

# Carnap on empirical significance

Sebastian Lutz

Received: 30 April 2014 / Accepted: 15 September 2014 / Published online: 2 October 2014  
© Springer Science+Business Media Dordrecht 2014

**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 for definite results. His criteria for terms—introducibility by chains of 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 the assumptions needed for the Carnap sentence approach, verifiability, falsifiability, and translatability become equivalent. As a result of the Carnap sentence approach, metaphysics is rendered analytic.

**Keywords** Empirical significance · Carnap · Logical empiricism · Carnap sentence · Verifiability · Falsifiability

## 1 Introduction

The search for a criterion of empirical significance is generally considered an abject failure (see, e. g., [Ruja 1961](#); [Soames 2003](#), Chap. 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

---

S. Lutz (✉)  
Munich Center for Mathematical Philosophy,  
Ludwig-Maximilians-Universität München, Munich, Germany  
e-mail: [sebastian.lutz@gmx.net](mailto:sebastian.lutz@gmx.net)

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 suggested as an informal necessary condition for empirical significance that a sentence be translatable into a formal language. This criterion is arguably empty. On the formal side, Carnap was 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 its formal structure very close to earlier suggestions. The biggest difference is the successful criterion's presumption of a very inclusive notion of 'observation sentence', which would have been at odds with Carnap's earlier views. And there are two further plausible reasons why Carnap did not develop his successful criterion earlier. First, the criterion relies on the Carnap sentence, which Carnap discovered late in his career. Second, the criterion requires a very loose relation between theoretical and observational terms, which Carnap only gradually came to allow. A sort of corollary of the latter requirement is that Carnap's successful criterion cannot be used to criticize metaphysical sentences as meaningless, contrary to his initial aim. Rather, metaphysical sentences are rendered analytic.

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'. As in formal logic, *expressions* may be ill-formed and are thus not always 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  $\mathcal{L}$ . It also connects easily to the logical notion of a formula. All of Carnap's technical criteria furthermore assume a distinguished sublanguage of  $\mathcal{L}$ , which I will call the '*basic language*'  $\mathcal{B}$ . In the texts discussed here, Carnap calls it the 'autopsychological basis', the 'physical language', 'protocol sentences', 'observation language', and more. I will assume that  $\mathcal{B}$  can be identified with a set of sentences, and so I will speak of ' *$\mathcal{B}$ -sentences*'. Typically, Carnap identifies  *$\mathcal{B}$ -sentences* by their logical structure and the terms that occur in them, in which case I will speak of ' *$\mathcal{B}$ -terms*' (rather than 'observational' or 'elementary term').<sup>1</sup> I will call non- $\mathcal{B}$ -terms '*auxiliary*' or ' *$\mathcal{A}$ -terms*' rather than 'theoretical' or 'abstract terms'.

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

## 2 Informal translatability

In “Overcoming Metaphysics Through Logical Analysis of Language”,<sup>2</sup> Carnap (1931b, p. 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, p. 69). Such pseudo-statements can be hard to identify if they have the same historical-grammatical form as significant sentences. Carnap (1931b, p. 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. One sort results from “type confusion” (p. 75), where the types are those of Russell’s type theory.<sup>3</sup> 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 an ill-formed expression. An example of the other sort 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).

To criticize ‘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

<sup>2</sup> Arthur Pap translates the title more loosely as “The Elimination of Metaphysics Through Logical Analysis of Language” (Carnap 1959b).

<sup>3</sup> Rolf A. George’s 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).

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.<sup>4</sup>

It is a major drawback of this translatability condition of significance that it is informal: There is no formal way of deciding 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<sup>5</sup>] 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, p. 79) states: "I do not know how such questions could be translated into [any] unambiguous and clear mode; 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, p. 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, p. 259) responds to this problem

<sup>4</sup> 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.

<sup>5</sup> "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, p. 69; cf. Heidegger 1931, pp. 9–10).

like Carnap: “I therefore put to the succeeding symposiasts the simple central question, ‘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, p. 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 of empirical significance. As Carnap (1937, p. 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, Carnap (1937, p. 5) states that 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, if not empty, then at the least problematic, and the hope has to rest on criteria of empirical significance that apply to sentences in logical syntax.<sup>6</sup> 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, p. 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) on the one hand and of his later positions on the other. For in his earlier works, roughly those published while he was in Europe, Carnap relied on the assumption that it is possible to develop all the terms of science starting

<sup>6</sup> 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.

out from basic terms or sentences. His later works, published during his time in the United States, explicitly assume that this is not always expedient or even possible.<sup>7</sup>

With respect to the relation between basic and auxiliary sentences, 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” (p. 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 in the following, however, will concern the relations between the many different criteria.

### 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.<sup>8</sup> Translatability typically requires some background assumptions which, for convenience of notation, I will treat as one long conjunction  $\vartheta$ .

**Definition 1** Sentence  $\sigma$  is (non-trivially) translatable into language  $\mathcal{B}$  by sentence  $\vartheta$  if and only if there is a sentence  $\beta$  in  $\mathcal{B}$  such that  $\vartheta \vdash \sigma \leftrightarrow \beta$  (and  $\vartheta \not\vdash \beta$ ,  $\vartheta \not\vdash \neg\beta$ ).

In some of Carnap’s elucidations, the background assumptions are expressed by 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, p. 166). This is inessential for most of what follows, but will be picked up again in Sect. 6.1. In the *Aufbau*, the background assumptions are explicitly given as the definitions of the constitutional system, and thus they are determined by the reconstructed sciences (Carnap 1928a, §179). As far as the reconstruction of scientific theories is concerned, the basic sentences are those containing only logical (including set theoretical) terms (§107) and a single basic relation interpreted as a recollection of similarity (§108). To include values in the reconstruction, emotions and possibly volitions have to be included in  $\mathcal{B}$  as well. These then can be used to construct value experiences, which in turn can be used to construct values (Carnap 1928a, §152; cf. Mormann 2007, §2). Carnap (1928a, §133) is explicit, however, that emotions and volitions are probably too varied to be useful for establishing intersubjective agreement. Expressions that cannot be translated into  $\mathcal{B}$ -sentences according to Definition 1 are not significant. For instance, Carnap (1928a, §176) states that the concept of reality that is at issue in the debate between realism, idealism, and phenomenalism “cannot be constructed in an experiential constructional system; this characterizes it as a nonrational, metaphysical concept.” Those expressions about which the three

<sup>7</sup> Even in Carnap’s early works, however, his later position occurs as an undercurrent (Lutz 2012, §3.6.1).

<sup>8</sup> 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).

positions seem to disagree are therefore all in the field of metaphysics (§178), where ‘metaphysics’ refers to “the result of a non-rational, purely intuitive process” (§182).<sup>9</sup> Structurally, Carnap here relies on the criterion of significance as spelled out in “Testability and Meaning”: An expression must belong to the language of the constructional system, and it must fulfill the criterion of significance of this constructional system, in this case translatability. Purely intuitive, non-rational processes do not fulfill this criterion according to Carnap.

Later, Carnap (1931a, p. 452, all translations are mine) criticizes metaphysical expressions not for lack of translatability, but based on a different criterion. He states that “logical analysis comes to the conclusion [...] that the so-called metaphysical sentences are pseudo-sentences, since they stand in no deductive relation (neither positive nor negative) to the sentences of the protocol-language.”<sup>10</sup> The protocol sentences here are literals (atomic sentences or negations thereof) about “experiences, perceptions, but also feelings, thoughts, etc.” (p. 437).<sup>11</sup> Protocols are finite sets of such literals. In the following, I will treat the (finite) conjunctions of the members of a protocol as a single  $\mathcal{B}$ -sentence (hence  $\mathcal{B}$ -sentences are not protocol sentences, but conjunctions thereof). As in the *Aufbau*,  $\mathcal{B}$ -sentences can thus contain emotional terms, but unlike in the *Aufbau*, the sentences are restricted in their logical form; they do not, for instance, contain quantifiers.

Introducing some more a-historical terminology for the sake of precision, I will say that according to Carnap, metaphysical sentences are pseudo-statements because they are neither verifiable nor falsifiable by protocols<sup>12</sup>:

**Definition 2** Sentence  $\sigma$  is (non-trivially) verifiable in language  $\mathcal{B}$  relative to  $\vartheta$  if and only if there is a sentence  $\beta$  in  $\mathcal{B}$  such that  $\vartheta \wedge \beta \not\vdash \perp$  and  $\vartheta \wedge \beta \vdash \sigma$  (and  $\vartheta \not\vdash \sigma$ ).

**Definition 3** Sentence  $\sigma$  is (non-trivially) falsifiable in language  $\mathcal{B}$  relative to  $\vartheta$  if and only if there is a sentence  $\beta$  in  $\mathcal{B}$  such that  $\vartheta \wedge \beta \not\vdash \perp$  and  $\vartheta \wedge \beta \vdash \neg\sigma$  (and  $\vartheta \not\vdash \neg\sigma$ ).

The background assumptions  $\vartheta$  here contain, as in the *Aufbau*, the laws of nature. Carnap is silent about whether the laws must be known. If they must be known, significance depends on the current state of science. If they do not have to be known, significance is not so dependent. However, at any point in time one can then only make preliminary claims about significance.<sup>13</sup> The demand that  $\beta$  be compatible with  $\vartheta$  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

<sup>9</sup> Carnap (1928a, §179) of course does not think that metaphysics in the sense of “basic knowledge”, i. e., “logical, experiential, constructional order” is non-significant. In this case, however, he suggests using the terms ‘basic science’ or ‘cosmology’.

<sup>10</sup> “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.”

<sup>11</sup> “Erlebnisse, Wahrnehmungen, aber auch Gefühle, Gedanken usw.”

<sup>12</sup> This terminology follows that of Hempel (1950, pp. 45–48).

<sup>13</sup> As Justus’s criticism (2011, pp. 429–430) of a criterion by Sober (2008, §2.14) shows, this dilemma is still under discussion.

basic sentences be logically possible (Carnap 1936, p. 423; cf. Friedl and Rutte 2008). In the following, I will suppress the references to  $\vartheta$  and  $\mathcal{B}$  when this does not lead to ambiguity.

Shortly after demanding verifiability or falsifiability from significant sentences, Carnap (1932/1933, p. 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  $\sigma$  has (non-trivial) content in  $\mathcal{B}$  relative to sentence  $\vartheta$  if and only if there is a sentence  $\beta$  in  $\mathcal{B}$  such that  $\vartheta \not\vdash \beta$  and  $\vartheta \wedge \sigma \vdash \beta$  (and  $\vartheta \not\vdash \neg\sigma$ ).

The  $\mathcal{B}$ -sentences are again conjunctions of literals (pp. 165–166). The rules of inference (here given by  $\vartheta$ ) must be “independent of space and time” (p. 166), which suggests that they can express laws of nature, where Carnap is again silent about whether these laws must be known.

Somewhat surprisingly, Definition 4 provides but one of the conditions under which Carnap (1931a, p. 452) claimed a sentence to be significant shortly before:

**Claim 1** If  $\mathcal{B}$  contains with every sentence also its negation,<sup>14</sup> then a sentence is (non-trivially) falsifiable relative to  $\vartheta$  if and only if it has (non-trivial) content relative to  $\vartheta$ .

*Proof*  $\vartheta \wedge \beta \vdash \neg\sigma$  if and only if  $\vartheta \wedge \sigma \vdash \neg\beta$ .  $\vartheta \wedge \beta \not\vdash \perp$  if and only if  $\vartheta \not\vdash \neg\beta$ .  $\square$

And at around the same time, Carnap (1931b, p. 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 protocol sentences. Call the weakest  $\mathcal{B}$ -sentence that entails  $S$  the ground  $G(S)$  of  $S$ , and the strongest  $\mathcal{B}$ -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  $\mathcal{B}$ -sentence: The ground of  $S$  is equivalent to the disjunction of all protocol sentences entailing  $S$ , and the content is equivalent to the set of all protocol sentences

<sup>14</sup> Note that this puts additional restrictions on protocols.



entailed by  $S$ .<sup>15</sup> The relation between (1) and (4) then suggests that the ground and the content together determine a sentence's meaning, so that somehow, any sentence  $\sigma$  can be translated into its ground  $G(\sigma)$  and its content  $C(\sigma)$ . However, by definition,

$$G(\sigma) \wedge \vartheta \vdash \sigma \wedge \vartheta \vdash C(\sigma) \wedge \vartheta, \quad (1)$$

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

Put in a slightly different way, a sentence that is verifiable or has content may not be translatable. In fact, one can state something even stronger:

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

*Proof* Choose  $\vartheta$  to be a logical truth and some  $\mu$  that is not verifiable and has no content (say, a logically contingent sentence containing only terms not occurring in  $\mathcal{B}$ ), and choose two sentences  $\beta$  and  $\beta'$  from  $\mathcal{B}$  whose disjunction is also in  $\mathcal{B}$ . Then  $(\mu \wedge \beta') \vee \beta$  can be derived from  $\beta$  and entails  $\beta \vee \beta'$ , but it is not equivalent to a  $\mathcal{B}$ -sentence.  $\square$

Thus some sentences are even verifiable *and* have content without being translatable into  $\mathcal{B}$ . There are, then, 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. And it is as if Carnap assumes that all of these relations are equivalent, even though they are clearly not.<sup>16</sup>

In addition, in *Pseudoproblems of Philosophy* Carnap (1928b, p. 327) 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  $E$ . A statement  $p$  is said to be 'testable' if conditions can be indicated under which an experience  $E$  would occur which supports  $p$  or the contradictory of  $p$ . A statement  $p$  is said to have 'factual content', if experiences which would support  $p$  or the contradictory of  $p$  are at least conceivable, and if their characteristics can be indicated.

It seems that conditions under which an experience would occur are indicated by  $\mathcal{B}$ -sentences with spatio-temporally restricted quantifiers (Carnap's example is 'In

<sup>15</sup> For the ground of  $S$ , the equivalence can be expressed only in languages that allow disjunctions with infinitely many disjuncts.

<sup>16</sup> There is an additional wrinkle whose implications I will not discuss in the following: According to Carnap (1932), it was Neurath (1932) who first suggested treating protocol sentences and system-sentences as from the same language, which suggests that Carnap assumed in earlier discussions that the protocol sentences are in one language, the remaining ("system-")sentences in another. But Carnap already 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.

the next room is a three-legged table'). Conceivable experiences seem to be described by spatio-temporally unrestrictedly quantified sentences (Carnap's example is 'There is a certain red color whose sight causes terror'). By allowing the inference of  $p$  or  $\neg p$  through 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, p. 58; Reichenbach 1951, §1), but as discussed below, Carnap would later also suggest this approach.

The standard definitions of probabilistic confirmation and disconfirmation (e.g., Howson and Urbach 1993, p. 117) can be used to define confirmability and disconfirmability as follows:

**Definition 5** Assuming all occurring probabilities are well-defined, sentence  $\sigma$  is probabilistically confirmable in  $\mathcal{B}$  relative to sentence  $\vartheta$  if and only if there is a sentence  $\beta$  in  $\mathcal{B}$  such that

$$P(\sigma \mid \beta \wedge \vartheta) > P(\sigma \mid \vartheta) .$$

A sentence  $\sigma$  is probabilistically disconfirmable in  $\mathcal{B}$  relative to sentence  $\vartheta$  if and only if there is a sentence  $\beta$  in  $\mathcal{B}$  such that

$$P(\sigma \mid \beta \wedge \vartheta) < P(\sigma \mid \vartheta) .$$

Note that this definition is not one of total confirmability, since the probability of  $\sigma$  might simply be raised minimally from a very low value to a value almost as low. Analogously, it is not a definition of total disconfirmability.

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  $\mathcal{B}$  contains with every sentence also its negation, then  $\sigma$  is disconfirmable if and only if  $\sigma$  is confirmable.*

As is often discussed (e.g., Howson and Urbach 1993, pp. 119–120), if inductive inferences are treated as probabilistic, falsifiability entails confirmability in all interesting cases:

**Corollary 4** *If all occurring probabilities are defined,  $\mathcal{B}$  contains with every sentence also its negation,  $P(\beta \mid \vartheta) \neq 0$  for each  $\mathcal{B}$ -sentence compatible with  $\vartheta$ ,  $P(\sigma \mid \vartheta) \neq 0$ , and  $\sigma$  is falsifiable relative to  $\vartheta$ , then  $\sigma$  is confirmable relative to  $\vartheta$ .*

*Proof* If  $\beta \wedge \vartheta \vdash \neg\sigma$ , then  $P(\sigma \mid \beta \wedge \vartheta) = 0 < P(\sigma \mid \vartheta)$  so that  $\sigma$  is disconfirmable. By claim 3, it is confirmable.  $\square$

Informally, a sentence  $\sigma$  is confirmed when a  $\mathcal{B}$ -sentence that would have falsified  $\sigma$  turns out false. This result and Claim 1 have the immediate

**Corollary 5** *If all occurring probabilities are defined,  $\mathcal{B}$  contains with every sentence also its negation,  $P(\beta | \vartheta) \neq 0$  for each  $\mathcal{B}$ -sentence compatible with  $\vartheta$ ,  $P(\sigma | \vartheta) \neq 0$ , and  $\sigma$  has content relative to  $\vartheta$ , then  $\sigma$  is confirmable relative to  $\vartheta$ .*

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

**Corollary 6** *If all occurring probabilities are defined,  $\mathcal{B}$  contains with every sentence also its negation,  $P(\beta | \vartheta) \neq 0$  for each  $\mathcal{B}$ -sentence compatible with  $\vartheta$ ,  $P(\sigma | \vartheta) \neq 1$ , and  $\sigma$  is verifiable relative to  $\vartheta$ , then  $\sigma$  is disconfirmable.*

*Proof* If  $\beta \wedge \vartheta \vdash \sigma$ , then  $P(\sigma | \beta \wedge \vartheta) = 1 > P(\sigma | \vartheta)$  so that  $\sigma$  is confirmable. By claim 3, it is disconfirmable.  $\square$

Informally, a sentence  $\sigma$  is disconfirmed when a  $\mathcal{B}$ -sentence that would have verified  $\sigma$  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, p. 420) speaks of confirmability,<sup>17</sup> 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: In empiricism, 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, p. 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, p. 35).

<sup>17</sup> 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.” As far as the deducibility from  $\mathcal{B}$ -sentences is concerned, Carnap’s argument for his claim is essentially that *universally quantified* sentences cannot be verified. Carnap’s more general point is that  $\mathcal{B}$ -sentences themselves are never absolutely secure and thus also holds for falsifiability (Carnap 1936, §6).

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, p. 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  $\gamma$ , then the confirmation of every sentence  $\sigma$  is reducible to that of  $C$ .*

*Proof* For any sentence  $\sigma$ , if  $\gamma$  is directly incompletely reducible to that of  $C$ , so is  $\gamma \wedge \sigma$ , which can therefore be in  $C_1$ . Then the confirmation of  $\sigma$  can be completely reduced to that of  $C_1 := \{\gamma \wedge \sigma\}$  because  $\{\gamma \wedge \sigma\} \models \sigma$  and  $\{\gamma \wedge \sigma\}$  is a finite subset of itself. Thus the confirmation of  $\sigma$  is directly reducible to that of  $C_1$ , whose confirmation is directly reducible to that of  $C$ , and therefore the confirmation of  $\sigma$  is reducible to that of  $C$ .  $\square$

If a language contains infinitely many constants  $\{c_i \mid i \in I\}$  for points in space-time,<sup>18</sup> the sentence ‘It will always be everywhere cold’ is an incompletely directly reducible sentence  $\gamma$ , since the temperature at each point in space-time is logically independent from the temperature at any other and thus  $\gamma$  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, pp. 435–436); call such a sub-class a *confirmation class*. Full sentences are literals, and a contravalid sentence is incompatible with the laws of nature, where Carnap is again silent about whether these laws have to be known (pp. 432–434). Because of Claim 7, if some confirmation class  $\Omega$  allows the direct incomplete reducibility of at least one sentence  $\gamma$ , the confirmation of any sentence  $\sigma$  is reducible to  $\Omega$ . (In the above example,  $\Omega^*$  is a confirmation class for  $\gamma$  if  $\{c_i \mid i \in I\} \cup \{\lambda x(\text{It is cold at } x)\} \subseteq C$ .) Thus the confirmation of any sentence  $\sigma$  is reducible to that of  $C$ . If now  $C$  contains only observable predicates ( $\mathcal{B}$ -predicates),  $\sigma$  is confirmable, because a “sentence  $S$  is called *confirmable* [...] if the confirmation of  $S$  is reducible [...] to that of a class of observable predicates” (p. 456). Since nothing was assumed about  $\sigma$ , the principle of empiricism is then met by any sentence whatever.

As Wagner (2014, pp. 40–41) has shown in response to the above, Carnap (1950b, p. 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 *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

<sup>18</sup> This is what Carnap (1936, pp. 433–434) seems to assume.

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.<sup>19</sup>

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 this was exactly Carnap’s intention: Five years before the reprint, Hempel (1945, pp. 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  $\varepsilon \vdash \rho$ , then  $\varepsilon$  confirms  $\rho$ . Thus, specifically, any sentence  $\gamma$  confirms itself. The *converse consequence condition* demands that if  $\varepsilon$  confirms  $\rho$  and  $\rho' \vdash \rho$ , then  $\varepsilon$  also confirms  $\rho'$ . Thus  $\gamma$  confirms  $\gamma \wedge \sigma$ , where  $\sigma$  is any sentence whatever. According to the *special consequence condition*, if  $\varepsilon$  confirms  $\rho$  and  $\rho \vdash \rho'$ , then  $\varepsilon$  confirms  $\rho'$ . Thus  $\gamma$  confirms  $\sigma$ . 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,<sup>20</sup> 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  $\varphi(a_i) \not\vdash \varphi(a_j)$ ,  $\neg\varphi(a_j)$  for any  $i, j \in I, i \neq j$ . Carnap’s intent here seems to be something along the lines of a confirmation by (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,  $\varphi$  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  $\varphi$  and craft a new formula  $\varphi^*$  that predicates  $\varphi$  of the objects used in the induction, but predicates a completely different formula of all other objects. This insight, which unfortunately came too late for Carnap to have taken it into account for the reprint, can be used to trivialize Carnap’s new criterion. Unlike in Goodman’s argument, it is not even necessary to use  $\varphi^*$  as the definiens of a new predicate;  $\varphi^*$  itself already suffices. 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 (1950b), if the class  $C$  of sentences allows the direct incomplete reducibility of the confirmation of at least one sentence  $\gamma$  (by  $\{\varphi(a_i) \mid i \in$

<sup>19</sup> I thank Pierre Wagner for very helpful discussions of this publication of Carnap’s.

<sup>20</sup> Carnap (1950a, §87) discusses the conditions at length.

$I$ ) and contains the sentences  $\lceil a_j \neq b \rceil$ ,  $j \in J$  for some  $J \subseteq I$  of infinite cardinality, then the confirmation of every sentence  $\sigma$  is reducible to that of  $C$ .

*Proof* If the confirmation of  $\gamma$  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  $\sigma$  is any sentence whatever. The confirmation of the latter sentences is thus completely reducible to that of  $C$ .

By construction, the set  $C_2 = \{[\varphi(a_j) \wedge a_j \neq b] \vee [a_j = b \wedge \sigma] \mid j \in J\}$  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  $\gamma'$  is directly incompletely reducible to the confirmation of  $C_2$ .  $\gamma'$  entails  $\sigma$ , and thus the confirmation of  $\sigma$  is completely reducible to the confirmation of  $C_1 = \{\gamma'\}$ . Therefore, the confirmation of every sentence  $\sigma$  is reducible to that of  $C$ .  $\square$

The additional assumption of the proof is fulfilled in the example given above: If a language contains infinitely many constants  $\{c_i \mid i \in I\}$  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  $c_g$ ,  $\sigma$  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 contribution in “Testability and Meaning” to the search for a criterion of empirical significance was not successful. His informal discussion of the relation between confirming sentences and confirmed sentence provide a tantalizing suggestion for a thorough and precise empiricism, however, also because it relates to his informally suggested equivalence of verifiability, falsifiability, and translatability. Carnap also indirectly contributed another informal insight. Or rather, he steered clear of an 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, pp. 38–39, cf. Lewis 1988a), Ayer (1946, p. 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  $\mathcal{B}$ -sentences, the sentence or its negation is indirectly verifiable, and thus Ayer’s amended criterion is close to trivial as well. The criterion was followed by a slew of further amendments and new trivialization proofs (Pokriefka 1983, 1984;

Wright 1986, 1989, §II; Lewis 1988b, §IV, n. 12; Yi 2001). Like Ayer’s criterion, all of these criteria for sentences are recursive in that the background assumptions  $\vartheta$  (the “other premises” with which a verifiable statement has to entail a directly verifiable statement) themselves only have to be verifiable.<sup>21</sup> Thus there is some reason to think that recursive criteria of this kind are at the very least a dangerous direction of the search for a criterion of empirical significance for sentences.<sup>22</sup> This is of course no proof that there cannot be a successful recursive criterion of empirical significance for sentences, but arguably a reason for trying other directions first.

In contrast to Ayer, Carnap (1935, p. 11) writes: “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  $\mathcal{B}$ -content relative to other sentences. But in contrast to Ayer’s criterion, the other sentences in his criterion are required to be not only 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  $\beta, \gamma$  with  $\beta \not\sim \gamma$ , then any non-tautological sentence  $\sigma$  can be verified according to Carnap (1935, p. 11).*

*Proof*  $\{(\sigma \rightarrow \gamma) \wedge \beta\} \vdash \beta$  and is thus verified by  $\beta$ . Since  $\{\sigma, (\sigma \rightarrow \gamma) \wedge \beta\} \vdash \gamma$  while  $(\sigma \rightarrow \gamma) \wedge \beta \not\sim \gamma$ ,  $\sigma$  is indirectly verifiable.  $\square$

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, p. 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

<sup>21</sup> Note the difference to Carnap’s recursive condition in “Testability and Meaning”, where the background assumptions are fixed, but (essentially) the set  $\mathcal{B}$  changes in each recursive step.

<sup>22</sup> This is not true for the strongest criterion, translatability: Assume  $\sigma$  can be translated into  $\mathcal{B}$ -sentence  $\beta$  by  $\vartheta$  and  $\vartheta$  can be translated into  $\mathcal{B}$ -sentence  $\beta'$  without background assumptions. If  $\mathcal{B}$  is only restricted by the terms it contains,  $\vartheta$  is then a  $\mathcal{B}$ -sentence and therefore  $\sigma$  itself is a  $\mathcal{B}$ -sentence. In this case, a definition of empirical significance as recursive translatability is safe, but also pointless.

explicitly defined in  $\mathcal{B}$ -terms (Carnap 1928a, §38),<sup>23</sup> 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, 1963b, p. 945). In other words, he claims that these explicit definitions are entailed by scientific theories.<sup>24</sup>

**Definition 6** A relation  $A$  is  $\mathcal{B}$ -definable in  $\vartheta$  if and only if there is a  $\mathcal{B}$ -formula  $\varphi$  such that

$$\vartheta \vdash \forall x_1 \dots x_n [Ax_1 \dots x_n \leftrightarrow \varphi(x_1, \dots, x_n)]. \tag{2}$$

When  $\mathcal{B}$ -sentences are, as in the *Aufbau*, unrestricted in their logical form,  $\mathcal{B}$ -definability relates to translatability in a very straightforward sense:

**Claim 10** If  $\mathcal{B}$  is only restricted by the terms it contains and if  $\sigma$  is a sentence of  $\mathcal{B}$ -terms and  $\mathcal{B}$ -definable relations, then  $\sigma$  is translatable into  $\mathcal{B}$ .

*Proof* If  $A$  is  $\mathcal{H}$ -definable in  $\vartheta$ , then for every  $\mathcal{H} \cup \{A\}$ -sentence  $\sigma$  there is a  $\mathcal{H}$ -sentence  $\kappa$  such that  $\vartheta \vDash \sigma \leftrightarrow \kappa$  (Essler 1982, p. 103). Therefore, if the  $\mathcal{A}$ -relations in  $\sigma$  are  $\{A_1, \dots, A_{k+1}\}$ ,  $\sigma$  can be translated into a  $\mathcal{B} \cup \{A_1, \dots, A_k\}$ -sentence  $\sigma_k$ , and for  $1 \leq l \leq k$ ,  $\sigma_l$  can be translated into a  $\mathcal{B} \cup \{A_1, \dots, A_{l-1}\}$ -sentence  $\sigma_{l-1}$ , with  $\sigma_0$  being a  $\mathcal{B}$ -sentence.  $\square$

With the assumption of the  $\mathcal{B}$ -definability of all  $\mathcal{A}$ -terms and a very inclusive notion of ‘ $\mathcal{B}$ -sentence’, Carnap therefore establishes the translatability of all sentences  $\sigma$  into  $\mathcal{B}$ -sentences, and hence the equivalence of the ground of  $\sigma$ , the content of  $\sigma$ , and  $\sigma$  itself, given the (reconstructed) theory. It is thus unfortunate that Carnap had to give up this assumption.

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 that it is impossible to define disposition terms explicitly in non-dispositional observational terms (Carnap 1936, p. 440). Instead, he suggests that new terms should be introduced by reduction pairs (p. 442)<sup>25</sup>:

A pair of sentences of the forms

$$Q_1 \supset (Q_2 \supset Q_3) \tag{R_1}$$

$$Q_4 \supset (Q_5 \supset \sim Q_3) \tag{R_2}$$

<sup>23</sup> 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).

<sup>24</sup> 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, p. 59), but also introduce technical subtleties that would lead the current discussion too far afield.

<sup>25</sup> 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.



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 ‘ $\forall x[Q_1x \rightarrow (Q_2x \rightarrow Q_3x)]$ ’ (Carnap 1936, p. 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” (p. 443). I will thus call  $Q_3$  ‘introducible by reduction pairs from  $\vartheta$  on the basis of  $\{Q_1, Q_2, Q_4, Q_5\}$ ’ (or ‘introducible’ for short)<sup>26</sup> and the conjunctions  $Q_1x \wedge Q_2x$  and  $Q_4x \wedge Q_5x$  ‘reduction formulas’ for  $Q_3$ . Note that in this case the background assumptions  $\vartheta$  have to be known and can be analytic (when the reduction pair is “laid down”) or synthetic (when the reduction pair is a consequence of “physical laws stated beforehand”).

It is far from clear that reduction pairs suffice for analyzing the meaning of disposition concepts (Belnap 1993, pp. 136–137; Malzkorn 2001, §2.1). But empirical significance differs from meaning,<sup>27</sup> and introducibility may still be a criterion of empirical significance. For one, it is obvious that every  $\mathcal{B}$ -definable relation is also introducible by reduction sentences (with the two reduction formulas being contradictory and thus ‘ $\forall x \neg[(Q_1x \wedge Q_2x) \vee (Q_4x \wedge Q_5x)]$ ’ a contradiction). 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  $\sigma$  containing only introducible predicates can be both unverifiable and unfalsifiable (as shown in the Appendix):

**Claim 11** *For some sentences  $\vartheta$  there are simply existentially quantified sentences  $\sigma_{\exists}$  and simply universally quantified sentences  $\sigma_{\forall}$  such that all terms of  $\sigma_{\exists}$  and  $\sigma_{\forall}$  are introducible by reduction pairs from  $\vartheta$ , but  $\sigma_{\exists}$  and  $\sigma_{\forall}$  are neither verifiable nor falsifiable relative to  $\vartheta$ .*

Carnap (1937, §25) was well aware that typically, universally quantified sentences are not verifiable, existentially quantified sentences are not falsifiable, and sentences with mixed quantifiers are neither verifiable nor falsifiable. Claim 11 however shows that once one allows introducible terms in a sentence, even some simply quantified sentences are neither verifiable nor falsifiable. For simply universally quantified sentences this means that none of their unquantified instantiations are falsifiable, which strongly suggests that the sentences are not empirical.

Conversely, some sentences containing only non-introducible predicates are translatable.

**Claim 12** *For some sentences  $\vartheta$  and sentences  $\sigma$  for whose terms  $\vartheta$  entails neither necessary nor sufficient conditions in  $\mathcal{B}$ ,  $\sigma$  can be non-trivially translated into  $\mathcal{B}$  by  $\vartheta$ .*

*Proof* Let  $\vartheta$  be the conjunction of (i)  $\forall x(A_2x \leftrightarrow x = b) \vee \forall x(A_2x \leftrightarrow x \neq b)$ , (ii)  $\forall x(A_1x \leftrightarrow Bx) \vee (A_1x \leftrightarrow \neg Bx)$ , and (iii)  $\exists x[Bx \wedge \forall y(By \rightarrow x = y)] \wedge \exists xy(\neg Bx \wedge$

<sup>26</sup> Carnap uses ‘reducible’ for another property of predicates. See the end of this section.

<sup>27</sup> In short, ‘being empirical significant’ is a categorical predicate, while ‘the meaning of’ is a function.

$\neg B y \wedge x \neq y$ ). Then there are neither necessary nor sufficient conditions for  $A_1$  or  $A_2$ , and  $\forall x(A_2x \leftrightarrow A_1x) \vee \forall x(A_2x \leftrightarrow \neg A_1x)$  can be translated into  $\mathcal{B}$ -sentence  $Bb$  by  $\vartheta$ :  $\forall x(A_2x \leftrightarrow A_1x) \vee \forall x(A_2x \leftrightarrow \neg A_1x)$  and (i) entail  $\forall x(A_1x \leftrightarrow x = b) \vee \forall x(A_1x \leftrightarrow x \neq b)$ , which with (ii) entails  $\forall x(Bx \leftrightarrow x = b) \vee \forall x(Bx \leftrightarrow x \neq b)$ . With (iii),  $Bb$  follows. Conversely,  $Bb$  and (i) entail  $\forall x(A_2x \leftrightarrow Bx) \vee \forall x(A_2x \leftrightarrow \neg Bx)$  and hence with (ii) entail  $\forall x(A_2x \leftrightarrow A_1x) \vee \forall x(A_2x \leftrightarrow \neg A_1x)$ .  $\square$

If a term has neither a necessary nor a sufficient condition, it is a fortiori not introducible.<sup>28</sup> Introducibility therefore provides neither a necessary nor a sufficient condition for either verifiability, falsifiability, or translatability, 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. It can of course be that this or one of the following criteria that turn out problematic can be modified in a satisfying way. But here and in the following, I am mainly interested in the viability of Carnap's contributions to the search for a criterion of empirical significance. Note also that a modification of introducibility would have to deviate significantly from Carnap's suggestion, since it can neither be a logical strengthening nor a logical weakening: Claim 11 shows that introducibility leads to the inclusion of some very plausibly non-empirical sentences, and Claim 12 shows that introducibility leads to the exclusion of some clearly empirical sentences.<sup>29</sup>

But things get still worse for criteria for terms, because Carnap (1936, p. 446) extends introducibility to include the following recursion:

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  $C$ .

Note that unlike in Ayer's criterion, where  $\vartheta$  is recursively defined, Carnap's criterion recursively defines the set  $C$  of terms.<sup>30</sup> And since chain-introducibility is relative to a theory  $\vartheta$ , not all terms can be chain-introduced on the basis of  $C$ , that is, the definition cannot be trivial in the way that Ayer's criterion and the other recursive criteria for sentences are. But there are good reasons to think that chain-introducibility is much too weak:

<sup>28</sup> Incidentally, Claim 12 can be cheaply proved by choosing  $\vartheta = \{\beta \leftrightarrow \sigma\}$  for any  $\mathcal{B}$ -sentence  $\beta$  and any contingent sentence  $\sigma$  containing only  $\mathcal{A}$ -terms. However, since  $\vartheta$  is a claim of purely propositional logic, this may go against the spirit of Carnap's suggestion.

<sup>29</sup> Besides explicit definitions and reduction pairs, Przelećki (1969, Chaps. 7–8) also discusses systematic weakenings of reduction pairs.

<sup>30</sup> This is just a shorthand:  $C$  stays the same, but for the recursive step, it does not matter whether a term is in  $C$  or introducible through a chain of reduction pairs on the basis of  $C$ .

**Claim 13** Some  $\vartheta$  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  $\vartheta = \forall x[B_1x \rightarrow (B_2x \leftrightarrow A_1x)] \wedge \forall x[\neg B_1x \rightarrow (A_1x \leftrightarrow A_2x)]$ . 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  $A_1$  is introducible on the basis of  $\{B_1, B_2\}$ , and  $\vartheta \not\vdash \neg\exists x(\neg B_1x \wedge A_1x)$  and  $\vartheta \not\vdash \neg\exists x(\neg B_1x \wedge \neg A_1x)$ , so that  $A_2$  is introducible on the basis of  $\{B_1, B_2, A_1\}$  and thus chain-introducible on the basis of  $\{B_1, B_2\}$ . But the interpretation of  $A_2$  is completely unrestricted: For any interpretation of  $B_1$  and  $B_2$ , the extension of  $A_1$  is only determined within the extension of  $B_1$ , but it determines the extension of  $A_2$  only within the anti-extension of  $B_1$ . It is clear that therefore  $A_2$  is also not introducible on the basis of  $C$ .  $\square$

Carnap (1936, p. 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 a literal involving the predicate (Carnap 1936, p. 436). Since the reducibility of sentences is trivial, so is the reducibility of predicates.<sup>31</sup>

Thus chain-introducibility is close to trivial, and there is no reason to believe that all or only 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, pp. 442–443), that is, conditional definitions. These are ubiquitous in the empirical sciences (Schurz 2014, pp. 248–249), and even in mathematics, they 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

On New Year’s Eve in 1935, Carnap presented a paper with the title “Testability and Meaning” at a meeting of the Eastern Division of the American Philosophical Association (Benson 1963, item 1936-10), and less than a year later, he emigrated to the United States (Carnap 1963a, p. 34). In a short contribution to the *Unity of Science Forum*, Carnap (1938, fn. 1) summarizes “Testability and Meaning” and points to

<sup>31</sup> There could, of course, again be modifications of Carnap’s recursive step that do not trivialize his criterion, but since even the recursion base is problematic, the search may not be worth the effort.

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 ( $\mathcal{B}$ -terms) as primitive and successively introduces “abstract” terms ( $\mathcal{A}$ -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, p. 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.<sup>32</sup> 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

$$\forall x (Bx \leftrightarrow [(A_1x \wedge A_2x) \vee (A_3x \wedge A_4x)]) . \quad (3)$$

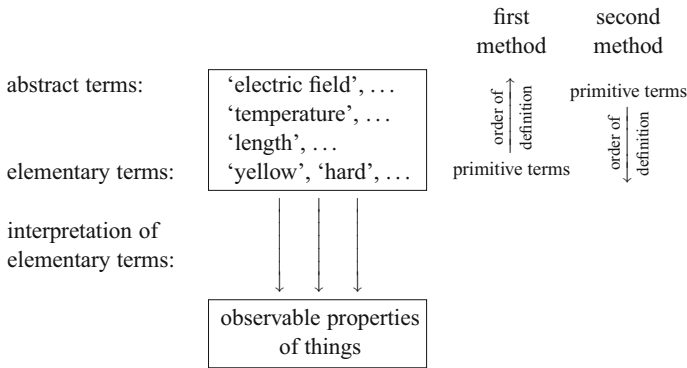
The applicability of  $B$  to any object is neither necessary nor sufficient for the applicability of any of the four abstract terms,<sup>33</sup> which are therefore also not introducible by the definition (3) on the basis of  $\{B\}$ . Furthermore, the second method does not demand that all abstract 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 (Fig. 1). While the first method describes the observational import of abstract terms very clearly, scientists “are inclined to choose the second method” (Carnap 1939, p. 206, emphasis removed). In “The Methodological Character of Theoretical Concepts”, Carnap (1956, p. 53) puts this more forcefully:

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 observa-

<sup>32</sup> Presumably because according to Carnap (1936, p. 447, thm. 7), it ensures that the abstract terms are reducible to elementary terms.

<sup>33</sup> Each abstract term is an INUS condition for  $B$  (Mackie 1980, p. 62).



**Fig. 1** Carnap’s two methods of giving an empirical interpretation to theoretical terms (adapted from Carnap 1939, p. 205)

tion terms and the terms of 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 (1956) goes on to develop a weaker criterion for cases in which one cannot assume anything about the relation between the  $\mathcal{B}$ -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 (pp. 51, 61), a theory  $\vartheta$  consisting of the conjunction  $T.C$  of theoretical postulates  $T$  and correspondence rules  $C$ ,<sup>34</sup> and a language with an observational sublanguage  $L_O$  with sentences containing only observational terms ( $\mathcal{B}$ -terms) and a theoretical sublanguage  $L_T$  with sentences that contain only terms from  $V_T$  ( $\mathcal{A}$ -terms). The logical structure of  $L_O$  (that is,  $\mathcal{B}$ ) is much more inclusive compared to “Testability and Meaning”, since it can contain quantifiers as long as they range only over observable objects.  $\mathcal{B}$  here seems to be very similar to the language Carnap assumed for the description of “conceivable experiences” in *Pseudoproblems of Philosophy* (Carnap 1928b, p. 327). Carnap’s suggestion for a new criterion of significance is the following (Carnap 1956, p. 51):

A term ‘ $M$ ’ is *significant relative to the class  $K$*  of terms, with respect to  $L_T$ ,  $L_O$ ,  $T$ , and  $C =_{Df}$  the terms of  $K$  belong to  $V_T$ , ‘ $M$ ’ 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  $S_K.T.C$ .

<sup>34</sup> The assumption that  $\vartheta$  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 1956, p. 53), it is not logically weaker than introducibility. For assume

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

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

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

$A$  is even  $\mathcal{B}$ -definable in  $\vartheta$ , but the only sentences that contain only  $A$  are either incompatible with  $\vartheta$  (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 (1956, p. 59) considers a specific example in which one might think that his definition 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 that take space-time points as arguments. This assumption seems to lead to a host of problems. For instance, 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  $M$ . Thus  $S_M$  is replaced by an interpretation  $M^{\mathfrak{A}}$  of some structure  $\mathfrak{A}$ , and the consistency of  $S_M.S_K.T.C$  is expressed by  $\mathfrak{A}$  being a model of  $S_K.T.C$ . Read like this, Carnap’s criterion of significance can be rephrased as follows<sup>35</sup>:

**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  $\beta$ , an extension  $\mathbf{M}$  of  $M$ , and a  $K$ -sentence  $\kappa$  such that

- (1) there is a model  $\mathfrak{A}$  of  $\kappa \wedge \vartheta$  with  $M^{\mathfrak{A}} = \mathbf{M}$ ,
- (2) every model  $\mathfrak{A}$  of  $\kappa \wedge \vartheta$  with  $M^{\mathfrak{A}} = \mathbf{M}$  is also a model of  $\beta$ , and
- (3)  $\kappa \wedge \vartheta \not\models \beta$ .

<sup>35</sup> Since Carnap (1956, pp. 51, 62) assumes that logical implication is given by semantic entailment, I use ‘ $\models$ ’ rather than ‘ $\vdash$ ’.

This can be stated more briefly:

**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  $\beta$  and a  $K$ -sentence  $\kappa$  such that*

$$\emptyset \subset \{M^{\mathfrak{M}} \mid \mathfrak{M} \models \kappa \wedge \vartheta \wedge \neg\beta\} \subset \{M^{\mathfrak{M}} \mid \mathfrak{M} \models \kappa \wedge \vartheta\} \tag{5}$$

*Proof* The first proper subset relation holds if and only if there is a model of  $\kappa \wedge \vartheta \wedge \neg\beta$ , that is,  $\kappa \wedge \vartheta \not\models \beta$ . The second proper subset relation holds if and only if for some  $\mathbf{M}$  there is a model of  $\kappa \wedge \vartheta$  with  $M^{\mathfrak{M}} = \mathbf{M}$  but no such model is also a model of  $\neg\beta$ . This holds if and only if for some  $\mathbf{M}$  there is a model of  $\kappa \wedge \vartheta$  with  $M^{\mathfrak{M}} = \mathbf{M}$  and every such model is also a model of  $\beta$ .  $\square$

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. The inclusion of an extension is a special case of the exclusion of extensions: If all extensions but one are excluded, the remaining extension is included. Conversely, only one extension can be included, since then all others are excluded. In a sense, Carnap has moved the criterion for significance of terms one (type-theoretic) order higher, providing a necessary condition for an extension being that of  $M$  (which sometimes amounts to a sufficient condition as well).

Unfortunately, Carnap goes further and gives a recursive definition of empirical significance (Carnap 1956, p. 51).<sup>36</sup>

A term ‘ $M_n$ ’ is *significant* with respect to  $L_T, L_O, T$ , and  $C =_{Df}$  there is a sequence of terms ‘ $M_1$ ’, ..., ‘ $M_n$ ’ of  $V_T$ , such that every term ‘ $M_i$ ’ ( $i = 1, \dots, n$ ) 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 with completely unrestricted interpretations can be significant (under 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  $A_{1c}$  and  $A_{2c}$  are still considered to contain only  $A_1$  and  $A_2$ , respectively, some  $\vartheta$  contain relations that are significant according to Carnap (1956), but are completely unrestricted in their interpretation.*

*Proof* Choose  $\vartheta = \forall x(Bx \rightarrow A_1x) \wedge \forall x(\neg Bx \wedge A_1x \rightarrow A_2x)$ . Then  $\vartheta \wedge \neg A_{1c} \vdash \neg Bc$  while  $\vartheta \not\vdash \neg Bc$ , and  $\vartheta \wedge A_{1c} \wedge \neg A_{2c} \vdash \neg Bc$  while  $\vartheta \wedge A_{1c} \not\vdash \neg Bc$ . But, similar to the proof of Claim 13,  $A_2$  is completely unrestricted in its interpretation.  $\square$

<sup>36</sup> Note that here, too, it is the set  $K$  of terms that is recursively defined, not  $\vartheta$  (here:  $T.C$ ).

In response to Kaplan (1975) and Van Cleve (1971),<sup>37</sup> Creath (1976) suggests a recursive criterion of empirical significance for terms formulated in analogy to Carnap's criterion, but weaker (see Justus, forthcoming, §4). Since it seems that Carnap's definition is already too weak, this direction of the search for a criterion does not seem very fruitful.<sup>38</sup>

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 extension of 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 the extension of a truth-predicate.<sup>39</sup> Thus the recursion base of Carnap's criterion is again an interesting and important concept, although not for empirical significance.

Carnap (1956, p. 49) also suggests 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 would, if worked out, lead 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 (1956, p. 60) also gives a criterion for sentences:

An expression  $A$  of  $L_T$  is a significant sentence of  $L_T =_{\text{Df}}$

- (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).

<sup>37</sup> Formula (3) is closely related to the method of "de-Occamization" suggested by Van Cleve (reported by Creath 1976, 397). Using Claim 14, it is easy to see that in formula (3), none of the  $\mathcal{A}$ -terms is Carnap-significant relative to  $\emptyset$ .

<sup>38</sup> However, the recursion base of Creath's criterion seems to me more promising than the full definition, because it already allows for determining whether a whole set of terms is significant. It thus could arguably be a non-trivial substitute for Carnap's recursive definition.

<sup>39</sup> In Tarski's fortunate case, the necessary condition is so strong that it provides a sufficient condition as well.



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.<sup>40</sup>

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** *For some sentences  $\vartheta$  there are simply existentially quantified sentences  $\sigma_{\exists}$  and simply universally quantified sentences  $\sigma_{\forall}$  such that all terms of  $\sigma_{\exists}$  and  $\sigma_{\forall}$  are significant relative to  $\varnothing$  but  $\sigma_{\exists}$  and  $\sigma_{\forall}$  are neither verifiable nor falsifiable relative to  $\vartheta$ .*

*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.<sup>41</sup>

There is also the worry that Carnap's criterion is incompatible with the motivation that he provides for it. Carnap (1956, p. 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, the conditions (c), (d), and (e) of Carnap's criterion for terms 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 to define empirically significant sentences so that all subformulas of a significant sentence are themselves

<sup>40</sup> In “Testability and Meaning”, Carnap (1937, p. 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, p. 457, thm. 10) also proves only that a sentence containing exclusively confirmable predicates is itself confirmable. Hempel (1950, p. 51) 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 (definition 3.2), 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, p. 292) claims.

<sup>41</sup> Accordingly, I do not agree with the optimistic evaluation of Carnap's criterion by Justus (forthcoming), although I am optimistic about some of Carnap's other criteria.

significant.<sup>42</sup> In that case, every significant sentence must have  $L_O$ -content, but the inverse would not have to hold; 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.<sup>43</sup>

In a discussion of meaning and verifiability, Carnap (1963b, p. 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 (p. 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, p. 327) early on in *Pseudoproblems of Philosophy*. It seems, then, that Carnap's philosophical position has changed very little, although decades of technical work lie between these two statements of empiricism. It is this position that, for example, led Skyrms (1984, pp. 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. His criteria for terms do not seem to be usable for identifying significant sentences or are next to trivial, the suggested equivalence between the ground and the content of a sentence seems false, and the claim of translatability by stipulation is but a tantalizing suggestion. 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 finding a criterion of significance as well, albeit only for deductive inferences.

In response to Hempel's criticism of the analytic-synthetic distinction (Carnap 1963b, p. 964), Carnap (1958, §4) argued that, without taking background assumptions into account, the synthetic component of a sentence  $\sigma(B_1, \dots, B_m, A_1, \dots, A_n)$  that

<sup>42</sup> I thank Richard Creath for this point.

<sup>43</sup> 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 1956, p. 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 (1950a, p. 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 (p. 7). Finally, Carnap provides a justification for his criterion for terms, but gives no justification at all for his criterion for sentences. (I thank Christopher French for discussions and his spirited defense of Carnap in this matter).

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

$$R_{\mathcal{B}}(\sigma) := \exists X_1 \dots X_n \sigma(B_1, \dots, B_m, X_1, \dots, X_n) , \tag{6}$$

which results from  $\sigma$  by existentially generalizing on all  $\mathcal{A}$ -terms in  $\sigma$ .<sup>44</sup>  $R_{\mathcal{B}}(\sigma)$  entails the same  $\mathcal{B}$ -sentences as  $\sigma$  (Rozeboom 1962, pp. 291–293) so that it is the content of  $\sigma$ ,  $R_{\mathcal{B}}(\sigma) \vdash C(\sigma)$ . The underlying assumption is that  $\mathcal{B}$  is only restricted in its non-logical symbols; quantifiers and connectives can occur in any combination and can range over any objects. For this reason, Carnap speaks of the ‘extended observation language’ (Psillos 2000b, pp. 158–159). This is a significant change from Carnap’s earlier assumptions about  $\mathcal{B}$ , which typically required  $\mathcal{B}$  to contain only conjunctions of literals and even in “The Methodological Character of Theoretical Concepts” contains only sentences whose quantifiers range over observable objects. It is the notion of  $\mathcal{B}$ -sentence that Carnap last used in the *Aufbau*.

After having loosened the relation between  $\mathcal{B}$ - and  $\mathcal{A}$ -terms in *Foundations of Logic and Mathematics*, Carnap has taken another essential step towards the unification of verifiability, falsifiability, and translatability. The final step makes use of the fact that the Ramsey sentence provides a closed expression for the empirical content of any sentence. It consists in Carnap’s suggestion to use the (by now) so-called Carnap sentence

$$C_{\mathcal{B}}(\sigma) = R_{\mathcal{B}}(\sigma) \rightarrow \sigma \tag{7}$$

as the analytic component of  $\sigma$ .<sup>45</sup> As analytic sentence,  $C_{\mathcal{B}}(\sigma)$  can be treated as a background assumption. After all, it is not under scrutiny when testing empirical claims. Taking  $C_{\mathcal{B}}(\sigma)$  as background assumption provides an immediate advantage over the choice of not further specified “laws of nature” or similar, since the status of  $C_{\mathcal{B}}(\sigma)$  is completely clear: It is an analytic sentence, and specifically the analytic component of a known (or conjectured) theory.<sup>46</sup> And with this choice of the background assumptions, Carnap’s claims over the decades come together in one clean expression. As noted in connection with the general relation between ground and content of a sentence (1), ground and content are not equivalent for every sentence  $\sigma$  and all background assumptions  $\vartheta$ . But the Carnap sentence  $C_{\mathcal{B}}(\sigma)$  renders a component of  $\sigma$  itself a background assumption, and as is easily shown (see also Winnie 1970, Eq. 6),

$$C_{\mathcal{B}}(\sigma) \vdash \sigma \leftrightarrow R_{\mathcal{B}}(\sigma) , \tag{8}$$

that is,  $\sigma$  is translatable into  $\mathcal{B}$  relative to  $C_{\mathcal{B}}(\sigma)$ , and

$$G(\sigma) \wedge C_{\mathcal{B}}(\sigma) \vdash C(\sigma) \wedge C_{\mathcal{B}}(\sigma) , \tag{9}$$

<sup>44</sup> Psillos (2000a, §4) argues that Carnap rediscovered the Ramsey sentence while trying to generalize Craig’s theorem to type theory.

<sup>45</sup> For the conjunction of two reduction pairs, Carnap (1936, §10) already describes an equivalent method in “Testability and Meaning” (cf. Carnap 1963b, pp. 964–965).

<sup>46</sup> That this does not render claims of empirical significance preliminary or knowledge dependent is shown in Sect. 6.1.

that is, the ground of  $\sigma$  is equivalent to the content of  $\sigma$ . Thus the weakest sentence that entails  $\sigma$  is also the strongest one entailed by  $\sigma$  and the content of a sentence is trivial if and only if its ground is trivial. And now, after a search of thirty years for a technical expression of empirical significance, after a change of the semantics of scientific terms, a change of the notion of a basic sentence, and the rediscovery of the Ramsey sentence, the following holds for the criteria of empirical significance for sentences:

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

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

*Proof* Choose  $\mathbf{R}_{\mathcal{B}}(\sigma)$  as the translation, the content, and the verifying sentence of  $\sigma$ , and choose  $\neg\mathbf{R}_{\mathcal{B}}(\sigma)$  as its falsifying sentence. Since formula (8) holds, the conditions for non-trivial translatability, verifiability, falsifiability, and having non-trivial content are equivalent to  $\not\vdash \mathbf{R}_{\mathcal{B}}(\sigma)$ , which entails  $\mathbf{C}_{\mathcal{B}}(\sigma) \not\vdash \sigma$ .  $\square$

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. As Carnap (1936, p. 420) had put it over 20 year earlier, “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”.<sup>47</sup>

## 6 Concluding thoughts on the criteria

Carnap (1963b, p. 962) defines the “O-content of  $S$  relative to  $TC$ ” as the  $\mathcal{B}$ -sentences entailed by  $S \wedge TC$  (so that a sentence with empty O-content has no non-trivial content in  $\mathcal{B}$  according to Definition 4) and notes that  $S$  may entail inductive relations beyond those entailed by its O-content. But “although it cannot replace  $S$  completely, the O-content of  $S$  relative to a given theory  $TC$  may still be taken as an explication for the experiential import (or, if one prefers, the deductive experiential import) of  $S$ ”. Thus for Carnap, the Ramsey sentence of  $\sigma$  does not cover the inductive content of  $\sigma$ , and hence the success of the Carnap sentence is restricted to deductive inferences. As far as criteria of empirical significance are concerned, this encompasses verification, falsification, and translation. The search for a closed expression of inductive content, an inductive analogue to the Carnap sentence, and thus a solution to the problem of inductive confirmability is not over. In the remainder of this essay, however, I will instead focus on the features of Carnap’s solution for deductive inferences.

<sup>47</sup> The criticism by Demopoulos (2007, VII) that, essentially, the Carnap sentence renders a theory analytically equivalent to its content in  $\mathcal{B}$  is therefore but an expression of Carnap’s conception of empiricism.

## 6.1 The role of the Ramsey and the Carnap sentence

To a certain extent, Carnap's use of the Ramsey sentence even provides a response to an influential argument by Hempel (1951, p. 74), who shows the intuitive inadequacy of a number of extant criteria of empirical significance under the assumption that  $\mathcal{B}$  contains only conjunctions of literals. He concludes 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 in a way, Carnap's use of the Ramsey sentence 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 as the degree to which a sentence is empirically significant (or, better: the degree to which it has content), with a tautological Ramsey sentence a mark of non-falsifiability:

**Claim 18** *If  $\mathcal{B}$  is only restricted by the terms it contains,  $\sigma$  has non-trivial content in  $\mathcal{B}$  if and only if  $\neg R_{\mathcal{B}}(\sigma)$ .*

*Proof*  $R_{\mathcal{B}}(\sigma) \vdash R_{\mathcal{B}}(\sigma \wedge C_{\mathcal{B}}(\sigma))$  and thus  $R_{\mathcal{B}}(\sigma)$  entails the same  $\mathcal{B}$ -sentences as  $\sigma$  relative to  $C_{\mathcal{B}}(\sigma)$ .  $\square$

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:<sup>48</sup>

- (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 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',<sup>49</sup> 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 disprove the existence of

<sup>48</sup> Whether his interpretation of Thomas Aquinas is correct is irrelevant for my discussion.

<sup>49</sup> The term 'God' thus must be defined as entailing a first cause.

God with ‘There is no first cause’, this would be cold comfort for someone who expects God to be benevolent and all-knowing.

Claim 18 also establishes that the empirical significance of  $\sigma$  is independent of the background assumptions, or rather, since  $\sigma$  determines the background assumptions  $C_{\mathcal{B}}(\sigma)$ ,  $\sigma$  itself alone already determines whether it is significant. Thus in spite of the background assumptions being the analytic component of a known (or conjectured) theory, one can make definitive, non-preliminary claims about  $\sigma$ ’s significance.

The price to pay for the connection between content, background assumptions, and empirical significance is, as pointed out, that the  $\mathcal{B}$ -sentences must not be restricted to conjunctions of literals, but must include unrestrictedly quantified sentences. This solution should not have been anathema to Hempel, since he himself uses type theory to arrive at explicit definitions for real-valued measurement results in observational terms (Hempel 1958, pp. 64–67). He notes that this “price will be considered too high” by some, but adds that “it would no doubt be generally considered a worthwhile advance in clarification if for a set of theoretical scientific expressions explicit definitions in terms of observables can be constructed at all” (Hempel 1958, pp. 65–66). Similarly, having a criterion of empirical significance (and empirical content) at all should be worth the price of extending one’s basic language as well.

Nonetheless, one may hope for a more general account of the empirical significance (and the content) of a sentence  $\sigma$ , one that allows placing more restrictions on the sentences in  $\mathcal{B}$ . For this aim, one could rely on a new definition of the content  $C(\sigma)$  of sentence  $\sigma$  in  $\mathcal{B}$  relative to  $\vartheta$ , defining it as the set<sup>50</sup>

$$\{\rho \mid \sigma \wedge \vartheta \vdash \rho \text{ and } \rho \in \mathcal{B}\} . \tag{10}$$

The problem is that  $C(\sigma)$  may be an infinite set, and thus cannot be used as the antecedent of an implication in the same way that the Ramsey sentence is used in the Carnap sentence. This problem can be solved by treating the background assumptions not as sets of sentences but rather as inference rules, as Carnap often did.<sup>51</sup> Then one can define the analytic component of  $\sigma$  as the inference rule

$$C(\sigma) \vdash \sigma . \tag{11}$$

The resulting inference system is not particularly graceful, for instance because it depends on the sentence  $\sigma$ , so that it differs from sentence to sentence. But in this inference system,

$$\sigma \vdash C(\sigma) , \tag{12}$$

which can be defined as the translatability of  $\sigma$  into a set of  $\mathcal{B}$ -sentences, and which is non-trivial if and only if  $C(\sigma)$  is neither contradictory nor tautologous. Corre-

<sup>50</sup> The discussion here and in the following ignores problems stemming from the incompleteness of deduction in higher order logic for Tarski semantics. One could avoid any such problems by switching either to semantic entailment or to Henkin semantics.

<sup>51</sup> This strategy is not that uncommon, for example when  $\ulcorner x \neq y \urcorner$  is defined in the metalanguage of a logic as an abbreviation for  $\ulcorner \neg x = y \urcorner$ , without introducing a corresponding axiom into the language.

spondingly, one can define  $\sigma$  to be (non-trivially) verifiable if and only if there is a consistent set  $\Omega$  of  $\mathcal{B}$ -sentences such that  $\Omega \vdash \sigma$  (and  $\not\vdash \sigma$ ). One can define  $\sigma$  to be (non-trivially) falsifiable if and only if there is a consistent set  $\Omega$  of  $\mathcal{B}$ -sentences such that  $\Omega \vdash \neg\sigma$  (and  $\not\vdash \neg\sigma$ ). As with the Carnap sentence, non-trivial verifiability, falsifiability, and translatability then are equivalent given the analytic component of  $\sigma$ . This solution allows more freedom with respect to the choice of  $\mathcal{B}$ , although it still does not establish the equivalence for verifiability, falsifiability, and translatability in the sense of Carnap's early works, where Carnap relied on single conjunctions of literals as  $\mathcal{B}$ -sentences.

## 6.2 Metaphysics and the criteria for terms

Maybe the most impressive aspect of the Carnap sentence is that it ensures the translatability of *any* theory. Carnap's only previous attempt at establishing the translatability of a theory was in the *Aufbau*. Like the Carnap sentence approach, the *Aufbau* relies on the very inclusive notion of  $\mathcal{B}$  that only restricts the terms that  $\mathcal{B}$  may contain, not the logical structure of  $\mathcal{B}$ -sentences. But the *Aufbau* places enormous restrictions on the logical form of scientific theories, since it requires the definability of all  $\mathcal{A}$ -terms in  $\mathcal{B}$ -terms. By contrast, the Carnap sentence approach places no restrictions whatsoever on a theory  $\sigma$ . However, this very flexibility is also the most threatening aspect for Carnap's original program, the criticism of metaphysics:  $\sigma$  may be any sentence, and thus also a sentence 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:  $Pa$ ) contains no  $\mathcal{B}$ -terms, and thus its Ramsey sentence is the tautology  $\exists Y \exists x Yx$ . The sentence's verifiability and translatability are thus trivial: It is an analytic truth, and thus trivially verified by and trivially 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  $\mathcal{B}$ -sentences, they may not be speaking nonsense, but also make no claims about the world. They are engaged in language choice. Hence Carnap's criticism of metaphysics only has to be reformulated. Metaphysical sentences are problematic, as Carnap originally claimed, but not because they are meaningless, but because they are pseudo-synthetic (cf. [Diamond 1975](#), pp. 16–20).

Thus it seems that Carnap's system has a place for all sentences, metaphysical ones and also sentences about values: If the latter connect to value experiences and thus to emotions as in the *Aufbau*, they amount to empirical statements. If they have no contact at all to  $\mathcal{B}$ -sentences, they are analytic.<sup>52</sup> Given its inclusiveness, the different criteria of empirical significance for terms may find their place in Carnap's system as well. For even if some non-significant sentence contains only terms that fulfill one of the different criteria, since that sentence is simply analytic, its terms do not have to be treated as meaningless, but rather as conventionally chosen. Thus

<sup>52</sup> Thus it seems that "pure optatives" (proposals, requests, demands, etc. that do not assert any matters of fact) are analytic, contrary to Carnap's claims (cf. [Carnap 1963b](#), p. 32.B).

one obvious role for the criteria is the identification of terms that can be chosen for specific roles. Introducible terms can be used for identifying 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. Terms that are significant according to other criteria may identify specific groups of properties.<sup>53</sup> Thus in spite of their questionable help in identifying significant sentences 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 is 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 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.

### 7 Further proofs

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

*Proof*

$$\begin{aligned}
 P(\sigma \mid \vartheta) > P(\sigma \mid \beta \wedge \vartheta) &= \frac{P(\beta \mid \sigma \wedge \vartheta)P(\sigma \mid \vartheta)}{P(\beta \mid \vartheta)} = \frac{1 - P(\neg\beta \mid \sigma \wedge \vartheta)P(\sigma \mid \vartheta)}{1 - P(\neg\beta \mid \vartheta)} \\
 &\Leftrightarrow P(\neg\beta \mid \vartheta) < P(\neg\beta \mid \sigma \wedge \vartheta) \\
 &\Leftrightarrow P(\sigma \mid \vartheta) < \frac{P(\neg\beta \mid \sigma \wedge \vartheta)P(\sigma \mid \vartheta)}{P(\neg\beta \mid \vartheta)} = P(\sigma \mid \neg\beta \wedge \vartheta)
 \end{aligned}
 \tag{13}$$

□

**Claim 11** *For some sentences  $\vartheta$  there are simply existentially quantified sentences  $\sigma_{\exists}$  and simply universally quantified sentences  $\sigma_{\forall}$  such that all terms of  $\sigma_{\exists}$  and  $\sigma_{\forall}$  are introducible by reduction pairs from  $\vartheta$ , but  $\sigma_{\exists}$  and  $\sigma_{\forall}$  are neither verifiable nor falsifiable relative to  $\vartheta$ .*

*Proof* Choose  $\vartheta = \forall x[B_1x \rightarrow (B_2x \leftrightarrow A_1x)] \wedge \forall x[\neg B_1x \rightarrow (B_2x \leftrightarrow A_2x)]$ . Then  $\sigma_{\exists} = \exists x(A_1x \wedge A_2x)$  is not falsifiable relative to  $\vartheta$ :

$$\begin{aligned}
 &\vdash \forall x[B_1x \rightarrow (B_2x \leftrightarrow \lambda y[B_2y \vee \neg B_1y]x)] \\
 &\quad \wedge \forall x[\neg B_1x \rightarrow (B_2x \leftrightarrow \lambda y[B_2y \vee B_1y]x)] \\
 &\quad \wedge \exists x[\lambda y(B_2y \vee \neg B_1y)x \wedge \lambda y(B_2y \vee B_1y)x] \\
 &\vdash \exists X_1 X_2 (\forall x[B_1x \rightarrow (B_2x \leftrightarrow X_1x)])
 \end{aligned}
 \tag{14a}$$

<sup>53</sup> The recursion base of the criterion by Creath (1976) may fit the bill.



$$\begin{aligned} & \wedge \forall x[\neg B_1x \rightarrow (B_2x \leftrightarrow X_2x)] \\ & \wedge \exists x[X_1x \wedge X_2x] \end{aligned} \tag{14b}$$

$$\vdash \mathbf{R}_{\mathcal{B}}(\vartheta \wedge \sigma_{\exists}) \tag{14c}$$

Since the Ramsey sentence of  $\sigma_{\exists} \wedge \vartheta$  entails the same  $\mathcal{B}$ -sentences as  $\sigma_{\exists} \wedge \vartheta$ ,  $\sigma_{\exists} \wedge \vartheta$  and specifically  $\sigma_{\exists}$  does not entail any  $\mathcal{B}$ -sentences not already entailed by  $\vartheta$  alone. Since a sentence is verifiable if and only if its negation is falsifiable (Hempel 1950, p. 48), a similar argument can be used to show that  $\sigma_{\exists}$  is also not verifiable relative to  $\vartheta$ :

$$\begin{aligned} & \vdash \forall x[B_1x \rightarrow (B_2x \leftrightarrow \lambda y[B_2y \wedge B_1y]x)] \\ & \wedge \forall x[\neg B_1x \rightarrow (B_2x \leftrightarrow \lambda y[B_2y \wedge \neg B_1y]x)] \\ & \wedge \neg \exists x[\lambda y(B_2y \wedge B_1y)x \wedge \lambda y(B_2y \wedge \neg B_1y)x] \end{aligned} \tag{15a}$$

$$\begin{aligned} & \vdash \exists X_1X_2(\forall x[B_1x \rightarrow (B_2x \leftrightarrow X_1x)] \\ & \wedge \forall x[\neg B_1x \rightarrow (B_2x \leftrightarrow X_2x)] \\ & \wedge \neg \exists x[X_1x \wedge X_2x]) \end{aligned} \tag{15b}$$

$$\vdash \mathbf{R}_{\mathcal{B}}(\vartheta \wedge \neg \sigma_{\exists}) \tag{15c}$$

Similarly, it can be shown that  $\sigma_{\forall} = \forall x(A_1x \leftrightarrow A_2x)$  is not falsifiable relative to  $\vartheta$ :

$$\begin{aligned} & \vdash \forall x[B_1x \rightarrow (B_2x \leftrightarrow B_2x)] \wedge \forall x[\neg B_1x \rightarrow (B_2x \leftrightarrow B_2x)] \\ & \wedge \forall x(B_2x \leftrightarrow B_2x) \end{aligned} \tag{16a}$$

$$\begin{aligned} & \vdash \exists X_1X_2(\forall x[B_1x \rightarrow (B_2x \leftrightarrow X_1x)] \wedge \forall x[\neg B_1x \rightarrow (B_2x \leftrightarrow X_2x)] \\ & \wedge \forall x[X_1x \leftrightarrow X_2x]) \end{aligned} \tag{16b}$$

$$\vdash \mathbf{R}_{\mathcal{B}}(\vartheta \wedge \sigma_{\forall}) \tag{16c}$$

And it can be shown that  $\sigma_{\forall}$  is not verifiable relative to  $\vartheta$ :

$$\begin{aligned} & \vdash \forall x[B_1x \rightarrow (B_2x \leftrightarrow \lambda y[(B_2y \wedge B_1y) \vee (\neg B_2y \wedge \neg B_1y)]x)] \\ & \wedge \forall x[\neg B_1x \rightarrow (B_2x \leftrightarrow \lambda y[(B_2y \wedge \neg B_1y) \vee (\neg B_2y \wedge B_1y)]x)] \\ & \wedge \neg \forall x(\lambda y[(B_2y \wedge B_1y) \vee (\neg B_2y \wedge \neg B_1y)]x \\ & \quad \leftrightarrow \lambda y[(B_2y \wedge \neg B_1y) \vee (\neg B_2y \wedge B_1y)]x) \end{aligned} \tag{17a}$$

$$\begin{aligned} & \vdash \exists X_1X_2(\forall x[B_1x \rightarrow (B_2x \leftrightarrow X_1x)] \\ & \wedge \forall x[\neg B_1x \rightarrow (B_2x \leftrightarrow X_2x)] \\ & \wedge \neg \forall x[X_1x \leftrightarrow X_2x]) \end{aligned} \tag{17b}$$

$$\vdash \mathbf{R}_{\mathcal{B}}(\vartheta \wedge \neg \sigma_{\forall}) \tag{17c}$$

□

**Acknowledgements** 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* under the title “The Criteria for the Empirical Significance of Terms”. I thank the audiences for helpful comments and discussions, and I especially thank two anonymous reviewers for a multitude of excellent comments, corrections, and suggestions. Research for this article was supported by the Alexander von Humboldt Foundation.

## References

- Adams, C. (2011a). Is there a God? Archived by WebCite® at <http://www.webcitation.org/6PC0xRqj6>, *Chicago Reader*, November 11. <http://www.straightdope.com/columns/read/3021/is-there-a-god>.
- Adams, C. (2011b). Is there a God (revisited)? Archived by WebCite® at <http://www.webcitation.org/6PCI3RkoE>, *Chicago Reader*, December 2. <http://www.straightdope.com/columns/read/3025/is-there-a-god-revisited>.
- Ayer, A. J. (1936). *Language, truth and logic* (1st ed.). References are to the second edition (Ayer 1946). London: Victor Gollanz.
- Ayer, A. J. (1946). *Language, truth and logic* (2nd ed.). London: Victor Gollanz.
- Ayer, A. J. (Ed.). (1959). *Logical positivism*. New York: The Free Press.
- Belnap, N. (1993). On rigorous definitions. *Philosophical Studies*, 72, 115–146.
- Benson, A. J. (1963). Bibliography of the writings of Rudolf Carnap. In Schilpp 1963 (pp. 1016–1070).
- Carnap, R. (1928a). *Der logische Aufbau der Welt*. References are to the translation (Carnap 1967a). Berlin-Schlachtensee: Weltkreis-Verlag.
- Carnap, R. (1928b). *Scheinprobleme in der Philosophie: Das Fremdpsychische und der Realismusstreit*. References are to the translation (Carnap 1967a). Berlin-Schlachtensee: Weltkreis-Verlag.
- Carnap, R. (1931a). Die physikalische Sprache als Universalsprache der Wissenschaft. Nominal publication date incorrect. Published in 1932. *Erkenntnis*, 2(5/6), 432–465.
- Carnap, R. (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).
- Carnap, R. (1932). Über Protokollsätze. References are to the translation (Carnap 1987). *Erkenntnis*, 3(1), 215–228.
- Carnap, R. (1932/1933). Psychologie in physikalischer Sprache. References are to the translation (Carnap 1959a). *Erkenntnis*, 3(1), 107–142.
- Carnap, R. (1934). *Logische Syntax der Sprache* (Vol. 8). Schriften zur wissenschaftlichen Weltauffassung. References are to the corrected translation (Carnap 1967b). Wien: Springer-Verlag.
- Carnap, R. (1935). *Philosophy and logical syntax*. Psyche Miniatures, General Series 70. London: Kegan Paul, Trench, Trubner & Co.
- Carnap, R. (1936). Testability and meaning. *Philosophy of Science*, 3(4), 420–468.
- Carnap, R. (1937). Testability and meaning—Continued. *Philosophy of Science*, 4(1), 2–35.
- Carnap, R. (1938). Empiricism and the language of science. *Synthese*, 3(1).
- Carnap, R. (1939). *Foundations of logic and mathematics* (Vol. I, No. 3, pp. 139–213). Foundations of the unity of science. Toward an International Encyclopedia of Unified Science. References are to the two-volume edition. Chicago: University of Chicago Press.
- Carnap, R. (1950a). *Logical foundations of probability*. References are to the second edition (Carnap 1962). Chicago: University of Chicago Press.
- Carnap, R. (1950b). *Testability and meaning*. New Haven, CT: Yale University.
- Carnap, R. (1956). The methodological character of theoretical concepts. In H. Feigl & M. Scriven (Eds.), *The foundations of science and the concepts of psychology and psychoanalysis* (Vol. 1, pp. 38–75). Minnesota Studies in the Philosophy of Science. Minneapolis, MN: University of Minnesota Press.
- Carnap, R. (1958). Beobachtungssprache und theoretische Sprache. *Dialectica*, 12(3–4), 236–248.
- Carnap, R. (1959a). Psychology in physical language. Translated by G. Schick. In A. J. Ayer (Ed.), *Logical positivism* (pp. 165–198). New York: Free Press.
- Carnap, R. (1959b). The elimination of metaphysics through logical analysis of language. Translated by A. Pap. In A. J. Ayer (Ed.), *Logical positivism* (pp. 60–81). New York: The Free Press.
- Carnap, R. (1962). *Logical foundations of probability* (2nd ed.). Chicago: University of Chicago Press.
- Carnap, R. (1963a). Intellectual autobiography. In Schilpp 1963 (pp. 1–84). La Salle, IL: The Open Court.
- Carnap, R. (1963b). Replies and systematic expositions. In Schilpp 1963 (pp. 859–1016). La Salle, IL: The Open Court.

- Carnap, R. (1967a). *The logical structure of the world. Pseudoproblems of philosophy*. Translated by R. A. George. Berkeley, CA: University of California Press.
- Carnap, R. (1967b). *The logical syntax of language*. Translated by A. Smeaton. Reprinted with corrections. London: Routledge & Kegan Paul Ltd.
- Carnap, R. (1987). On protocol sentences. Translated by R. Creath & R. Nollan. *Noûs*, 21(4), 457–470.
- Church, A. (1949). Review: Alfred Jules Ayer, *Language, truth, and logic* (2nd ed.). *The Journal of Symbolic Logic*, 14(1), 52–53.
- Creath, Richard. (1976). On Kaplan on Carnap on significance. *Philosophical Studies*, 30, 393–400.
- Demopoulos, W. (2007). Carnap on the rational reconstruction of scientific theories. In M. Friedman & R. Creath (Eds.), *The Cambridge companion to Carnap* (pp. 248–272). Cambridge: Cambridge University Press.
- Diamond, M. L. (1975). The challenge of contemporary empiricism. *Diamond and Litzenburg*, 1975, 3–62.
- Diamond, M. L., & Litzenburg, T. V. (Eds.). (1975). *The logic of God: Theology and verification*. Indianapolis, IN: Bobbs-Merill.
- Essler, W. K. (1982). Wissenschaftstheorie I. Definition und Reduktion. In *Kolleg Philosophie* (2nd ed.). München: Verlag Karl Alber.
- Flew, A. (1950). Theology and falsification. References are to the reprint (Diamond and Litzenburg, 1975, 257–259). *University*, 1, 1–8.
- Friedl, J., & Rutte, H. (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, N. (1965). *Fact, fiction, and forecast* (2nd ed.). Indianapolis, IN: Bobbs-Merrill.
- Gupta, A. (2009). Definitions. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*, Spring 2009. Stanford, CA: The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University.
- Heidegger, M. (1931). *Was ist Metaphysik?*. Bonn: Verlag von Friedrich Cohen.
- Hempel, C. G. (1945). Studies in the logic of confirmation (II.). *Mind*, 54(214), 97–121.
- Hempel, C. G. (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.
- Hempel, C. G. (1958). The theoretician's dilemma. In: H. Feigl, M. Scriven, & G. Maxwell (Eds.), *Concepts, theories, and the mind-body problem* (Vol. 2, pp. 173–226). Minnesota Studies in the Philosophy of Science. Minneapolis, MN: University of Minnesota Press.
- Hodges, W. (1993). Model theory. In *Encyclopedia of mathematics and its applications* (Vol. 42). Digitally printed in 2008. Cambridge: Cambridge University Press.
- Howson, C., & Urbach, P. (1993). *Scientific reasoning: The Bayesian approach* (2nd ed.). Second printing 1996. Chicago: Open Court.
- Justus, J. (forthcoming). Carnap's forgotten criterion of empirical significance. *Mind*. doi:10.1093/mind/fzu042.
- Justus, J. (2011). Evidentiary inference in evolutionary biology. Review of (Sober 2008). *Biology and Philosophy*, 26, 419–437. doi:10.1007/s10539-010-9205-7.
- Kaplan, D. (1975). Significance and analyticity: A comment on some recent proposals of Carnap. In J. Hintikka (Ed.), *Rudolph Carnap, logical empiricist: Materials and perspectives*. Dordrecht: D. Reidel.
- Lewis, D. (1988a). Ayer's first empiricist criterion of meaning: Why does it fail? *Analysis*, 48(1), 1–3.
- Lewis, D. (1988b). Statements partly about observation. References are to the reprint (Lewis 1998, 125–155). *Philosophical Papers*, 17, 1–31.
- Lewis, D. (1998). *Papers in philosophical logic*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press.
- Lutz, S. (2012). Criteria of empirical significance: foundations, relations, applications. PhD dissertation, Utrecht University. <http://philsci-archive.pitt.edu/id/eprint/9117>.
- Mackie, J. L. (1980). *The cement of the universe. A study of causation* (2nd ed.). Reprinted 2002. Oxford: Clarendon Press.
- Malzkorn, W. (2001). Defining disposition concepts: A brief history of the problem. *Studies in History and Philosophy of Science Part A*, 32(2), 335–353.
- Marhenke, P. (1949/1950). The criterion of significance. *Proceedings and addresses of the American Philosophical Association*, 23, 1–21.

- Mormann, T. (2007). Carnap's logical empiricism, values, and American pragmatism. *Journal for General Philosophy of Science*, 38, 127–146. doi:10.1007/s10838-007-9033-x.
- Neurath, O. (1932). Protokollsätze. *Erkenntnis*, 3(1), 204–214.
- Pokriefka, M. L. (1983). Ayer's definition of empirical significance revisited. *Analysis*, 43(4), 166–170.
- Pokriefka, M. L. (1984). More on empirical significance. *Analysis*, 44(2), 92–93.
- Przełęcki, M. (1969). *The logic of empirical theories*. Monographs in Modern Logic Series. London/New York: Routledge & Kegan Paul/Humanities Press.
- Psillos, S. (2000a). Carnap, the Ramsey-sentence and realistic empiricism. *Erkenntnis*, 52, 253–279. doi:10.1023/a:1005589117656.
- Psillos, S. (2000b). Rudolf Carnap's theoretical concepts in science. *Studies in History and Philosophy of Science Part A*, 31(1), 151–172. doi:10.1016/s0039-3681(99)00031-x.
- Reichenbach, H. (1951). The verifiability theory of meaning. *Proceedings of the American Academy of Arts and Sciences*, 80(1), 46–60.
- Rozeboom, W. W. (1962). The factual content of theoretical concepts. In H. Feigl, M. Scriven, & G. Maxwell (Eds.), *Scientific explanation, space, and time* (Vol. 3, pp. 273–357). Minnesota Studies in the Philosophy of Science. Minneapolis, MN: University of Minnesota Press.
- Ruja, H. (1961). The present status of the verifiability criterion. *Philosophy and Phenomenological Research*, 22(2), 216–222.
- Schilpp, P. A. (Ed.). (1963). The philosophy of Rudolf Carnap. In *The library of living philosophers* (Vol. 11). Chicago: Open Court Publishing Company.
- Schurz, G. (2014). *Philosophy of science. A unified approach*. New York: Routledge.
- Skyrms, B. (1984). *Prediction and experience*. New Haven, CT: Yale University Press.
- Soames, S. (2003). The dawn of analysis. In *Philosophical analysis in the twentieth century* (Vol. 1). Princeton, NJ: Princeton University Press.
- Sober, E. (2008). *Evidence and evolution: The logic behind the science*. Cambridge: Cambridge University Press.
- Van Cleve, J. (1971). Empiricist criteria of meaningfulness. Unpublished, cited by Creath (1976).
- Wagner, P. (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.
- Winnie, J. A. (1970). Theoretical analyticity. In R. C. Buck & R. S. Cohen (Eds.), *In Memory of Rudolf Carnap: Proceedings of the biennial meeting of the Philosophy of Science Association* (Vol. VIII, pp. 289–305). Boston Studies in the Philosophy of Science. Dordrecht: D. Reidel Publishing Company.
- Wright, C. (1986). Scientific realism, observation and the verification principle. In G. Macdonald & C. Wright (Eds.), *Fact, science and morality: Essays on A. J. Ayer's language, truth and logic* (pp. 247–274). Oxford: Basil Blackwell.
- Wright, C. (1989). The verification principle: Another puncture—another patch. *Mind*, 98(392), 611–622.
- Yi, B.-U. (2001). Compact entailment and Wright's verification principle. *Mind*, 110(437), 414–421.