 at the Groningen/Munich Summer School *Formal Methods in Philosophy*. I thank the audiences for helpful comments and discussions. Research for this article was supported by the Alexander von Humboldt Foundation.

1 Introduction

The search for a criterion of empirical significance is generally considered an abject failure (see, e. g., Ruja 1961; Soames 2003, ch. 13). However, while it is true that the ratio of successful to unsuccessful criteria is abysmally low, it obviously just needs one successful criterion out of the many for the search to succeed. And I will argue in the following that over the course of his career, Rudolf Carnap not only contributed many of the unsuccessful criteria, but also the successful one.

Carnap was in fact engaged in two logically distinct but historically closely connected searches. On the one hand, he tried developing a criterion of empirical significance for terms. This search failed on a number of levels, leading to criteria that were unmotivated, trivial, or both. On the other hand, he tried developing a criterion for sentences; this search remained well-motivated throughout his career, and the successful criterion is in fact very close to earlier suggestions. That Carnap did not develop his successful criterion earlier has probably two main reasons. For one, the criterion relies on the Carnap sentence, which Carnap discovered late in his career. Furthermore, the criterion requires the explicit conventionalism that Carnap also only developed later. A sort of corollary of the latter requirement is that Carnap's successful criterion cannot be used to criticize all of metaphysics as meaningless, contrary to his initial aim.

Since Carnap changed his terminology over the course of his career, and I am interested in the relations between his different accounts of empirical significance, I will translate Carnap's different terms into my own. For one, I will use the term 'empirical significance' or, when this does not lead to confusion, 'significance', while Carnap instead used 'meaningfulness', 'cognitive significance' and a number of other terms. I will further speak systematically of sentences that are empirically significant as 'statements', not as, for instance, 'propositions'. 'Expressions' may be ill-formed and thus not sentences. A 'pseudo-statement' may be either a non-significant sentence or an ill-formed expression. The use of the term 'sentence' has the advantage that it connects to the technical criteria that Carnap has suggested, which all apply to sentences in some logical language $\mathscr{L}$. It also connects easily to the logical notion of a formula. All of Carnap's technical criteria furthermore assume a distinguished sublanguage of $\mathscr{L}$, which I will call the basic language $\mathscr{B}$. In the texts discussed here, Carnap calls it the autopsychological basis, the physical language, protocol sentences, observation language, and more. I will assume that $\mathscr{B}$ can be identified with a set of sentences, and so I will speak of $\mathscr{B}$-sentences. Sometimes, Carnap identifies $\mathscr{B}$-sentences by their logical structure and the terms that occur in them, in which case I will speak of $\mathscr{B}$-terms.[1] I will call non-$\mathscr{B}$-terms auxiliary or $\mathscr{A}$-terms.

2 Informal Translatability

In his initial, programmatic criticism of metaphysics, Carnap (1931b, 61) states that there are two ways for an expression to lack empirical significance.

> A language consists of a vocabulary and a syntax, i. e. a set of words which have meanings and rules of sentence formation. These rules indicate how sentences may be formed out of the various sorts of words. Accordingly, there are two kinds of pseudo-statements: either they contain a word which is erroneously believed to have meaning, or the constituent words are meaningful, yet are put together in a counter-syntactical way, so that they do not yield a meaningful statement. [M]etaphysics in its entirety consists of such pseudo-statements.

Thus the first kind of pseudo-statements consists of sentences containing non-significant terms, the second kind of ill-formed expressions.

The second kind of pseudo-statement occurs when expressions accord with the rules of historical-grammatical syntax, but violate the rules of logical syntax (Carnap 1931b, 69). Carnap (1931b, 68) writes:

> If grammatical syntax corresponded exactly to logical syntax, pseudo-statements could not arise. If grammatical syntax differentiated not only the wordcategories of nouns, adjectives, verbs, conjunctions etc., but within each of these categories made the further distinctions that are logically indispensable, then no pseudo-statements could be formed.

The examples Carnap discusses in the remainder of his article are of two sorts. There is Heidegger's well-known 'The Nothing itself nothings', which Carnap considers a pseudo-statement because 'nothing' marks the negation of an existentially quantified sentence and cannot be identified with a constant symbol (§5). The other sort results from "type confusion" (75), where the types are those of Russell's type theory.[2] In a type confusion, a word of one type is used at a position in a formula that can only be used by a word of another type, resulting in

1. More or less in keeping with Carnap's and common terminology, I will use 'term' synonymously with 'non-logical symbol'.

2. The translation of the *Aufbau* (Carnap 1928a) uses the more literal translations 'object sphere' for 'type' (Carnap 1967a, §29), and 'confusion of spheres' for 'type confusion' (§30).

an ill-formed expression. Pseudo-statements can be hard to identify if they have the same historical-grammatical form as significant sentences.

To argue against Heidegger's 'The Nothing itself nothings', Carnap compares three kinds of expressions: Expressions in logical syntax (IIIA), expressions in historical-grammatical syntax that can be translated into statements in logical syntax (IIA), and expressions that cannot be translated into logical syntax (IIB).

> Sentence form IIA [...] does not, indeed, satisfy the requirements to be imposed on a logically correct language. But it is nevertheless meaningful, because it is translatable into correct language. This is shown by sentence IIIA, which has the same meaning as IIA. [T]he meaningless sentence forms IIB, which are taken from [Heidegger's text ...] cannot even be constructed in the correct language.

Thus Carnap considers it a necessary condition for significance that a sentence can be translated into a sentence in logical syntax. If all terms that occur in the translation are also significant (as is assumed for sentence IIIA), the condition is also sufficient.[3]

It is a major drawback of this translatability condition of significance that it is informal: There is no formal way to decide whether a statement in logical syntax is a correct translation of some expression in historical-grammatical syntax. Thus Carnap in effect has to engage in the interpretation of Heidegger's text:

> [W]e might be led to conjecture that perhaps the word 'nothing' has in Heidegger's treatise a meaning entirely different from the customary one. [...] But the first sentence of the quotation [of Heidegger's][4] proves that this interpretation is not possible. The combination of 'only' and 'nothing else' shows unmistakably that the word 'nothing' here has the usual meaning of a logical particle that serves for the formulation of a negative existential statement.

Thus, in the end, critics of allegedly non-significant expressions must show that they did not simply fail to grasp the meaning of perfectly fine statements. If the language and the assumptions in the context of an expression are not fixed, they hence have to guess the intention of the one who proposed the expression to determine whether it is a significant sentence. But this problem can also be posed to the proponents of such expressions. Speaking about a specific kind of question in philosophy that may be non-significant, Carnap (1935, 79) states: "I do not know how such questions could be translated into [any] unambiguous and clear mode;

3. Note that the translation into logical syntax is different from the translation into the formal mode of speech (Carnap 1934, §74): Logical syntax is used to phrase object-sentences, which express claims of the empirical sciences, while the formal mode of speech is used to phrase syntactical sentences, which express philosophical claims.

4. "What is to be investigated is being only and—*nothing* else; being alone and further—*nothing*; solely being, and beyond being—*nothing*. *What about this Nothing?*" (Carnap 1931b, 69; cf. Heidegger 1931, 9–10).

and I doubt whether the philosophers themselves who are dealing with them are able to give us any such precise formulation. Therefore it seems to me that these questions are metaphysical pseudo-questions." Carnap in effect turns around the burden of proof: Rather than showing an expression non-significant, he demands that its significance be shown by translation into logical syntax. This strategy would later be used by Flew (1950, 258) in an influential argument against the significance of theological expressions. As Flew puts it, someone may utter 'God loves us as a father loves his children' with the standard meaning of 'God', 'love', and so on, and thus with straightforward implications for the world (say, the absence of undeserved suffering). But in light of counterexamples, he may qualify the hypothesis more and more and finally "may dissipate his assertion completely without noticing that he has done so. A fine brash hypothesis may thus be killed by inches, the death by a thousand qualifications." If it is difficult for the proponent of a hypothesis to realize its non-significance, it is much more difficult for the critic to show its non-significance. Flew (1950, 259) responds to this problem like Carnap: "I therefore put to the symposiasts the simple central questions, 'What would have to occur or to have occurred to constitute for you a disproof of the love of, or of the existence of, God?".

Flew combines Carnap's informal translatability condition with the demand that every significant sentence must be falsifiable. In "Testability and Meaning", Carnap (1937, 3) distinguishes clearly between the two aspects: The question about the criterion of empirical significance "refers to a given language-system L and concerns an expression E of L [. . .]. The question is, whether E is meaningful or not. This question can be divided into two parts: a) 'Is E a sentence of L'?, and b) 'If so, does E fulfill the empiricist criterion of meaning'?" Flew's question thus assumes that the empiricist criterion of meaning is that of falsifiability. However, the question whether E is meaningful (i. e., significant) is but *one kind* of question about the criterion. As Carnap (1937, 4) puts it, a question of the second kind

> concerns a language-system L which is being proposed for construction. In this case the rules of L are not given, and the problem is how to choose them. We may construct L in whatever way we wish. There is no question of right or wrong, but only a practical question of convenience or inconvenience of a system form, i. e. of its suitability for certain purposes.

For instance, the sentence S_1, 'This stone is now thinking about Vienna', would have been declared meaningless because it cannot be translated into logical syntax (presumably because of a type confusion). "But at present I should prefer to construct the scientific language in such a way that it contains a sentence S_2 corresponding to S_1. (Of course I should then take S_2 as false, and hence $\sim S_2$ as true.)" However, with that much leeway in translating sentences, it is not obviously impossible to translate 'The nothing itself nothings' into the logical syntax of *some* language. Thus the informal condition of translatability is rather problematic, and the hope has to rest on criteria of empirical significance that apply to sentences in logical syntax.[5] In the remainder of this essay, I will look at Carnap's suggestions for such criteria.

3 Europe

Since Carnap's criteria of empirical significance are connected to the notion of meaning and the scientific language (Carnap 1935, 32), they are directly related to his positions on the semantics of scientific theories. As far as his explicit proclamations are concerned, this leads to a natural grouping of his positions up to "Testability and Meaning" (Carnap 1936, 1937) and of his later positions. For in his earlier works, Carnap relied on the assumption that it is possible to develop all the terms of science starting out from basic terms or sentences. His later works explicitly assume that this is not always expedient or even possible.[6]

With respect to the relation between basic and auxiliary terms, Carnap (1963a, §9) describes in his "Intellectual Autobiography" the development of logical empiricism as a gradual liberalization. Initially, every kind of knowledge "was supposed to be firmly supported" by the experiences as described by Wittgenstein's principle of verifiability, "which says that it is in principle possible to obtain either a definite verification or a definite refutation for any meaningful sentence" (57). But even the early Carnap was more tolerant than that, for some of his criteria also allowed (non-definite) confirmation and disconfirmation. One central question, however, will concern the relations between the many different conceptions.

3.1 Criteria for Sentences

In the *Aufbau*, Carnap (1928a, cf. §§38–39) describes how to translate every scientific sentence into a basic ("autopsychological") sentence.[7] Note for the following that translatability typically requires some background assumptions which, for convenience of notation, I will treat as one long conjunction ϑ. In Carnap's elucidations, the role of background assumptions is typically included in the rules of inference: "We say of a sentence P that it is translatable (more precisely, that it is reciprocally translatable) into a sentence Q if there are rules, independent of space and time, in accordance with which Q may be deduced from P and P from Q" (Carnap 1932/1933, 166). In the *Aufbau*, the background assumptions are the

5. Marhenke (1949/1950) comes to the opposite conclusion because he ignores the problems with the informal translatability condition and, contrary to Carnap, assumes that inference rules are only applicable to significant sentences.

6. Even in Carnap's early works, however, his later position occurs as an undercurrent (Lutz 2012, §3.6.1).

7. He also claims that it is in general possible to translate all scientific sentences into a variety of different basic languages, for example the physical language (Carnap 1928a, §54–60).

definitions of the constitutional system, but in general, any set of sentences is allowed:

Definition 1. Sentence σ is (non-trivially) translatable into language $\mathscr{B}$ by sentence ϑ if and only if there is a sentence β in $\mathscr{B}$ such that $\vartheta \vdash \sigma \leftrightarrow \beta$ (and $\vartheta \nvdash \beta$, $\vartheta \nvdash \neg\beta$).

Carnap does not explicitly claim that only translatable sentences are significant. Carnap (1931a, 453, all translations are mine) later does argue for physicalism, the thesis that "every scientific sentence can be translated into the physicalistic language"[8]; the physicalistic language is the set of protocol-sentences, which play the role of $\mathscr{B}$-sentences. However, Carnap (1931a, 452) states in the same text that "logical analysis comes to the conclusion [...] that the so-called metaphysical sentences are pseudo-sentences, since they stand in no inferential relation (neither positive nor negative) to the sentences of the protocol-language."[9] Introducing some more a-historical terminology for the sake of precision, I will say that according to Carnap, metaphysical sentences are neither verifiable nor falsifiable:[10]

Definition 2. Sentence σ is (non-trivially) verifiable in language $\mathscr{B}$ relative to ϑ if and only if there is a sentence β in $\mathscr{B}$ such that $\vartheta \wedge \beta \nvdash \bot$ and $\vartheta \wedge \beta \vdash \sigma$ (and $\vartheta \nvdash \sigma$).

Definition 3. Sentence σ is (non-trivially) falsifiable in language $\mathscr{B}$ relative to set ϑ of sentences if and only if there is a sentence β in $\mathscr{B}$ such that $\vartheta \wedge \beta \nvdash \bot$ and $\vartheta \wedge \beta \vdash \neg\sigma$ (and $\vartheta \nvdash \neg\sigma$).

The demand that β be compatible with ϑ stems from Carnap's position that the basic sentences must be possible according to the laws of nature; this stance is in opposition to Schlick, who only demands that basic sentences be logically possible (Carnap 1936, 423; cf. Friedl and Rutte 2008).

In an article arguing that every sentence of psychology can be translated into physical language, Carnap (1932/1933, 166) states that a person "*tests* (verifies) a system-sentence by deducing from it sentences of his own protocol language, and comparing these sentences with those of his actual protocol. The possibility of such a deduction of protocol sentences constitutes the content of a sentence. If a sentence permits no such deductions, it has no content, and is meaningless." This suggests

Definition 4. Sentence σ has (non-trivial) content relative to sentence ϑ if and only if there is a sentence β in $\mathscr{B}$ such that $\vartheta \nvdash \beta$ and $\vartheta \wedge \sigma \vdash \beta$ (and $\vartheta \nvdash \neg\sigma$).

8. "Unsere Überlegungen [...] führen somit zu dem Ergebnis, daß jeder wissenschaftliche Satz in die physikalische Sprache übersetzbar ist."

9. "Aber die logische Analyse kommt zu dem Ergebnis [...], daß die sog. metaphysischen Sätze Scheinsätze sind, da sie in keinem Ableitungsverhältnis (weder einem positiven noch einem negativen) zu den Sätzen der Protokollsprache stehen."

10. This terminology follows that of Hempel (1950, 45–48).

However, this is but one of the conditions under which Carnap (1931a, 452) claimed a sentence to be meaningful just a year earlier:

Claim 1. *If $\mathscr{B}$ contains with every sentence also its negation, then a sentence is (non-trivially) falsifiable relative to ϑ if and only if it has (non-trivial) content relative to ϑ.*

Proof. $\vartheta \wedge \beta \vDash \neg\sigma$ if and only if $\vartheta \wedge \sigma \vDash \neg\beta$. $\vartheta \wedge \beta \nvDash \bot$ if and only if $\vartheta \nvDash \neg\beta$. □

And in another article of the same year, Carnap (1931b, 62) discusses the significance of a word (like 'stone') using elementary sentences S (like 'This diamond is a stone'):

> [F]or an elementary sentence S containing the word an answer must be given to the following question, which can be formulated in various ways:
>
> 1. What sentences is S *deducible* from, and what sentences are deducible from S?
> 2. Under what conditions is S supposed to be *true*, and under what conditions false?
> 3. How is S to be *verified*?
> 4. What is the *meaning* of S?
>
> (1) is the correct formulation; formulation (2) accords with the phraseology of logic, (3) with the phraseology of the theory of knowledge, (4) with that of philosophy (phenomenology).

In (1), the sentences entailing S and entailed by S are subsequently restricted to $\mathscr{B}$-sentences (protocol sentences). Call the weakest $\mathscr{B}$-sentence that entails S the *ground* $G(S)$ of S, and the strongest $\mathscr{B}$-sentence that is entailed by S the *content* $C(S)$ of S. Then (1) identifies the meaning of S with the ground and the content of S, if they can be expressed in a single sentence: The ground of S is equivalent to the disjunction of all $\mathscr{B}$-sentences entailing S, and the content is equivalent to the set of all $\mathscr{B}$-sentences entailed by S.[11] The relation between (1) and (4) then suggest that the ground and the content together determine a sentence's meaning, so that somehow, any sentence σ can be translated into its ground $G(\sigma)$ and its content $C(\sigma)$. However, by definition,

$$G(\sigma) \wedge \vartheta \vdash \sigma \wedge \vartheta \vdash C(\sigma) \wedge \vartheta \ , \tag{1}$$

where I have already taken the background assumption into account. By assumption, $G(\sigma)$ and $C(\sigma)$ are $\mathscr{B}$-sentences, and so σ is translatable into $\mathscr{B}$ if and only if $G(\sigma) \wedge \vartheta \dashv\vdash C(\sigma) \wedge \vartheta$. Obviously, this is not always fulfilled.

One might also also expect that a sentence that is verifiable or has content is translatable. But his is not the case either:

11. For the ground of S, the equivalence can be expressed only in languages that allow disjunctions with infinitely many disjuncts.

Claim 2. *There are sentences that are non-trivially verifiable in some $\mathscr{B}$ and have non-trivial content relative to some ϑ without being translatable into $\mathscr{B}$ by ϑ.*

Proof. Choose $\vartheta = \varnothing$ and some μ that is not verifiable and has not content (say, a logically contingent sentence containing only terms not occurring in $\mathscr{B}$), and choose two sentences β and β' from $\mathscr{B}$ whose disjunction is also in $\mathscr{B}$. Then $(\mu \land \beta') \lor \beta$ can be derived from β and entails $\beta \lor \beta'$, but it is not equivalent to a $\mathscr{B}$-sentence. □

Thus some sentences are even verifiable *and* have content without being translatable into $\mathscr{B}$.

Thus there are three different kinds of relations (falsifiability, verifiability or falsifiability, and translatability), all of which seem to determine *on their own* whether a sentence is significant. There is also an additional wrinkle whose implications I will not discuss in the following: In early discussions, Carnap assumes that the protocol sentence are in one language, the remaining ("system-")sentences in another. According to Carnap (1932), it was Neurath (1932) who first suggested treating protocol sentences and system-sentences as from the same language.[12]

In addition, Carnap (1928b, 327–328) already allows early on not only verification and falsification, but also confirmation and disconfirmation: "If a statement p expresses the content of an experience E, and if the statement q is either the same as p or can be derived from p and prior experiences, either through deductive or inductive arguments, then we say that q is 'supported by' the experience E." By allowing inference of p or $\neg p$ by deductive arguments, Carnap stipulates a sentence to be significant if it is verifiable or falsifiable. But beyond that, he also allows inductive inferences. Since he does not spell out what kind of inductive inferences he has in mind, it is hard to say how much of a deviation from verifiability and falsifiability this addition is, but for the sequel, it will be informative to look at probabilistic inference. During Carnap's early years, its use was especially championed by Reichenbach (Carnap 1963b, 58), but as discussed below, Carnap would later also suggest this approach.

The standard definition of probabilistic confirmation is the following (cf. Howson and Urbach 1993, 117):

Definition 5. Assuming all occurring probabilities are well-defined, sentence σ is probabilistically confirmable in $\mathscr{B}$ relative to sentence ϑ if and only if there is a sentence β in $\mathscr{B}$ such that

$$\Pr(\sigma \mid \beta \land \vartheta) > \Pr(\sigma \mid \vartheta) .$$

12. A further, odd wrinkle that I will not pursue is that Carnap assumed in the *Aufbau* that all scientific terms are definable in basic terms, and thus the basic sentences are a subset of the scientific language. Thus it is not clear in what way Neurath's suggestion was novel.

A sentence σ is probabilistically disconfirmable in $\mathscr{B}$ relative to sentences ϑ if and only if there is a sentence β in $\mathscr{B}$ such that

$$\Pr(\sigma \mid \beta \land \vartheta) < \Pr(\sigma \mid \vartheta) .$$

Note that this definition is not one of total confirmation, since the probability of σ might simply be raised minimally from a very low value to a value almost as low. Analogously, it is not a definition of total disconfirmation.

Unlike verifiability and falsifiability, which do not entail each other, confirmability and disconfirmability are equivalent (see appendix):

Claim 3. *If all occurring probabilities are defined and $\mathscr{B}$ contains with every sentence also its negation, then σ is disconfirmable if and only if σ is confirmable.*

As is often discussed (cf. Howson and Urbach 1993, 119–20), if inductive inferences are treated as probabilistic, falsifiability entails confirmability in all interesting cases:

Corollary 4. *If all occurring probabilities are defined and $\{\Pr(\neg\beta \mid \vartheta), \Pr(\sigma \mid \vartheta)\} \nsubseteq \{0, 1\}$, $\mathscr{B}$ contains with every sentence also its negation, and σ is falsifiable relative to ϑ, then ϑ is confirmable.*

Proof. If $\beta \land \vartheta \vdash \neg\sigma$, then $\Pr(\sigma \mid \beta \land \vartheta) = 0 < \Pr(\sigma \mid \vartheta)$ so that σ is disconfirmable. By claim 3, it is confirmable. □

Informally, a sentence σ is confirmed when a $\mathscr{B}$-sentence that would have falsified σ turns out false. Note that this result and claim 1 have the immediate

Corollary 5. *If all occurring probabilities are defined and $\{\Pr(\neg\beta \mid \vartheta), \Pr(\sigma \mid \vartheta)\} \nsubseteq \{0, 1\}$, $\mathscr{B}$ contains with every sentence also its negation, and σ has content relative to ϑ, then ϑ is confirmable.*

A less often mentioned consequence of probabilistic inferences is that verifiability entails disconfirmability:

Corollary 6. *If all occurring probabilities are defined and $\{\Pr(\neg\beta \mid \vartheta), \Pr(\sigma \mid \vartheta)\} \nsubseteq \{0, 1\}$, $\mathscr{B}$ contains with every sentence also its negation, and σ is verifiable relative to ϑ, then σ is disconfirmable.*

Proof. If $\beta \land \vartheta \vdash \sigma$, then $\Pr(\sigma \mid \beta \land \vartheta) = 1 > \Pr(\sigma \mid \vartheta)$ so that σ is confirmable. By claim 3, it is disconfirmable. □

Informally, a sentence σ is disconfirmed when a $\mathscr{B}$-sentence that would have verified σ turns out false. Together, claim 3 and its corollaries 4 and 6 show that, with the right choice of inductive inference, speaking of confirmability already includes disconfirmability, verifiability, and falsifiability.

In "Testability and Meaning", Carnap (1936, 420) speaks of confirmability,[13] and again claims translatability, this time of an inductive kind.

> Obviously we must understand a sentence, i. e. we must know its meaning, before we can try to find out whether it is true or not. But, from the point of view of empiricism, [if] we knew what it would be for a given sentence to be found true then we would know what its meaning is. And if for two sentences the conditions under which we would have to take them as true are the same, then they have the same meaning. Thus the meaning of a sentence is in a certain sense identical with the way we determine its truth or falsehood; and a sentence has meaning only if such a determination is possible.

Thus it seems that having meaning is identical to being confirmable or disconfirmable (and thus confirmable *and* disconfirmable) and also identical to being translatable, albeit by fiat: The meaning of a sentence is *stipulated* to be given by the set of sentences that confirms it and the set of sentences that disconfirms it.

I now want to show that the technical aspect of Carnap's account in "Testability and Meaning" does not illuminate this relationship. Carnap (1936, 435) calls the confirmation of a sentence S "directly reducible to a class C of sentences" if "S is a consequence of a finite subclass of C" (complete reducibility of confirmation) or "if the confirmation of S is not completely reducible to that of C but if there is an infinite subclass C' of C such that the sentences of C' are mutually independent and are consequences of S" (direct incomplete reducibility of confirmation). This definition is the first in a long chain that eventually leads to the requirement of confirmability, which "suffices as a formulation of the principle of empiricism" (Carnap 1937, 35).[14] Carnap's path to the principle of empiricism is somewhat circuitous, but significantly simplified when taking into account that it becomes trivial with the next link: Carnap (1936, 435) calls the confirmation of S "*reducible* to that of [a class of sentences] C, if there is a finite series of classes $C_1, C_2, \dots, C_n$ such that the relation of directly reducible confirmation subsists 1) between S and C_1, 2) between every sentence of C_i and C_{i+1} ($i = 1$ to $n - 1$), and 3) between every sentence of C_n and C." And this leads to

Claim 7. *If the class C of sentences allows the direct incomplete reducibility of at least one sentence γ, then the confirmation of every sentence σ is reducible to that of C.*

Proof. For any sentence σ, if γ is directly incompletely reducible to that of C, so is $\gamma \wedge \sigma$, which can therefore be in C_1. Then the confirmation of σ can be

13. He adds: "If by verification is meant a definitive and final establishment of truth, then no (synthetic) sentence is ever verifiable, as we shall see. We can only confirm a sentence more and more." However, Carnap's argument for his claim is essentially that *universally quantified* sentences cannot be verified (Carnap 1936, §6). His argument does not hold in general.

14. Note that the incomplete reducibility of confirmation corresponds to the relation between confirmability and having content (claim 5).

completely reduced to that of $C_1 := \{\gamma \wedge \sigma\}$ because $\{\gamma \wedge \sigma\} \vDash \sigma$ and $\{\gamma \wedge \sigma\}$ is a finite subset of itself. Thus the confirmation of σ is directly reducible to that of C_1, whose confirmation is directly reducible to that of C, and therefore the confirmation of σ is reducible to that of C. □

If a language contains infinitely many constants $\{c_i \mid i \in I\}$ for points in space-time,[15] the sentence 'It will always be everywhere cold' is an incompletely directly reducible sentence γ, since the temperature at each point in space-time is logically independent from the temperature at any other and thus γ entails the infinite set of logically independent sentences $\Omega^* := \{\ulcorner\text{It is cold at } c_i\urcorner \mid i \in I\}$.

Since the reducibility of confirmation to a class of sentences is trivial, all other definitions that build on it collapse, too: The confirmation of a sentence S is reducible to that of a class C of predicates if the confirmation of S "is reducible [. . .] to a not contravalid sub-class of the class which contains the full sentences of the predicates of C and the negations of these sentences" (Carnap 1936, 435–436); call such a sub-class a *confirmation class*. Full sentences are atomic sentences, and a contravalid sentence is incompatible with the laws of nature (432–434). Because of claim 7, if some confirmation class Ω allows the direct incomplete reducibility of at least one sentence γ, the confirmation of any sentence σ is reducible to Ω. (In the above example, Ω^* is a confirmation class for γ if $\{c_i \mid i \in I\} \cup \{\lambda x(\text{It is cold at } x)\} \subseteq C$.) Thus the confirmation of any sentence σ is reducible to that of C. If now C is contains only observable predicates ($\mathscr{B}$-predicates), σ is confirmable, because a "sentence S is called *confirmable* [. . .] if the confirmation of S is reducible [. . .] to that of a class of observable predicates" (456). Since nothing was assumed about σ, the principle of empiricism is then met by any sentence whatever.

As Wagner (2014, 40–41) has shown in response to the above, Carnap (1950c, 40A) changes the definition of direct incomplete reducibility in a reprint of "Testability and Meaning" in a way that blocks the above trivialization proof: Now, "the confirmation of a *a non-contravalid sentence* S is directly incompletely reducible to that of C, if the confirmation of S is not completely reducible to that of C but if there is an infinite subclass C' of C such that the sentences of C' are mutually independent and are consequences of S *by substitution alone*." This restricts the entailment needed for direct incomplete reductions to universal instantiations, that is, S must be a universally quantified formula $\forall x \varphi(x)$ and specifically cannot be a conjunction as assumed in the proof of claim 7.

It is not known why Carnap made these two changes, but one can make educated guesses: The first addition avoids an obvious trivialization: If S is contravalid, it entails every sentence, and thus specifically those of C'. Thus it is directly incompletely confirmed and, being contravalid, can be used to completely confirm any sentence whatever. The second addition avoids the less obvious trivialization of claim 7 and there is a somewhat speculative reason to think that

15. This is what Carnap (1936, 433–434) seems to assume.

this was exactly Carnap's intention: Five years before the reprint, Hempel (1945, 103–104) had pointed out that the conjunction of three intuitively plausible conditions of adequacy for confirmation is trivial. According to the *entailment condition*, if $\varepsilon \vdash \varrho$, then ε confirms ϱ. Thus, specifically, any sentence γ confirms itself. The *converse consequence condition* demands that if ε confirms ϱ and $\varrho' \vdash \varrho$, then ε also confirms ϱ'. Thus γ confirms $\gamma \wedge \sigma$, where σ is any sentence whatever. According to the *special consequence condition*, if ε confirms ϱ and $\varrho \vdash \varrho'$, then ε confirms ϱ'. Thus γ confirms σ. It is easy to see that direct incomplete reducibility fulfills the converse consequence condition and complete reducibility fulfills the special consequence condition. The proof of claim 7 essentially follows Hempel's trivialization proof, skipping the use of the entailment condition by assuming that there is a directly incompletely confirmed sentence. Since in all likelihood Carnap had analyzed Hempel's conditions of adequacy before preparing "Testability and Meaning" for the reprint,[16] he could easily have seen this connection.

Unfortunately, Carnap's modification does not avoid the trivialization of his criterion. To see why, note first that if S must have the form $\forall x \varphi(x)$, then C' must have the form $\{\varphi(a_i) \mid i \in I\}$, where C has infinite cardinality and $\varphi(a_i) \nvdash \varphi(a_j), \neg\varphi(a_j)$ for any $i, j \in I, i \neq j$. Carnap's intent here seems to be something along the lines of an (infinite) enumerative induction. In other words, he seems to presume that $\ulcorner a_i \neq a_j \urcorner \in C$ if $i \neq j$, that is, φ is to be predicated of infinitely many objects. But as became clear through Goodman's "new riddle of induction" (Goodman 1965, §III.4), it is always possible to take a formula φ and craft a new one that predicates φ of the objects used in the induction, but predicates a completely different formula of all other objects. This insight, which unfortunately came to late for Carnap to have taken it into account for the reprint, can be used to trivialize Carnap's new criterion. The only additional assumption is that it is possible to identify at least one object that is not used in (or can be left out of) the induction or, in Carnap's terms, that is not used for (or can be left out of) the direct incomplete confirmation of a sentence.

Claim 8. *According to Carnap (1950a), if the class C of sentences allows the direct incomplete reducibility of the confirmation of at least one sentence γ (by $\{\varphi(a_i) \mid i \in I\}$) and contains the sentences $\ulcorner a_j \neq b \urcorner, j \in J$ for some $J \subseteq I$ of infinite cardinality, then the confirmation of every sentence σ is reducible to that of C.*

Proof. If the confirmation of γ is reducible to that of C, then there is a set $C' \subset C$ of infinite cardinality of the form $\{\varphi(a_i) \mid i \in I\}$. By assumption, $\{a_j \neq b \mid j \in J\} \subset C$. Thus for each $j \in J$, $\varphi(a_j) \wedge a_j \neq b$ is entailed by a finite subclass of C (namely $\{\varphi(a_j), a_j \neq b\}$, and so is $[\varphi(a_j) \wedge a_j \neq b] \vee [a_j = b \wedge \sigma]$, where σ is any sentence whatever. The confirmation of the latter sentences is thus completely reducible to that of C.

16. Carnap (1950b, §87) discusses the conditions at length.

By construction, the set $C_2 = \{[\varphi(a_j) \wedge a_j \neq b] \vee [a_j = b \wedge \sigma] \,|\, j \in J\}$ has infinite cardinality. Since each of its elements is a universal instantiation of the sentence $\gamma' = \forall x([\varphi(x) \wedge x \neq b] \vee [x = b \wedge \sigma])$, the confirmation of γ' is directly incompletely reducible to the confirmation of C_2. γ' entails σ, and thus confirmation of σ is completely reducible to confirmation of $C_1 = \{\gamma'\}$. Therefore, the confirmation of every sentence σ is reducible to that of C. □

The additional assumption of the proof is fulfilled in the example given above: If a language contains infinitely many constants $\{c_i \,|\, i \in I\}$ for points in space-time, the sentence 'It will always be everywhere cold' is incompletely directly reducible, and one can choose any constant $c_g, g \in I$ to build the sentence 'It is cold at every space-time point different from c_g, and for c_g, σ holds'. And again, since the reducibility of confirmation to a class of sentences is trivial, all other definitions that build on it collapse.

In conclusion, Carnap's technical contributions to the search for a criterion of empirical significance up to this point were not successful. His informal discussion of the relation between confirming sentences and confirmed sentence are fascinating, however, and he also indirectly contributed another informal insight. Or rather, he steered clear of a very unfortunate development in the search for a criterion of significance that started, as far as I can tell, with A. J. Ayer.

After one unsuccessful attempt at defining a criterion of empirical significance (Ayer 1936, 38–39, cf. Lewis 1988a), Ayer (1946, 13) proposes two definitions. The first essentially stipulates that a sentence is directly verifiable if and only if it has content relative to any other observational sentence. In his second definition, Ayer proposes saying that

> a statement is indirectly verifiable if it satisfies the following conditions: first, that in conjunction with certain other premises it entails one or more directly verifiable statements which are not deducible from these other premises alone; and secondly, that these other premises do not include any statement that is not either analytic, or directly verifiable, or capable of being independently established as indirectly verifiable.

In a review, Church (1949) showed that for any sentence, as long as there are three logically independent $\mathscr{B}$-sentences, the sentence or its negation is indirectly verifiable, and thus Ayer's amended criterion is close to trivial as well. It was followed by a slew of further amendments and new trivialization proofs succinctly summarized and extended by Pokriefka (1983), who cuts out the middleman and proves the triviality of his amendment himself (Pokriefka 1984). All of these criteria share one crucial feature with Carnap's criterion from "Testability and Meaning" and later criteria (Wright 1986, 1989) that also turned out to be trivial (Lewis 1988b, §IV, n. 12; Wright 1989, §II; Yi 2001): They are all recursive. Thus there is some reason to think that recursive criteria are at the very least a dangerous direction of the search for a criterion of empirical significance.[17] Another reason to question the search for a recursive criterion is that in such a criterion, the background assumptions can contain any empirically significant sentence, even those that are known to be false.

In contradistinction, Carnap (1935, 11) writes in "Philosophy and Logical Syntax": "A proposition P which is not directly verifiable can only be verified by direct verification of propositions deduced from P together with other already verified propositions." Like Ayer's definition of indirect verifiability, Carnap here essentially defines a sentence as verifiable if and only if it has $\mathscr{B}$-content relative to other sentences. But in contrast to Ayer's criterion, the other sentences in his criterion are not only required to be verifiable, but actually verified. Unlike Ayer, Carnap does not define 'verifiability' recursively, but rather relative to a set of confirmed sentences.

Unfortunately, Carnap's criterion fails for a different reason:

Claim 9. *If there are at least two directly verified sentences β, γ with $\beta \nvdash \gamma$, then every non-tautological sentence P can be verified according to Carnap (1935, 11).*

Proof. $\{(P \rightarrow \gamma) \wedge \beta\} \vdash \beta$ and is thus verified by β. Since $\{P, (P \rightarrow \gamma) \wedge \beta\} \vdash \gamma$ while $(P \rightarrow \gamma) \wedge \beta \nvdash \gamma$, P is indirectly verifiable. □

The problem, I surmise, is that Carnap implicitly confuses absolute and relative confirmation: Being verifiable is expressed by relative confirmability (via claim 5), but the background assumptions used in deriving the content of a sentence should not only have been relatively, but absolutely confirmed.

3.2 Criteria for Terms

Parallel to his criteria of empirical significance for sentences, Carnap also developed criteria for terms. Whenever he discusses these, he tries to make sure that they run in parallel to his criteria for sentences. In the *Aufbau*, for instance, every meaningful sentence is supposed to be translatable into a sentence about experiences, and this means that "the concepts of science are explicitly definable on the basis of observation concepts" (Carnap 1963a, 59). It is thus unsurprising that Carnap also assumes for his criteria for terms that the background assumptions are verified rather than verifiable sentences. For instance, when suggesting that every scientific term can be explicitly defined in $\mathscr{B}$-terms (Carnap 1928a, §38),[18] Carnap (1928a, §67, §122) does not intend these definitions to follow from the meanings of the terms outside of any empirical theory, but rather from the regularities that are described by empirical theories (cf. Carnap 1967a, ix; 1963, 945).

17. This is not true for the strongest criterion, translatability: Obviously, if σ can be translated into γ and γ into β, then σ can be translated into β. Thus a definition of empirical significance as recursive translatability is safe, but also pointless.

18. Carnap (1928a, §38) also discusses the need for "definitions in use". As far as terms (i. e., non-logical symbols) are concerned, these are equivalent to explicit definitions because of the eliminability theorems (cf. Gupta 2009, §2.3).

In other words, he claims that these explicit definitions are entailed by scientific theories.[19]

Definition 6. A relation P is $\mathscr{B}$*-definable* in ϑ if and only if there is a $\mathscr{B}$-formula φ such that

$$\vartheta \vdash \forall x_1 \ldots x_n [P x_1 \ldots x_n \leftrightarrow \varphi(x_1, \ldots, x_n)] \tag{2}$$

$\mathscr{B}$-definability relates to translatability in a very straightforward sense:

Claim 10. *If σ is a sentence of $\mathscr{B}$-terms and $\mathscr{B}$-definable relations, then σ is translatable into $\mathscr{B}$.*

Proof. If P is $\mathscr{K}$-definable in ϑ, then for every $\mathscr{K} \cup \{P\}$-sentence σ there is a $\mathscr{K}$-sentence $\varkappa$ such that $\vartheta \vDash \sigma \leftrightarrow \varkappa$ (Essler 1982, 103). Therefore, if the $\mathscr{A}$-relations in σ are $\{P_1, \ldots, P_{k+1}\}$, σ can be translated into a $\mathscr{B} \cup \{P_1, \ldots, P_k\}$-sentence σ_k, and for $1 \leq l \leq k$, σ_l can be translated into a $\mathscr{B} \cup \{P_1, \ldots, P_{l-1}\}$-sentence σ_{l-1}, with σ_0 being a $\mathscr{B}$-sentence. $\square$

In "Testability and Meaning" (Carnap 1936, 1937), Carnap relaxes his claim of explicit definability of all scientific terms because he has come to the opinion (but does not prove) that it is impossible to define disposition terms explicitly in non-dispositional observational terms (Carnap 1936, 440). Instead, he suggests that new terms should be introduced by reduction pairs (442):[20]

> A pair of sentences of the forms
>
> $$Q_1 \supset (Q_2 \supset Q_3) \tag{R$_1$}$$
> $$Q_4 \supset (Q_5 \supset Q_3) \tag{R$_2$}$$
>
> is called a reduction pair for 'Q_3' provided '$\sim[(Q_1 \cdot Q_2) \vee (Q_4 \cdot Q_5)]$ is not valid.

Here, (R_1), for instance, stands for '$\forall x[Q_1 x \to (Q_2 x \to Q_3 x)]$' (434). A reduction pair is "either laid down in order to introduce 'Q_3' on the basis of Q_1, Q_2, Q_4, and Q_5, or consequences of physical laws stated beforehand" (443). I will thus call Q_3 'introducible by reduction pairs from ϑ on the basis of $\{Q_1, Q_2, Q_4, Q_5\}$ (or 'introducible' for short);[21] I will call the conjunctions $Q_1 x \wedge Q_2 x$ and $Q_4 x \wedge Q_5 x$ *reduction formulas* for Q_3.

19. The definitions of the definability of constant or function symbols additionally contain uniqueness conditions for the constants and function values, respectively. The conditions are philosophically interesting because they introduce restrictions on the sets of sentences in which constant and function symbols can be defined (Essler 1982, §14, §15; Hodges 1993, 59), but also introduce technical subtleties that would lead the current discussion too far afield.

20. Carnap also defines single reduction sentences, but these will not be relevant in the following. As in the *Aufbau*, Carnap (1936, §16) assumes that there are different "sufficient bases" for the reduction of scientific terms.

21. Carnap uses 'reducible' for another property of predicates. See the end of this section.

It is far from clear that reduction pairs suffice for analyzing the meaning of disposition concepts (cf. Belnap 1993, 136–137; Malzkorn 2001, §2.1). But empirical significance differs from meaning,[22] and introducibility may still be a criterion of empirical significance. For one, it is obvious that every $\mathscr{B}$-definable relation is also introducible by reduction sentences (with the two reduction formulas being contradictories and thus $\forall x \neg[(Q_1 x \wedge Q_2 x) \vee (Q_4 x \wedge Q_5 x)]$ a tautology). Introducibility is thus a straightforward weakening of a criterion of empirical significance that is usually considered too strong.

The relation between introducible terms and the criteria for sentences discussed so far is complicated. For instance, a sentence σ containing only introducible predicates can be both unverifiable and unfalsifiable (as shown in the appendix):

Claim 11. *For some sentences ϑ and some sentence σ whose predicates are introducible by reduction pairs from ϑ, σ is neither verifiable nor falsifiable relative to ϑ.*

Conversely, some sentences that containing exclusively introducible predicates are translatable.

Claim 12. *For some sentences ϑ and sentences σ whose relations are not introducible by reduction pairs from ϑ, σ can be non-trivially translated into $\mathscr{B}$ by ϑ.*

Proof. Choose $\mathscr{B} = \{B, b\}$ and $\vartheta = \forall x(Bx \leftrightarrow P_1 x \vee P_2 x)$. Then P_1 and P_2 are not introducible by a reduction pair but $\vartheta \vdash P_1 b \vee P_2 b \leftrightarrow Bb$. □

Introducibility therefore provides neither a necessary nor a sufficient condition for either verifiability or falsifiability, at least if the condition is to hold for all sentences and if it is to be based solely on the introducibility or non-introducibility of the predicates occurring in the sentences.

But things get still worse for criteria for terms. The culprit is again a recursive definition. For Carnap attempts to extend introducibility to include a recursion, even though on the face of it, introducibility can be applied indiscriminately to any term that occurs in ϑ. Carnap (1936, 446) gives the following definitions:

> A (finite) chain of (finite) sets of sentences is called an *introductive chain* based upon the class C of predicates if the following conditions are fulfilled. Each set of the chain consists either of one definition or of one or more reduction pairs for one predicate, say 'Q'; every predicate occurring in the set, other than 'Q', either belongs to C or is such that one of the previous sets of the chain is either a definition for it or a set of reduction pairs for it.
>
> [...] If the last set of a given introductive chain based upon C either consists in a definition for 'Q' or in a set of reduction pairs for 'Q', 'Q' is said to be introduced by this chain on the basis of C.

22. In short, 'being empirical significant' is a categorical predicate, while 'the meaning of' is a function.

Since chain-introducibility is relative to a theory ϑ, not all terms can be chain-introduced on the basis of C, that is, the definition will not be trivial in the way that the recursive criteria for sentences are. But there are good reasons to think that chain-introducibility is much too weak:

Claim 13. *Some ϑ contain relations that are chain-introducible on the basis of C, but are not introducible on the basis of C and are completely unrestricted in their interpretation by the interpretation of C.*

Proof. Choose $C = \{B_1, B_2\}$ and $\vartheta = \{\forall x[B_1x \rightarrow (B_2 \leftrightarrow P_1x)], \forall x[\neg B_1x \rightarrow (P_1x \leftrightarrow P_2x)]$. Then $\vartheta \not\vdash \neg\exists x(B_1x \wedge B_2x)$ and $\vartheta \not\vdash \neg\exists x(B_1x \wedge \neg B_2x)$, so that P_1 is introducible on the basis of $\{B_1, B_2\}$, and $\vartheta \not\vdash \neg\exists x(\neg B_1x \wedge P_1x)$ and $\vartheta \not\vdash \neg\exists x(\neg B_1x \wedge \neg P_1x)$, so that P_2 is introducible on the basis of $\{B_1, B_2, P_1\}$ and thus chain-introducible on the basis of $\{B_1, B_2\}$. But the interpretation of P_2 is completely unrestricted: For any interpretation of B_1 and B_2, the extension of P_1 is only determined within the extension of B_1, but it determines the extension of P_2 only within the anti-extension of B_1. It is clear that therefore P_2 is also not introducible on the basis of C. □

Carnap (1936, 447, theorem 7) proves what seems to be, in contradiction to the above result, the non-triviality of reductive chains. But his proof that "[i]f 'P' is introduced by an introductive chain based upon C, 'P' is reducible to C" turns out to be empty: The reducibility of a predicate is defined as the reducibility of the confirmation of the predicate, which in turn is defined via the reducibility of an atomic sentence involving the predicate Carnap (1936, 436). Since the reducibility of sentences is trivial, so is the reducibility of predicates.

Thus chain-introducibility is close to trivial, and there is no reason to believe that sentences containing only introducible terms are empirically significant. But these negative results should not detract from the importance of reduction pairs. In many cases in which an explicit definition for a term cannot be given, one sufficient and one (different) necessary condition will often do, and these can be phrased as reduction pairs. A special case of reduction pairs are "bilateral reduction sentences" (Carnap 1936, 442–43), which express conditional definitions. Even in mathematics, these seem to be more prevalent than definitions. For instance, one does not define for any object that it is continuous in such and such a case. Rather, one defines that a *function* is continuous in such and such a case. And this is a conditional definition. Thus reduction pairs *are* important. Only, it seems, not for empirical significance.

4 The United States

4.1 Criteria for Terms

In a short contribution to the *Unity of Science Forum*, Carnap summarizes "Testability and Meaning" (Carnap 1938, fn. 1) and points to the *Foundations of Logic* *and Mathematics* (Carnap 1939) for an elaboration of two methods of constructing a scientific language. One method starts with "elementary" terms ($\mathscr{B}$-terms) as primitive and successively introduces "abstract" terms ($\mathscr{A}$-terms) through reduction sentences as in "Testability and Meaning". The second method starts with abstract terms that are already related to each other through the postulates of a theory. These abstract terms are taken as primitive and further abstract terms are successively introduced to arrive eventually at elementary terms. In the second method, Carnap suggests, it may be possible to explicitly define all terms. In both methods, only the elementary terms are directly interpreted. Carnap (1938, 34) claims:

> The first way is interesting from the point of view of empiricism because it allows a closer check-up with respect to the empirical character of the language of science. By beginning our construction at the bottom, we see more easily whether and how each term proposed for introduction is connected with possible observations.

With its reliance on reduction sentences, the first method is supposed to relate abstract terms more easily to elementary terms.[23] It is easy to see that the conditions for abstract terms given by the second method can be very complicated. For not all definitions of elementary terms in abstract terms lead to necessary or sufficient conditions for the abstract terms. As an example, consider the definition of an elementary term B by four abstract terms A_1, A_2, A_3, A_4 in

$$\forall x \big(Bx \leftrightarrow [(A_1 x \wedge A_2 x) \vee (A_3 x \wedge A_4 x)]\big) \ . \tag{3}$$

The applicability of B to any object is neither necessary nor sufficient for the applicability of any of the four abstract terms. Also, the second method does not demand that all abstract terms occur non-trivially in the definition of an elementary term and some abstract terms may only be related to other abstract terms through the postulates of the theory, which are not further restricted.

In the *Foundations of Logic and Mathematics* itself, Carnap elaborates on the distinction between the two methods for relating abstract and elementary terms (figure 1). While the first method describes the observational import of abstract terms very clearly, scientists "are inclined to choose the second method" (Carnap 1939, 206, emphasis removed). Similarly, Carnap (1956b, 53) writes in "The Methodological Character of Theoretical Concepts":

> At the time of ["Testability and Meaning"], I still believed that all scientific terms could be introduced as disposition terms on the basis of observation terms either by explicit definitions or by so-called reduction sentences. Today I think, in agreement with most empiricists, that the connection between the observation terms and the terms of

23. Presumably because according to Carnap (1936, 447, thm. 7), it ensures that the abstract terms are reducible to elementary terms.

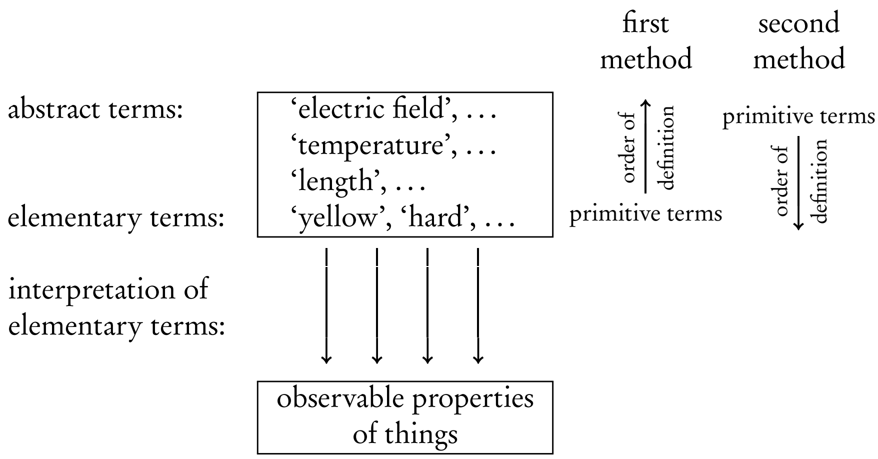

Figure 1: Carnap's diagram of two methods of giving an empirical interpretation to theoretical terms (adapted from Carnap 1939, 205).

> theoretical science is much more weak than it was conceived [...] in my earlier formulations [...].

The second method of constructing terms made it necessary to find a new criterion of empirical significance. Accordingly, Carnap (1956b) goes on to develop a weaker criterion for cases in which one cannot assume anything about the relation between the $\mathscr{B}$-terms and the rest of the scientific terms. In other words, Carnap develops a criterion that works in a very general framework and for arbitrary sentences. He only assumes higher order logic with semantic entailment (51, 61), a theory ϑ consisting of the conjunction $T.C$ of theoretical postulates T and correspondence rules C,[24] and a language with an observational sublanguage L_O with sentences containing only observational terms ($\mathscr{B}$-terms) and a theoretical sublanguage L_T with sentences that contain only terms from V_T ($\mathscr{A}$-terms). His suggestion is the following (51):

> A term 'M' is *significant relative to the class K of terms*, with respect to L_T, L_O, T, and C $=_{\mathrm{Df}}$ the terms of K belong to V_T, 'M' belongs to V_T but not to K, and there are three sentences, S_M and S_K in L_T and S_O in L_O, such that the following conditions are fulfilled:
>
> (a) S_M contains 'M' as the only descriptive term.
>
> (b) The descriptive terms in S_K belong to K.
>
> (c) The conjunction $S_M.S_K.T.C$ is consistent (i. e., not logically false).
>
> (d) S_O is logically implied by the conjunction $S_M.S_K.T.C$.
>
> (e) S_O is not logically implied by $S_K.T.C$.

24. The assumption that ϑ is a conjunction is not essential.

A major problem with the definition as it is stated is that, in contradiction to Carnap's intent (Carnap 1956b, 53), it is not logically weaker than introducibility. For assume

$$\vartheta = \bigwedge\{\exists x_1 x_2[x_1 \neq x_2 \wedge \forall y(y = x_1 \vee y = x_2)], \tag{4a}$$

$$\exists x[Bx \wedge \forall y(By \rightarrow x = y)], \tag{4b}$$

$$\forall x(Px \leftrightarrow Bx)\}\ . \tag{4c}$$

P is explicitly defined through a $\mathscr{B}$-term, but the only sentences that contain only P are either incompatible with ϑ (and thus fall afoul of (c)) or imply no new sentence in L_O (and thus fall afoul of (d) and (e)).

The solution to this apparent inconsistency in Carnap's claims is that Carnap, first, treats mathematical constants as logical constants and, second, allows for mathematical constants to have physical meaning and appear as arguments of V_T-relations. This becomes clear in Carnap's argument that his definition is not too narrow. To show this, Carnap (1956b, 59) considers a specific example in which one might think that it is too narrow, and argues that in this case,

> there must be a possible distribution of values of M for the space-time points of the region a', which is compatible with T, C, and S. Let 'F' be a logical constant, designating a mathematical function which represents such a value distribution. Then we take the following sentence as S_M: "For every space-time point in a', the value of M is equal to that of F." [...] Then S_M contains 'M' as the only descriptive term[.]

Carnap thus assumes that all mathematical terms are logical terms and can be identified with theoretical terms. This assumption seems to lead to a host of problems. For one, if two theoretical functions have the same values, they are identified with the same mathematical function and are thus identical, which may lead to trouble if they are related to different observation terms. It may thus be difficult to individuate theoretical terms, and may require a reformulation of many scientific theories, assuming that a consistent reformulation is even possible.

Avoiding the threat of inconsistency, one can read Carnap's proof as relying on the possibility of giving a direct interpretation to M. Read like this, his criterion of significance can be rephrased as follows:

Definition 7. A term M is *Carnap-significant relative to* the class K of terms with respect to L_T, L_O, T, and C if and only if $K \subseteq V_T$, $M \in V_T$, $M \notin K$, and there are an L_O-sentence β, an extension $\mathbf{M}$ of M, and a K-sentence $\varkappa$ such that

1. there is a model $\mathfrak{A}$ of $\varkappa \wedge \vartheta$ with $M^{\mathfrak{A}} = \mathbf{M}$,

2. every model $\mathfrak{A}$ of $\varkappa \wedge \vartheta$ with $M^{\mathfrak{A}} = \mathbf{M}$ is also a model of β, and

3. $\varkappa \cup \vartheta \nvDash \beta$.

This can be phrased shorter:

Claim 14. *A term M is* Carnap-significant relative to *the class K of terms with respect to L_T, L_O, T, and C if and only if $K \subseteq V_T$, $M \in V_T$, $M \notin K$, and there are an L_O-sentence β, an extension* $\mathbf{M}$ *of M, and a K-sentence $\varkappa$ such that*

$$\varnothing \subset \{M^{\mathfrak{A}} \mid \mathfrak{A} \vDash \varkappa \wedge \vartheta \wedge \neg\beta\} \subset \{M^{\mathfrak{A}} \mid \mathfrak{A} \vDash \varkappa \wedge \vartheta\} \tag{5}$$

Proof. $\varkappa \wedge \vartheta \nvDash \beta$ if and only if there is a model of $\varkappa \wedge \vartheta \wedge \neg\beta$, which holds if and only if the first proper subset relation holds. There is a model of $\varkappa \wedge \vartheta$ with $M^{\mathfrak{A}} = \mathbf{M}$ but no such model is also a model of $\neg\beta$ if and only if the second proper subset relation holds. □

The change from the notion of introducibility by reduction pairs is now clear: With reduction pairs, some *objects* in the domain are included in the extension of the introduced predicate and some objects are excluded from the extension of the predicate. By contrast, in definition 7 some *extensions* for a predicate are excluded.[25] In a sense, Carnap has moved the criterion for significance of terms one (type-theoretic) order higher, providing a necessary condition for a predicate being identical to M.

Unfortunately, Carnap goes further and gives a recursive definition of empirical significance (Carnap 1956b, 51):

> A term 'M_n' is *significant* with respect to L_T, L_O, T, and C $=_{\mathrm{Df}}$ there is a sequence of terms 'M_1',..., 'M_n' of V_T, such that every term 'M_i' ($i = 1, \ldots, n$) is significant relative to the class of those terms which precede it in the sequence, with respect to L_T, L_O, T, and C.

Given that significance was meant to be weaker than reducibility, it is not surprising that relations that with completely unrestricted interpretations can be significant (with Carnap's assumptions for his proof that his criterion is not too narrow):

Claim 15. *Assuming that there is a constant symbol c so that M_1c and M_2c are still considered to contain only M_1 and M_2, respectively, according to Carnap (1956b), some ϑ contain relations that are significant, but are completely unrestricted in their interpretation.*

Proof. Choose $\vartheta = \bigwedge\{\forall x(Bx \to P_1x), \forall x(\neg Bx \wedge P_1x \to P_2x)\}$. Then $\vartheta \wedge \neg P_1c \vDash \neg Bc$ while $\vartheta \nvDash \neg Bc$, and $\vartheta \wedge P_1c \wedge \neg P_2c \vDash \neg Bc$ while $\vartheta \wedge P_1c \nvDash \neg Bc$. But, similar to the proof of claim 13, P_2 is completely unrestricted in its interpretation. □

25. Note that excluding extensions is the same as *including* extensions: In both cases, one determines which extensions a predicate can have and which it cannot have.

In response to Van Cleve (1971) and Kaplan (1975), Creath (1976) suggests a recursive criterion of empirical significance for terms formulated in analogy to Carnap's criterion, but weaker. Since it seems that Carnap's definition is already to weak, this direction of the search for a criterion does not seem very fruitful.[26]

As in the case of introducibility and chain-introducibility, that Carnap's full criterion is close to trivial should not distract from its interesting recursion base. Having a necessary condition for the identity with some predicate M is often a significant step forward, and is sometimes all that is needed. The best example here is possibly Tarski's necessary condition for something being a truth-predicate.[27] Thus the recursion base of Carnap's criterion is again an interesting and important concept, although not for empirical significance.

Carnap (1956b, 49) also provides another interesting concept. After giving examples of correspondence rules (C-rules), he states:

> In the above examples, the C-rules have the form of universal postulates. A more general form would be that of statistical laws involving the concept of statistical probability [...]. A postulate of this kind might say, for example, that, if a region has a certain state specified in theoretical terms, then there is a probability of 0.8 that a certain observable event occurs [...]. Or it might, conversely, state the probability for the theoretical property, with respect to the observable event.

This generalization of the correspondence rules leads to a generalization of his criterion of significance as well. Thus in "The Methodological Character of Theoretical Concepts", Carnap hints at the to my knowledge only probabilistic criterion for the empirical significance of terms.

4.2 Criteria for Sentences

In "The Methodological Character of Theoretical Concepts", Carnap (1956b, 60) also gives a criterion for sentences:

> An expression A of L_T is a significant sentence of $L_T =_{\mathrm{Df}}$
>
> (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).

26. However, the recursion base of Creath's criterion seems to me more promising than the full definition, because it already allows determining whether a whole set of terms is significant. This arguably makes the recursive step of the definition superfluous.

27. In Tarski's fortunate case, the necessary condition is so strong that it even provides a sufficient condition up to extensional equivalence.

D2 is just the recursive definition of significance for terms that Carnap gives in the same article. This seems to be the first time that Carnap explicitly defines a sentence as significant if and only if it contains only significant terms.[28]

Even using only the recursion base of the definition of significant terms, however, it is possible to construct out of significant terms sentences that are not verifiable or falsifiable:

Claim 16. *There is a theory ϑ and a sentence σ such that σ contains only terms that are significant relative to $\varnothing$ but is not verifiable or falsifiable relative to ϑ.*

Proof. As the proof of claim 11. □

Obviously, things will not get better when the recursion step of the definition of significant terms is taken into account.[29]

There is also the worry that Carnap's criterion is incompatible with the motivation that he provides for it. Carnap (1956b, 49) writes:

> My task is to explicate the concept of empirical meaningfulness of theoretical terms. [. . .] In preparation for the task of explication, let me try to clarify the explicandum somewhat more, i. e., the concept of empirical meaningfulness in its presystematic sense. [. . .] What does it mean for 'M' to be *empirically meaningful*? Roughly speaking, it means that a certain assumption involving the magnitude M makes a difference for the prediction of an observable event. More specifically, there must be a certain sentence S_M about M such that we can infer with its help a sentence S_O in L_O.

So it seems that Carnap already makes a substantial assumption about what makes a sentence significant: We must be able to "infer with its help a sentence S_O in L_O", which essentially means that the sentence has to have L_O-content, and indeed, Carnap's conditions (c), (d), and (e) are exactly those of definition 4. But in that case, there is no need for any further definitions, which can at best be redundant, and at worst (as in this case) incompatible with the definition of empirical significance as having L_O-content.

It would be puzzling if Carnap had not seen this tension. And indeed, there is a possible solution to this puzzle. Carnap's intent may have been do define

28. In "Testability and Meaning", Carnap (1937, 34) states that in confirmable sentences (which are used to define the principle of empiricism), "[p]redicates which are confirmable [. . .] are admitted", but this does not mean that *only* confirmable predicates are admitted. As far as I can tell, Carnap (1936, 457, thm. 10) also only proves that a sentence containing exclusively confirmable predicates is itself confirmable. Hempel (1950, §3.1) may have intended his claim that the approach defining significant sentences via significant terms "has its origin in Carnap's essay, Testability and Meaning" as a historical rather than logical remark. The criterion he discusses, which demands that a significant sentence must contain only observable terms or those definable in observable terms, is his own. It did not originate with "Testability and Meaning", as Soames (2003, 292) claims.

29. Accordingly, I do not agree with a recent optimistic evaluation of Carnap's criterion (Justus, forthcoming), although I am optimistic about some of Carnap's other criteria.

empirically significant sentences so that all subformulas of a significant sentence are themselves significant.[30] In that case, every significant sentence must have L_O-content, but the inverse would not have to hold. Then the sentence S_M would be significant because it has L_O-content and no subformulas. It is only because there are sentences that are significant according to Carnap's definition but do not have L_O-content that *another* tension arises. But this one is not particularly obvious, and so might have been overlooked by Carnap.[31]

In a discussion of meaning and verifiability, Carnap (1963b, 887) remarks that the above criterion for the significance of terms "represents an explication of the requirement of confirmability in a modified form", where the requirement of confirmability is one thesis of empiricism (874):

> *Principle of confirmability*. If it is in principle impossible for any conceivable observational result to be either confirming or disconfirming evidence for a linguistic expression A, then expression A is devoid of cognitive meaning.

Since verifiability and falsifiability entail confirmability, this principle of confirmability is equivalent to the one suggested by Carnap (1928b, 327–328) early on. It seems, then, that Carnap's philosophical position has changed very little, although decades of technical work lies between these two statements of empiricism. It is this position that, for example, led Skyrms (1984, 14–15) to a Bayesian criterion of empirical significance that is equivalent to the demand that a significant sentence be confirmable or disconfirmable in the probabilistic sense.

5 Success

I have argued that despite his unshaken position on the form of a criterion of empirical significance, Carnap's technical endeavours bore mixed results at best. I now want to argue that his more successful technical endeavours in a slightly different context were so successful that they solve the problem of a criterion of significance as well.

In response to Hempel's criticism of the analytic-synthetic distinction (Carnap 1963b, 964), Carnap (1958, §4) argued that, without taking background assumptions into account, the synthetic component of a sentence

30. I thank Richard Creath for this point.

31. Besides this tension, the passage quoted above is connected to another lacuna. Carnap follows it up by in effect spelling out the different conditions of his criterion, and concludes with: "On the basis of the preceding considerations, I shall now give definitions for the concept of significance[.]" (Carnap 1956b, 51). But as a first step of an explication, the determination of the explicandum, it is remarkably poor. There is no investigation of actual usage or practically clear cases as Carnap (1950b, 4) in his elucidation of explication demands, but rather the use of pure intuition on comparably technical matters of deductive inferences, which one may share or not. Furthermore, Carnap also does not show the criterion's fruitfulness for the formulation of many universal statements, which is arguably the central requirement of a good explication (7).

$\sigma(B_1,\dots,B_m,A_1,\dots,A_n)$ that contains the basic terms $B_1,\dots,B_m$ and auxiliary terms $A_1,\dots,A_n$ can be identified with its *Ramsey sentence*

$$\mathsf{R}_{\mathscr{B}}(\sigma) := \exists X_1 \dots X_n \sigma\left(B_1,\dots,B_m,X_1,\dots,X_n\right) , \tag{6}$$

which results from σ by existentially generalizing on all $\mathscr{A}$-terms in σ.[32] $\mathsf{R}_{\mathscr{B}}(\sigma)$ entails the same $\mathscr{B}$-sentences as σ (Rozeboom 1962, 291–293) so that it is the content of σ, $\mathsf{R}_{\mathscr{B}}(\sigma) \dashv\vdash C(\sigma)$. The underlying assumption is that $\mathscr{B}$ is only restricted in its non-logical symbols; quantifiers and connectives can occur in any combination. For this reason, Carnap speaks of the 'extended observation language' (Psillos 2000, 158–159). To be the right choice for the *translation* of σ, $\mathsf{R}_{\mathscr{B}}(\sigma)$ would also have to be the ground of σ, which, as pointed out, it is not. Carnap's brilliant suggestion was to use the (by now) so-called Carnap sentence

$$\mathsf{C}_{\mathscr{O}}(\sigma) = \mathsf{R}_{\mathscr{O}}(\sigma) \rightarrow \sigma \ . \tag{7}$$

as the analytic component of σ. As analytic sentence, $\mathsf{C}_{\mathscr{B}}(\sigma)$ can be treated as a background assumption. After all, it is not under scrutiny when testing empirical claims. And with this choice of the background assumptions, Carnap's claims over the decades come together in one clean expression. As is easily shown,

$$\mathsf{C}_{\mathscr{B}}(\sigma) \vdash \sigma \leftrightarrow \mathsf{R}_{\mathscr{B}}(\sigma) , \tag{8}$$

that is, σ is translatable into $\mathscr{B}$, and

$$G(\sigma) \wedge \mathsf{C}_{\mathscr{B}}(\sigma) \dashv\vdash C(\sigma) \wedge \mathsf{C}_{\mathscr{B}}(\sigma) , \tag{9}$$

that is, the ground of σ is equivalent to the content of σ. Thus the weakest sentence that entails σ is also the strongest one entailed by σ and the content of a sentence is trivial (a tautology) if and only if its ground is trivial (a contradiction). For the criteria of empirical significance for sentences, the following holds:

Claim 17. *If $\mathscr{B}$ is only restricted by the terms it contains, the following statements are equivalent:*

1. *σ is non-trivially translatable into $\mathscr{B}$ by $\mathsf{C}_{\mathscr{B}}(\sigma)$.*
2. *σ is non-trivially verifiable relative to $\mathsf{C}_{\mathscr{B}}(\sigma)$.*
3. *σ is non-trivially falsifiable relative to $\mathsf{C}_{\mathscr{B}}(\sigma)$.*
4. *σ has non-trivial content relative to $\mathsf{C}_{\mathscr{B}}(\sigma)$.*

Proof. Choose $\mathsf{R}_{\mathscr{B}}(\sigma)$ as the translation, the content, and the verifying sentence of σ, and choose $\neg\mathsf{R}_{\mathscr{B}}(\sigma)$ as its falsifying sentence. Since formula (8) holds, the conditions for non-trivial translatability, verifiability, falsifiability, and having non-trivial content are equivalent to $\nvdash \mathsf{R}_{\mathscr{B}}(\sigma)$, which entails $\mathsf{C}_{\mathscr{B}}(\sigma) \nvdash \sigma$. □

32. Psillos (2000, §1, n. 7) argues that Carnap rediscovered the Ramsey sentence while trying to generalize Craig's theorem to type theory.

Thus a sentence that is non-trivially significant, by either being verifiable, falsifiable, or having content, is also non-trivially translatable, and so indeed for deductive inferences, all criteria for sentences become one.

This result became possible because in "Testability and Meaning", Carnap had led go of the idea of a fixed language. Instead, a theory can determine its own language and, in this case, its own analytic sentences. Thus Carnap's eventual success required, for one, the technical results of the Carnap and Ramsey sentence, but also the philosophical result of a strong conventionalism. In a way, the possibility to choose the language allowed Carnap to identify ground and content, and thus achieve the translatability of the theory.

Maybe the most impressive aspect of the Carnap sentence is that it can be applied to *any* theory. This aspect is also the most threatening for Carnap's original program, that of the criticism of metaphysics: σ can contain any well-formed sentence, and thus also sentences like 'The Absolute is perfect', which is thus translatable and also verified by its translation. However, this qualification of Carnap's success must itself be qualified: On its own, 'The Absolute is perfect' (in symbols: Ap) contains no $\mathscr{B}$-terms, and thus its Ramsey sentence is the tautology $\exists Y \exists x Yx$. Its verifiability and translation are thus trivial: It is an analytic truth, and thus verified by and translated into a tautology. It also cannot be falsified, since any falsifying sentence would have to be incompatible with the background assumptions, which is impossible. Thus as long as metaphysicians do not connect their sentences to $\mathscr{B}$-sentences, they may not be speaking nonsense, but also make no claims about the world. They are engaged in language choice.

One technical qualification has to be pointed out as well: The Ramsey and Carnap sentence approach relies on a fixed logic, and a somewhat fixed $\mathscr{B}$-language: The logic is that of higher order predicate logic, and $\mathscr{B}$ is the extended observation language. In a way, this is a small price to pay for the philosophical success, since there is nothing (besides, possibly, technical ability) that hinders the development of analogues of the Ramsey sentence and the Carnap sentence for other logics and other $\mathscr{B}$-languages. The hope is that there is also an analogue for inductive inferences.

6 Some Concluding Thoughts on the Criteria

Hempel famously switched from an optimistic outlook on criteria of empirical significance (Hempel 1950) to a pessimistic one (Hempel 1951) within one year. In his pessimistic conclusion, Hempel (1951, 74) writes that "cognitive significance in a system is a matter of degree", and sees this as a reason for disposing of the concept altogether. Instead of "dichotomizing this array [of systems] into significant and non-significant systems", he states, one should compare systems of sentences by their precision, systematicity, simplicity, and level of confirmation. But Carnap's approach using Ramsey and Carnap sentences in a way makes empirical significance a matter of degree without retreating to such notoriously elusive concepts as systematicity or simplicity: The logical strength of the Ramsey sentence can be seen a the degree to which a sentence has empirical content (and ground). It is only that a sentence with minimal degree of empirical significance is not meaningless but analytic.

This feature of Carnap's approach is again nicely illustrated by a discussion about the existence of God, which also sheds some light on Flew's argument. Adams (2011a) presents, according to his own summary (Adams 2011b) the following argument:[33]

> 1. [. . .] Thomas Aquinas reasoned that the universe must have a First Cause, to which he assigned the name God.
> 2. Modern physicists in their way are likewise in search of a First Cause.
> 3. If the physicists succeed, one taking the Thomistic view of things might reasonably call that First Cause God.

In reply to a strongly-worded criticism by a pseudonymous author, Adams (2011b) points out the following implication of his argument:

> Can we identify some fundamental principle or essence at the root of the universe and define that as the deity? Sure. Does doing so provide us with grounds for belief in a benevolent, all-knowing Creator? Clearly not. [. . .] To put it another way, the more closely we examine arguments for the existence of God, the more surely traditional belief in the deity slips from our grasp.

The claim that God exists if and only if a first cause exists plays the role of a Carnap sentence, and it reduces the content of the claim 'God exists' to nothing more than the claim 'There is *some* first cause', which plays the role of the Ramsey sentence. And this Ramsey sentence may have a completely non-theistic instantiation. Thus even if someone might respond to Flew's challenge to name a sentence that would disproof the existence of God with 'There is no first cause', this would be cold comfort for someone who expects God to be benevolent and all-knowing.

If metaphysical claims are saved but relegated to definitional status, the different criteria of empirical significance for terms may find their place in Carnap's system as well. One obvious role is the identification of terms that fulfill specific roles. Introducible terms can be used to identify specific kinds objects, terms that are significant according to the recursion base of the definition given in "The Methodological Character of Theoretical Concepts" identify specific kinds of properties. Other criteria may identify specific *groups* of properties.[34] Thus

33. Whether the interpretation of Thomas Aquinas is correct is irrelevant for my discussion.
34. The criterion by Creath (1976) may fit the bill.

in spite of their questionable help in identifying significant sentence or even significant terms, criteria for the significance of terms may have important uses in analyzing the components of a theory.

It seems then that, as far as deductive inference are concerned, Carnap's search for a criterion of empirical significance for sentences was a success, and his search for a criterion for terms was useful in producing analytical tools for the analysis of terms. With respect to inductive inferences, Carnap suggested the only criterion for terms, and his position towards a criterion for sentences can be spelled out in Bayesian terms if inductive inferences are explicated probabilistically. Abject failures look different.

A Further Proofs

Claim 3. *If all occurring probabilities are defined and $\mathscr{B}$ contains with every sentence also its negation, then σ is disconfirmable if and only if σ is confirmable.*

Proof.

$$\begin{aligned}\Pr(\sigma\,|\,\vartheta) > \Pr(\sigma\,|\,\beta\wedge\vartheta) = \frac{\Pr(\beta\,|\,\sigma\wedge\vartheta)\Pr(\sigma\,|\,\vartheta)}{\Pr(\beta\,|\,\vartheta)} &= \frac{1-\Pr(\neg\beta\,|\,\sigma\wedge\vartheta)\Pr(\sigma\,|\,\vartheta)}{1-\Pr(\neg\beta\,|\,\vartheta)}\\ \Leftrightarrow \Pr(\neg\beta\,|\,\vartheta) &< \Pr(\neg\beta\,|\,\sigma\wedge\vartheta)\\ \Leftrightarrow \Pr(\sigma\,|\,\vartheta) &< \frac{\Pr(\neg\beta\,|\,\sigma\wedge\vartheta)\Pr(\sigma\,|\,\vartheta)}{\Pr(\neg\beta\,|\,\vartheta)} = \Pr(\sigma\,|\,\neg\beta\wedge\vartheta) \end{aligned} \tag{10}$$

□

Claim 11. *For some sentences ϑ and some sentence σ whose predicates are introducible by reduction pairs from ϑ, σ is neither verifiable nor falsifiable relative to ϑ.*

Proof. Choose $\mathscr{B} = \{B, b\}$, $\mathscr{A} = \{P_1, P_2, P_3, P_4\}$, $\vartheta \dashv\vdash \{\forall x(Bx \to P_1x \wedge P_2x), \forall x(\neg Bx \to P_3x \wedge P_4x)\}$ and $\sigma \dashv\vdash (P_1b \wedge P_3b) \vee (\neg P_2b \wedge \neg P_4b)$. Then all terms in σ are reducible to $\mathscr{B}$. But σ is not falsifiable relative to ϑ

$$\begin{aligned}&\vdash \forall x[Bx \to \lambda y(y=y)x \wedge Bx] \wedge \forall x[\neg Bx \to \lambda y(y=y)x \wedge Bx]\\ &\quad \wedge \big([\lambda y(y=y)b \wedge \lambda y(y=y)b] \vee [\neg Bb \wedge \neg Bb]\big)\end{aligned} \tag{11a}$$

$$\begin{aligned}&\vdash \exists \bar{X}\big(\forall x[Bx \to X_1x \wedge X_2x] \wedge \forall x[\neg Bx \to X_3x \wedge X_4x]\\ &\qquad \wedge [(X_1b \wedge X_3b) \vee (\neg X_2b \wedge \neg X_4b)]\big)\end{aligned} \tag{11b}$$

$$\vdash \mathsf{R}_{\mathscr{B}}(\vartheta \wedge \sigma)\,. \tag{11c}$$

Since the Ramsey sentence of $\sigma \wedge \vartheta$ entails the same $\mathscr{B}$-sentences as $\sigma \wedge \vartheta$, $\sigma \wedge \vartheta$ and specifically does not entail any $\mathscr{B}$-sentences not already entailed by ϑ alone.

σ is also not verifiable relative to ϑ:

$$\begin{aligned}&\vdash \forall x[Bx \rightarrow Bx \wedge \lambda y(y=y)x] \wedge \forall x[\neg Bx \rightarrow \neg Bx \wedge \lambda y(y=y)x] \\ &\quad \wedge [\neg Bb \vee Bb] \wedge [\lambda y(y=y)b \vee \lambda y(y=y)b]\end{aligned} \tag{12a}$$

$$\begin{aligned}&\vdash \exists \bar{X} \left(\forall x[Bx \rightarrow X_1 x \wedge X_2 x] \wedge \forall x[\neg Bx \rightarrow X_3 x \wedge X_4 x]\right. \\ &\qquad \left.\wedge [\neg X_1 b \vee \neg X_3 b] \wedge [X_2 b \vee X_4 b]\right)\end{aligned} \tag{12b}$$

$$\vdash \mathsf{R}_{\mathscr{B}}(\vartheta \wedge \neg\sigma)\,. \tag{12c}$$

Since σ is verifiable if and only if $\neg\sigma$ is falsifiable, an argument analogous to the previous can be used. □

References

Adams, Cecil. 2011a. "Is there a God?" Archived by WebCite® at http://www.webcitation.org/6PCoxRqj6, *Chicago Reader,* November 11. http://www.straightdope.com/columns/read/3021/is-there-a-god.

———. 2011b. "Is there a God (revisited)?" Archived by WebCite® at http://www.webcitation.org/6PC13RKoE, *Chicago Reader,* December 2. http://www.straightdope.com/columns/read/3025/is-there-a-god-revisited.

Ayer, Alfred Jules. 1936. *Language, Truth and Logic.* 1st. References are to the second edition (Ayer 1946). London: Victor Gollanz.

———. 1946. *Language, Truth and Logic.* 2nd. London: Victor Gollanz.

———, ed. 1959. *Logical Positivism.* New York, NY: The Free Press.

Belnap, Nuel. 1993. "On Rigorous Definitions." *Philosophical Studies* 72:115–146.

Carnap, Rudolf. 1928a. *Der logische Aufbau der Welt.* References are to the translation (Carnap 1967a). Berlin-Schlachtensee: Weltkreis-Verlag.

———. 1928b. *Scheinprobleme in der Philosophie: Das Fremdpsychische und der Realismusstreit.* References are to the translation (Carnap 1967a). Berlin-Schlachtensee: Weltkreis-Verlag.

———. 1931a. "Die physikalische Sprache als Universalsprache der Wissenschaft." Nominal publication date incorrect. Published in 1932, *Erkenntnis* 2 (5/6): 432–465.

———. 1931b. "Überwindung der Metaphysik durch logische Analyse der Sprache." Nominal publication date incorrect. Published in 1932. References are to the translation (Carnap 1959b), *Erkenntnis* 2 (4).

———. 1932. "Über Protokollsätze." References are to the translation (Carnap 1987), *Erkenntnis* 3 (1): 215–228.

Carnap, Rudolf. 1932/1933. "Psychologie in physikalischer Sprache." References are to the translation (Carnap 1959a), *Erkenntnis* 3 (1): 107–142.

———. 1934. *Logische Syntax der Sprache.* Schriften zur wissenschaftlichen Weltauffassung. References are to the corrected translation (Carnap 1967b). Wien: Springer-Verlag.

———. 1935. *Philosophy and Logical Syntax.* Psyche Miniatures, General Series 70. London: Kegan Paul, Trench, Trubner & Co.

———. 1936. "Testability and Meaning." *Philosophy of Science* 3 (4): 420–468.

———. 1937. "Testability and Meaning—Continued." *Philosophy of Science* 4 (1): 2–35.

———. 1938. "Empiricism and the Language of Science." *Synthese* 3 (1).

———. 1939. *Foundations of Logic and Mathematics.* I,3:139–213. Foundations of the Unity of Science. Toward an International Encyclopedia of Unified Science. References are to the two-volume edition. Chicago and London: University of Chicago Press.

———. 1950a. "Empiricism, Semantics, and Ontology." References are to the slightly modified reprint (Carnap 1956a, appendix A), *Revue Internationale de Philosophie* 4:20–40.

———. 1950b. *Logical Foundations of Probability.* References are to the 2nd edition (Carnap 1962). Chicago: University of Chicago Press.

———. 1950c. *Testability and Meaning.* New Haven, Conn.

———. 1956a. *Meaning and Necessity.* 2nd. Chicago: University of Chicago Press.

———. 1956b. "The Methodological Character of Theoretical Concepts." In *The Foundations of Science and the Concepts of Psychology and Psychoanalysis,* edited by Herbert Feigl and Michael Scriven. Minnesota Studies in the Philosophy of Science. Minneapolis, MN: University of Minnesota Press.

———. 1958. "Beobachtungssprache und theoretische Sprache." *Dialectica* 12 (3-4): 236–248.

———. 1959a. "Psychology in Physical Language." In Ayer 1959, 165–198.

———. 1959b. "The Elimination of Metaphysics Through Logical Analysis of Language." In Ayer 1959, 60–81.

———. 1962. *Logical Foundations of Probability.* 2nd. Chicago: University of Chicago Press.

———. 1963a. "Intellectual Autobiography." In Schilpp 1963, 11:1–84.

Carnap, Rudolf. 1963b. "Replies and Systematic Expositions." In Schilpp 1963, 11:859–1016.

———. 1967a. *The Logical Structure of the World. Pseudoproblems of Philosophy.* Translation by Rolf A. George. Berkeley and Los Angeles: University of California Press.

———. 1967b. *The Logical Syntax of Language.* Reprinted with corrections. Translation by Amethe Smeaton (Countess von Zeppelin). London: Routledge & Kegan Paul Ltd.

———. 1987. "On Protocol Sentences." Translated by Richard Creath and Richard Nollan. *Noûs* 21 (4): 457–470.

Church, Alonzo. 1949. "Review of Ayer's *Language, Truth, and Logic*, 2nd Edition." *The Journal of Symbolic Logic* 14 (1): 52–53.

Creath, Richard. 1976. "On Kaplan on Carnap on Significance." *Philosophical Studies* 30:393–400.

Diamond, Malcolm L., and Thomas V. Litzenburg, eds. 1975. *The Logic of God: Theology and Verification.* Indianapolis, IN: Bobbs-Merill.

Essler, Wilhelm K. 1982. *Wissenschaftstheorie I. Definition und Reduktion.* 2nd. Kolleg Philosophie. Freiburg and München: Verlag Karl Alber.

Flew, Antony. 1950. "Theology and Falsification." References are to the reprint (Diamond and Litzenburg 1975, 257–259), *University* 1:1–8.

Friedl, Johannes, and Heiner Rutte. 2008. "Meaning and Verification—Editorischer Bericht." In *Die Wiener Zeit. Aufsätze, Beiträge, Rezensionen 1926–1936,* I.6:703–708. Moritz Schlick. Gesamtausgabe. Edited and with introductions by Johannes Friedl and Heiner Rutte. Wien: Springer-Verlag.

Goodman, Nelson. 1965. *Fact, Fiction, and Forecast.* 2nd. Indianapolis: Bobbs-Merrill.

Gupta, Anil. 2009. "Definitions." In *The Stanford Encyclopedia of Philosophy,* Spring 2009, edited by Edward N. Zalta. The Metaphysics Research Lab, Center for the Study of Language / Information, Stanford University.

Heidegger, Martin. 1931. *Was ist Metaphysik?* Bonn: Verlag von Friedrich Cohen.

Hempel, Carl G. 1945. "Studies in the Logic of Confirmation (II.)" *Mind* 54 (214): 97–121.

———. 1950. "Problems and Changes in the Empiricist Criterion of Meaning." *Revue Internationale de Philosophie* 11:41–63.

Hempel, Carl G. 1951. "The Concept of Cognitive Significance: A Reconsideration." *Proceedings of the American Academy of Arts and Sciences* 80 (1): 61–77.

Hodges, Wilfrid. 1993. *Model Theory.* Encyclopedia of Mathematics and its Applications. Digitally printed in 2008. Cambridge: Cambridge University Press.

Howson, Colin, and Peter Urbach. 1993. *Scientific Reasoning: The Bayesian Approach.* Second edition. Second printing 1996. Chicago and La Salle, IL: Open Court.

Justus, James. Forthcoming. "Carnap's Forgotten Criterion of Empirical Significance." *Mind.*

Kaplan, David. 1975. "Significance and Analyticity: a Comment on Some Recent Proposals of Carnap." In *Rudolph Carnap, Logical Empiricist: Materials and Perspectives,* edited by Jaakko Hintikka. Dordrecht, The Netherlands, Boston, MA: D. Reidel.

Lewis, David. 1988a. "Ayer's First Empiricist Criterion of Meaning: Why does it Fail?" *Analysis* 48 (1): 1–3.

———. 1988b. "Statements Partly About Observation." References are to the reprint (Lewis 1998, 125–155), *Philosophical Papers* 17:1–31.

———. 1998. *Papers in Philosophical Logic.* Cambridge Studies in Philosophy. Cambridge: Cambridge University Press.

Lutz, Sebastian. 2012. "Criteria of Empirical Significance: Foundations, Relations, Applications." PhD diss., Utrecht University. http://philsci-archive.pitt.edu/id/eprint/9117.

Malzkorn, Wolfgang. 2001. "Defining disposition concepts: A brief history of the problem." *Studies In History and Philosophy of Science Part A* 32 (2): 335–353.

Marhenke, Paul. 1949/1950. "The Criterion of Significance." *Proceedings and Addresses of the American Philosophical Association* 23:1–21.

Neurath, Otto. 1932. "Protokollsätze." *Erkenntnis* 3 (1): 204–214.

Pokriefka, Michael L. 1983. "Ayer's Definition of Empirical Significance Revisited." *Analysis* 43 (4): 166–170.

———. 1984. "More on Empirical Significance." *Analysis* 44 (2): 92–93.

Psillos, Stathis. 2000. "Rudolf Carnap's 'Theoretical Concepts in Science'." *Studies in History and Philosophy of Science Part A* 31 (1): 151–172.

Rozeboom, William W. 1962. "The Factual Content of Theoretical Concepts." In *Scientific Explanation, Space, and Time,* edited by Herbert Feigl, Michael Scriven, and Grover Maxwell, 3:273–357. Minnesota Studies in the Philosophy of Science. Minneapolis, MN: University of Minnesota Press.

Ruja, Harry. 1961. "The Present Status of the Verifiability Criterion." *Philosophy and Phenomenological Research* 22 (2): 216–222.

Schilpp, Paul Arthur, ed. 1963. *The Philosophy of Rudolf Carnap.* The Library of Living Philosophers. Chicago and LaSalle, IL: Open Court Publishing Company.

Skyrms, Brian. 1984. *Prediction and Experience.* New Haven, CT: Yale University Press.

Soames, Scott. 2003. *The Dawn of Analysis.* Philosophical Analysis in the Twentieth Century. Princeton, NJ: Princeton University Press.

Van Cleve, James. 1971. "Empiricist Criteria of Meaningfulness." Unpublished, cited by Creath (1976).

Wagner, Pierre. 2014. "Quel empirisme? Le projet de Carnap dans *Testabilité et signification*." In *Testabilité et signification,* by Rudolf Carnap, translated by Yann Benétreau-Dupin, 7–53. Translation of (Carnap 1936, 1937). Paris: Vrin.

Wright, Crispin. 1986. "Scientific Realism, Observation and the Verification Principle." In *Fact, Science and Morality: Essays on A. J. Ayer's Language, Truth and Logic,* edited by Graham Macdonald and Crispin Wright, 247–274. Oxford: Basil Blackwell.

———. 1989. "The Verification Principle: Another Puncture—Another Patch." *Mind* 98 (392): 611–622.

Yi, Byeong-Uk. 2001. "Compact Entailment and Wright's Verification Principle." *Mind* 110 (437): 414–421.