THE REVIEW OF SYMBOLIC LOGIC
Volume 12, Number 2, June 2018

CARNAP'S DEFENSE OF IMPREDICATIVE DEFINITIONS

VERA FLOCKE
New York University

Abstract. A definition of a property P is *impredicative* if it quantifies over a domain to which P belongs. Due to influential arguments by Ramsey and Gödel, impredicative mathematics is often thought to possess special metaphysical commitments. The reason is that an impredicative definition of a property P does not have its intended meaning unless P exists, suggesting that the existence of P cannot depend on its explicit definition. Carnap (1937 [1934], p. 164) argues, however, that accepting impredicative definitions amounts to choosing a "form of language" and is free from metaphysical implications. This article explains this view in its historical context. I discuss the development of Carnap's thought on the foundations of mathematics from the mid-1920s to the mid-1930s, concluding with an account of Carnap's (1937 [1934]) non-Platonistic defense of impredicativity. This discussion is also important for understanding Carnap's influential views on ontology more generally, since Carnap's (1937 [1934]) view, according to which accepting impredicative definitions amounts to choosing a "form of language", is an early precursor of the view that Carnap presents in "Empiricism, Semantics and Ontology" (1956 [1950]), according to which referring to abstract entities amounts to accepting a "linguistic framework".

§1. Introduction. Impredicative definitions in some way invoke, reference or mention in the defining clause what they define. For example, a definition of a property P that uses a quantifier that ranges over a domain to which P belongs is impredicative. The notion of an impredicative definition is of central importance in mathematics and logic. For example, the definition of the property *being the least upper bound of a bounded class of real numbers* is impredicative. The reason why this definition is impredicative has to do with Dedekind's definition of the real numbers, according to which real numbers can be represented as downward closed sets of rational numbers with no largest element.[1] The least upper bound of a bounded class C of sets is the union of all elements in C. Given Dedekind's definition, the least upper bound of a bounded class C of reals is the set that contains every rational number q which is an element of one of the elements P of C, and the property of *being the least upper bound of a bounded class* C *of real numbers* can be defined as follows:

Received: September 6, 2017.
2010 *Mathematics Subject Classification*: 00A30, 01A60, 03B15.
Key words and phrases: Carnap, principle of tolerance, impredicative definitions, mathematical ontology, metaontology, philosophy of mathematics.

[1] In more detail, a *Dedekind cut* is a partition of the rational numbers into two sets A and B, such that A is downward closed, B is upward closed, every element of A is smaller than any element of B and A has no largest element. Dedekind (1963 [1872], sec. IV) represented real numbers as Dedekind cuts. Since one of the sets that is part of a Dedekind cut uniquely determines the other, we can more simply represent real numbers as downward closed sets of rational numbers with no largest element.

doi:10.1017/S1755020318000370

(1) $\lambda x.[\forall q(q \in x \leftrightarrow \exists P(P \in C \wedge q \in P))]$.[2]

Real numbers have the least upper bound property, meaning that any nonempty and bounded set of real numbers has a least upper bound in the real numbers.[3] The existential quantifier $\exists$ that appears in formula (1) hence binds a variable P that ranges over the property that this formula defines as one of its values.

Impredicative mathematics is often thought to possess special metaphysical commitments. An impredicative definition of a property P in a sense refers to the property that it defines. It hence seems that this property P cannot depend for its existence on its explicit definition, since this definition does not have its intended meaning unless P already exists. Many philosophers, from Ramsey (1931 [1926], p. 41) through Gödel (1984 [1944], p. 456) to scholars working today (Linnebo, 2004, pp. 155–156), therefore think that impredicative definitions are legitimate only if the defined properties can also be defined predicatively, or else exist independently of their explicit definition.

Carnap argues, however, that proponents of impredicative mathematics need not answer metaphysical questions. He puts forward the so-called principle of tolerance,[4] as formulated in the following passage:

> "*In logic, there are no morals.* Everyone is at liberty to build up his own logic, i.e., his own form of language, as he wishes. All that is required is that, if he wishes to discuss it, he must state his methods clearly, and give syntactical rules instead of philosophical arguments". (Carnap, 1937 [1934], p. 52).[5]

According to this principle, a choice between alternative logical systems does not require philosophical justification. Carnap directly applies this principle to the question of whether impredicative definitions are permissible, in the following passage:

> "The proper way of framing the question is not 'Are indefinite (or impredicative) symbols admissible?' for, since there are no morals in logic [...] what meaning can 'admissible' have here? The problem can only be expressed in this way: 'How shall we construct a particular language? Shall we admit symbols of this kind or not? And what are the consequences of either procedure?' It is therefore a question of choosing a form of language—that is, of the establishment of rules of syntax and the investigation of the consequences of these". (Carnap, 1937 [1934], p. 164).

[2] For maximum clarity: q is a variable that ranges over rational numbers, P is a variable that ranges over real numbers (which are represented as sets of rationals), and C is a constant that refers to a bounded class of real numbers.

[3] This is important because the real numbers, unlike the rational numbers, form a continuous line with no gaps.

[4] There is a disagreement in the literature on whether Carnap proclaims his principle of tolerance as a thesis, that would be true or false, or rather puts it forth as a methodological recommendation that wouldn't be true or false. Coffa (1993, p. 325) endorses the former view, while Goldfarb (1997, p. 61) endorses the latter. My choice of words, "putting forward", signals that I am following Goldfarb's line of interpretation.

[5] Carnap presumably intends this principle to apply to male and female logicians alike.

Carnap's principle of tolerance supposedly yields that accepting impredicative definitions amounts to choosing a "form of language", which is a choice that, according to Carnap, does not require the justification of a metaphysical viewpoint.

Carnap's principle of tolerance raises many questions. What sort of "methods" does Carnap have in mind? What are "syntactical rules"? Why should the choice of a system of logic or mathematics not require philosophical justification? Carnap is often understood as someone who dogmatically rejects the engagement with philosophical questions. On this interpretation, Carnap, in particular, simply refuses to answer metaphysical questions relating to the admissibility of impredicative definitions. However, *The Logical Syntax of Language* does contain a further going explanation of why proponents of impredicative mathematics are not required to answer metaphysical questions, which provides the rationale for a specific application of Carnap's principle of tolerance. The goal of this article is to explain this account in its historical context.

In more detail, Carnap provided not just one but two non-Platonistic defenses of the admissibility of using impredicative definitions in mathematics, one in the 1931 article "The Logicist Foundations of Mathematics" and one in *The Logical Syntax of Language* (1937 [1934]). I discuss the historical development of Carnap's thought on the foundations of mathematics and problems surrounding impredicativity from the mid-1920s to the mid-1930s, and explain, first, Carnap's 1931 and then his 1937 [1934] defense of the admissibility of impredicative definitions.

The discussion in this article constitutes a step in a larger research project that aims at explaining Carnap's influential views on ontology. Carnap's claim that accepting impredicative mathematics, i.e., mathematics that makes use of impredicative definitions, amounts to "choosing a form of language" (1937 [1934], p. 164) is an early precursor of a claim he later made in the famous article "Empiricism, Semantics and Ontology" (1956 [1950]), according to which referring to abstract entities amounts to accepting a "linguistic framework". Carnap's notion of a framework, and the related distinction between "internal" and "external" questions, thus appears to be a direct outcome of his reflections on the foundation of mathematics. Understanding Carnap's views on the foundations of mathematics therefore helps to understand his views on ontology more generally. In other work, I provide a detailed interpretation of "Empiricism, Semantics and Ontology" and the 1956 [1950] distinction between "internal" and "external" questions.[6]

My discussion will be structured as follows. I begin by explaining why impredicative definitions may seem to possess metaphysical implications (§2), and explain Carnap's pre-*Syntax* philosophy of mathematics in the late 1920s and early 1930s (§3). This exposition will allow me to explain Carnap's 1931 defense of impredicative definitions (§4). Carnap's views changed dramatically after he learned of Tarski's meta-mathematical approach and Gödel's incompleteness theorems in 1930. I go on to explain his later syntactic approach in *The Logical Syntax of Language* (1937 [1934]) (§5), provide an account of Carnap's distinction between syntax and semantics (§6), and then explain his 1937 [1934] defense of impredicative definitions (§7). I conclude by discussing the direct influence that Gödel had on the eventual shape of Carnap's view (§8).

§2. Impredicative definitions. From Carnap's viewpoint in the early 1930s, the main system of predicative mathematics was ramified type theory as proposed by Russell &

6 See Flocke (forthcoming).

https://doi.org/10.1017/S1755020318000370 Published online by Cambridge University Press

Whitehead (1927 [1910]), and the main system of impredicative mathematics was simple type theory as proposed by Ramsey (1931 [1926])). At that time, the decision between simple and ramified type theory seemed to involve thorny metaphysical questions. I will begin by explaining why the decision between simple and ramified type theory may seem to have metaphysical implications, and then go on to explain Carnap's attempt at avoiding these implications.

Type theories are formal systems that contain variables and quantifiers of different syntactic categories, the so-called *types*. Formal systems of this kind were originally developed with the goal of providing a solution to certain paradoxes, including Russell's paradox. The main idea of type theories to solve this paradox is that sentences such as '$A \in A$' (where A is a set) violate type restrictions on which terms can combine to form meaningful expressions, and are therefore meaningless. The main difference between simple and ramified type theories is that they assign types to terms in different ways. In short, simple type theory includes just one type of propositions and assigns types to propositional functions depending on only the free variables that they contain. Ramified type theories, by contrast, include several types of propositions and assign both a *type* and an *order* to propositional functions depending on the free and the bound variables that they contain. Functions of a given type hence divide into various orders.

In more detail, versions of both kinds of type theory typically require that all variables are assigned a unique type.[7] Simple type theories then distinguish between a type i of individuals, a type $\langle\rangle$ of propositions, and various types of propositional functions, so that $\langle t_1, \ldots, t_n \rangle$ is the type of an n-place propositional function taking arguments of type t_i at the ith argument place.[8] Ramified type theories, by contrast, require that propositions and propositional functions of a given type divide into various *orders*. Thus, $\langle m \rangle$ is the type of propositions of order m, and types of propositional functions are ordered $n + 1$-tuples of the form $\langle t_1, \ldots, t_n, m \rangle$, where $t_1, \ldots, t_n$ track the types of the arguments, and m is used to divide functions of a given type into orders. The order m of a proposition depends on the bound variables it contains, and the order m of a function depends on both its free and its bound variables. Ramified type theories then require that quantifiers that appear in a given propositional function bind only variables whose type ends with a number lower than the number with which the type of the function ends.[9]

The respective type assignments have an effect on how simple and ramified type theories treat impredicative definitions: simple type theorists accept impredicative definitions, but ramified type theorists reject them. According to a simple type theorist, the quantifiers and variables in formula (1) need to be indexed to specific types, as follows:

(1') $\lambda x_{\langle i \rangle}.[\forall_i q_i (q_i \in x_{\langle i \rangle} \leftrightarrow \exists_{\langle i \rangle} P_{\langle i \rangle} (P_{\langle i \rangle} \in C \wedge q_i \in P_{\langle i \rangle}))].$

Formula (1′) is unproblematic from the viewpoint of simple type theory. A ramified type theorist would insist, however, that the bound variable $P_{\langle i \rangle}$ and the free variable $x_{\langle i \rangle}$

---

7 Church (1940) initiated the approach that associates each variable with a unique type. A different approach, due to Curry, associates each free variable with a unique type but assigns types to bound variables "in situ", in the formulas in which they appear (see Curry, Feys, & Craig (1958), pp. 315–343).

8 Simple type theory was first proposed by Chwistek (2012 [1922]) and Ramsey (1931 [1926]), and then formalized using the λ-calculus by Church (1940). For contemporary versions of simple type theory see Hindley (1997) and Hodes (2015, sec. 1).

9 For contemporary versions of ramified type theory, see Hodes (2013) and (2015, sec. 6). Thanks to Harold Hodes for helpful comments on my presentation of the two kinds of type theory.

must be assigned different indices, and $x_{\langle i \rangle}$ assigned a higher order than $P_{\langle i \rangle}$, since the class that is defined by (1′) cannot be a value of the bound variable that appears in its definition.

This observation reveals a problem for ramified type theory, since it shows that the least upper bound of a class C of reals is of an order higher than the order of C. The real numbers hence split into various orders, even though reals, in contrast to rationals, are supposed to form a continuous line with no gaps. In the first edition of *Principia Mathematica*, Russell and Whitehead proposed to solve this problem by introducing the *Axiom of Reducibility*, according to which for every propositional function there is an equivalent predicative function (see section $*12$ of *Principia Mathematica*). A function of type $\langle t_1, \ldots, t_n, m\rangle$ is predicative iff $m = t_i + 1$, where t_i is the highest type of one of $t_1, \ldots, t_n$. This axiom essentially collapses the previously induced stratification of propositional functions. However, various philosophers (including Carnap, 1931, p. 97, and Ramsey, 1931 [1926], p. 57) criticized the axiom of reducibility as an *ad hoc* addition to the theory, and as being neither a tautology nor logically necessary.[10]

More recent developments in predicative mathematics have shown that large parts of mathematics, including a theory of real numbers based on a Dedekindian definition can after all be recovered on predicative grounds.[11] Some of the relevant results were in principle available already in the 1920s. In particular, Weyl (1918) argued that even though a proponent of predicative mathematics cannot derive the least upper bound property for *sets* of real numbers, a derivation that makes use of only predicative means is possible for *sequences* of real numbers.[12] However, it seems that, in the 1920s, Carnap was not aware of Weyl's work.[13] From his viewpoint at the time, ramified type theory presents a dilemma "between the Scylla that is the axiom of reducibility and the Charybdis that is the division of the real numbers into various orders" (1931, p. 104, my translation). Simple type theory, in contrast to ramified type theory, avoids this dilemma since it imposes no restriction on the bound variables that appear in functional terms. Carnap therefore preferred simple over ramified type theory.[14]

However, simple type theory also runs into certain difficulties. Russell's and Whitehead's official reason for ramification[15] is that Russell's paradox (and other paradoxes) arises from a violation of the so-called vicious circle principle.[16] They offer three different formulations of this principle (see Russell & Whitehead (1927 [1910], p. 37)):[17]

---

10 For a more extensive discussion of the axiom of reducibility and of various objections to it, see Linsky (1999, chap. 6). For a defense, see Myhill (1979).

11 See Feferman (2005) for an overview and references.

12 Thanks to Ian Rumfitt for drawing my attention to the work of Weyl.

13 Carnap does not cite Weyl (1918) in any of his pre-*Syntax* works on logic, which I list in footnote 23. The first time that Carnap cites Weyl (1918) is in *The Logical Syntax of Language* (1937 [1934]).

14 See, in particular, Carnap's use of simple type theory in his (1929) and his (2000).

15 See Goldfarb (1989) and Hodes (2015) for discussions of Russell's and Whitehead's *actual* reasons for ramification.

16 See Chihara (1973) and Hylton (2005, chap. 5) for discussions of the vicious circle principle.

17 Russell (1908, p. 237) offers yet another formulation: "Whatever contains an apparent variable must not be a possible value of that variable". An *apparent variable* is what we would today call a *bound variable*.

1. No totality can contain members that are *definable* only in terms of that totality.
2. No totality can contain members that *presuppose* that totality.
3. No totality can contain members that *involve* that totality.[18]

It is not clear how the vicious circle principle should be understood, and, as Gödel (1984 [1944], p. 455) points out, the three formulations of the principle are not necessarily equivalent, since the expressions "definable only in terms of", "involving" and "presupposing" do not mean the same. However, the important point for present purposes is that the vicious circle principle creates a challenge for proponents of impredicative mathematics. The challenge arises from two different directions simultaneously.

First, Russell & Whitehead (see 1927 [1910], p. 39) make an additional assumption, according to which a propositional function is well-defined only if all of its values are well-defined. As they put it: "every propositional function presupposes the totality of its values", and therefore also the totality of its possible arguments. This assumption motivates a further formulation of the vicious circle principle that applies specifically to propositional functions, according to which a propositional function cannot have anything as argument which is defined in terms of that function. Otherwise, the totality of the values of a function would have an element that presupposes that totality, in violation of the vicious circle principle (see Russell & Whitehead (1927 [1910], p. 39)).

Second, the first formulation of the vicious circle principle (using 'definable only in terms of') appears to directly rule out impredicative definitions. If a domain is a "totality", impredicative definitions imply that some totalities include members that are definable only in terms of that totality.[19] For instance, a Dedekindian definition of the real numbers implies the existence of real numbers that are definable only by reference to all real numbers.[20] The first formulation of the vicious circle principle suggests that this definition should be rejected as meaningless. Simple type theorists hence are under pressure to explain why impredicative definitions are nevertheless admissible.

From Carnap's viewpoint in the late 1920s and early 1930s, the most prominent defense of impredicative definitions is due to Ramsey (1931 [1926], p. 41). Ramsey first distinguishes between *semantic* and *logical* or *mathematical* paradoxes, points out that solving the logical paradoxes (including Russell's paradox) requires only simple and not ramified

[18] To quote Russell & Whitehead (1927 [1910], p. 37) directly, they say:

(a) "If, provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total".
(b) "[G]iven any set of objects such that, if we suppose the set to have a total, it will contain members which presuppose this total then such a set cannot have a total".
(c) "Whatever involves *all* of a collection must not be one of the collection".

I hope that my above paraphrases are somewhat easier to parse. For instance, Russell and Whitehead diverge from the common usage of 'set', by presupposing that a set may not have a total. Since Cantor's early work on set theory, however, sets are commonly viewed as totalities of entities which, as such, are entities in their own right (see Cantor (1895, sec. 1)). Russell's and Whitehead's point would be better expressed using 'aggregate' or 'collection' in the place of 'set', as they do in the other two formulations of the vicious circle principle.

[19] Contemporary philosophers, following Cartwright's (1994) rejection of the so-called all-in-one principle, often deny the assumption that domains of quantification are "totalities".

[20] Or so it seemed to Carnap in the 1920s. See my previous discussion of recent developments in predicative mathematics.

type theory, and provides a separate solution for the so-called semantic paradoxes (p. 20). He then argues that impredicative definitions are no more problematic than to “refer to a man as the tallest in a group, thus identifying him by means of a totality of which he is himself a member” (p. 41). Ramsey’s point here is that impredicative definitions merely identify certain properties, but do not bring them into existence. As the example of the tallest man in a group illustrates, such an identifying use of definitions is entirely unproblematic. In this view, impredicative definitions hence are unproblematic as long as the defined properties exist independently of their explicit definition.

Ramsey’s metaphysical defense of impredicative definitions was later popularized by Gödel (1984 [1944]). Gödel explicitly argues that impredicative definitions are problematic only if one assumes that impredicatively defined properties depend for their existence on their explicit definitions. In Gödel’s words:

> “it seems that the vicious circle principle [...] applies only if the entities involved are constructed by ourselves. In this case there must clearly exist a definition (namely the description of the construction) which does not refer to a totality to which the object defined belongs, because the construction of a thing can certainly not be based on a totality of things to which the thing to be constructed itself belongs”. (p. 456).

However, if one believes that impredicatively defined properties exist independently from how we define them, then impredicative definitions are unproblematic. As Gödel says:

> “If, however, it is a question of objects that exist independently of our constructions, there is nothing in the least absurd in the existence of totalities containing members, which can be described (i.e., uniquely characterized) only by reference to this totality”. (p. 456).

The upshot is that, assuming that impredicatively defined properties exist independently from how we define them, impredicative definitions are unproblematic.[21,22]

§3. Carnap’s pre-*syntax* philosophy of mathematics. Carnap proposed two distinct non-Platonistic defenses of impredicativity, one in 1931 and one in 1937 [1934]. These two defenses of impredicativity took place in very different historical contexts, and in different phases of Carnap’s intellectual development. Understanding his arguments requires an overview of Carnap’s more general ambitions in the philosophy of mathematics, and of the historical evolution of his thought. In what follows, I first provide an overview of Carnap’s pre-*Syntax* philosophy of mathematics, then explain his 1931 defense of impredicativity, go on to explain his approach in *The Logical Syntax of Language*, and then explain his 1937 [1934] defense of impredicativity.

---

21 In addition to this philosophical defense, Gödel (1984 [1944], p. 455) also gave a pragmatic defense of impredicative definitions, according to which the fact that “a good deal of modern mathematics” is incompatible with the vicious circle principle should be taken as a refutation of this principle, rather than as a refutation of modern mathematics.

22 An alternative view, often attributed to Bernays (1983 [1935]), is that impredicative definitions are admissible as long as the impredicatively defined entity can also be *specified* independently, by predicative means. For instance, in Ramsey’s example of ‘*the tallest man in a group*’, it is possible to specify the defined entity independently, by pointing to him. See Linnebo (2017) for an overview.

Carnap was actively engaged in a major research project on pure logic and metamathematics during large parts of the 1920s, parallel and subsequent to his work on the *Aufbau* (1998 [1928]).[23] He pursued a program that combines aspects of formalism and logicism. Like Hilbert, Carnap sought a formalization of all of mathematics in axiomatic form. Carnap regarded the axiomatic approach to mathematics as an instance of a more general method for the analysis of statements and concepts, which is in principle applicable in various areas of inquiry, including set theory, geometry, and physics (see Carnap (1929)).[24] This general axiomatic method has two main components. In the first step, one defines an axiomatic system that allows one to deduce all statements from the chosen area of research. The axioms implicitly define a certain range of basic concepts.[25] The second step then consists in providing explicit definitions of all concepts from the chosen area of research in terms of the implicitly defined basic concepts.

Unlike Hilbert, Carnap supplements the axiomatic approach with a version of logicism. Carnap's (1931, pp. 91–92) version of logicism is defined by the following two principles:

(a) All mathematical concepts can in a stepwise process be explicitly defined in terms of a small class of logical concepts.

(b) All mathematical theorems can be logically deduced from logical axioms.

Carnap calls the view (a) that all mathematical concepts can be defined in terms of a small class of logical concepts "constructivism" (p. 94). He thinks of a Dedekindian definition of real numbers as sections of rationals as a typical example of explicitly defining a complex mathematical concept in terms of more basic ones (p. 94). 'Definitionism' may be a better term for Carnap's view, since he evidently does not mean a form of intuitionism (in conflict with how the term 'constructivism' is now commonly used).[26] However, in what follows I will adopt Carnap's terminology in order to avoid a mismatch between quotations from Carnap's text and my discussion.

---

23 The main published studies to come out of this research project were the textbook *Abriss der Logistik* (1929) and the articles "Eigentliche und Uneigentliche Begriffe" (1927), "Die Alte und die Neue Logik" (2004 [1928]), "Die Mathematik als Zweig der Logik" (1930), and "Die Logizistische Grundlegung der Mathematik" (1931). Carnap, furthermore, worked on a major book manuscript around 1928, under the title *Untersuchungen zur allgemeinen Axiomatik* (2000). This manuscript was not published during Carnap's lifetime, though a short summary of the main results appeared in the article "Bericht über die Untersuchungen zur allgemeinen Axiomatik" (1930/1931). See Reck (2007) for a helpful introduction to Carnap's contributions to modern logic. See also Schiemer, Zach, & Reck (2017) and the further references in this article.

24 In the preface to the textbook *Abriss der Logistik* (1929), he describes his motivation as follows: "The survey at hand does not aim at presenting a theory, but rather at teaching a practice. If someone aims at an exact analysis of the statements and concepts in an area of philosophy or the special sciences, then he shall be given here the logistic means, and in particular those from the theory of relations, as sharp tools" (p. III, my translation).

25 An implicit definition of a term is given when a number of principles or axioms involving it are laid down. For instance, the Peano axioms implicitly define the concept of a natural number. An explicit definition, by contrast, is an equation that identifies a term with another term. The definition of *being the least upper bound of a bounded class of real numbers*, which I discussed in the introduction, is an example of an explicit definition.

26 The term 'definitionism' is due to Feferman (2000, p. 182). Intuitionists believe in mathematical intuition as a means of acquiring mathematical knowledge and deny the law of excluded middle (see Iemhoff (2016) for a useful presentation). Carnap, however, denies that mathematical

The important point about such a logicist reduction, from the viewpoint of Carnap, is that it reconciles empiricism with the existence of mathematical knowledge.[27] Logicism shows mathematics to be *analytic*, assuming that logic is analytic. Carnap thinks that, since analytic sentences do not describe the world as being one way or another, they do not stand in need of empirical justification. If mathematics is reducible to logic, it is hence exempt from the requirement of empirical justification, and we may have mathematical knowledge even though it is not empirically verifiable.[28,29]

However, a key problem for Carnap's combination of formalism and logicism is that axiom systems are multiply realizable by formal models or physical structures. Carnap was well aware of this problem (see Carnap (1927, p. 362)). As he points out, axiom systems *explicitly* define second-order concepts, that distinguish structures that are models of the system from those that are not (p. 368). For instance, the Peano axioms explicitly define the concept of a progression. But various structures can satisfy the so-defined second-order concepts, which is why the concepts *implicitly* defined by an axiom system do not possess a unique realization but have the character of variables (p. 371).

Carnap (2000) tried to solve this (and other problems) by proving the equivalence of three different notions of completeness of consistent axiom systems: categoricity,[30] decidability[31] and "nonforkability" ("Nichtgabelbarkeit"). An axiom system is "nonforkable", roughly, if any two of its models satisfy all the same sentences. As Awodey & Carus (2001, p. 147) explain:

---

knowledge rests on mathematical intuition (1930, p. 308) and accepts the law of excluded middle (2000, pp. 81–82).

27 According to the kind of empiricism that Carnap favored in the late 1920s, all scientific knowledge ultimately rests on immediately given empirical knowledge. He thought that there is a certain set of basic "perceptual-physical" concepts that have perceptual marks, so that perceptual experience may definitely verify whether they are applicable in a particular instance. Carnap (1998 [1928]) then argued that all concepts of the empirical sciences can be defined in terms of these basic "perceptual-physical" concepts. He was thus led to the view that all sentences of the empirical sciences can be definitely refuted or verified by experience. See, for instance, the manifesto of the Vienna circle that Carnap co-authored (Verein Ernst Mach, 1929, p. 19), Carnap (2004 [1928], pp. 26–27) and Carnap (1963a, p. 57).

28 See Carnap (2004 [1928], p. 76) for a discussion of this point. In his intellectual autobiography, Carnap credits the origin of this basic strategy for reconciling empiricism with the existence of mathematical knowledge to Wittgenstein (1921), who argues that logic is tautologous. Carnap (1963a, p. 47) comments: "What was important in [Wittgenstein's] conception from our point of view was the fact that it becomes possible for the first time to combine the basic tenet of empiricism with a satisfactory explanation of the nature of logic and mathematics".

29 So far, I have mentioned the development of formal tools for the analysis and clarification of debates and the reconciliation of empiricism with the existence of mathematical knowledge as central motivations of Carnap's early philosophy of mathematics. Carnap was interested in further problems, too. Chief among them, explaining the applicability of mathematical and logical concepts to the physical world (1929, part II), and showing that the sciences are unified in the sense that the concepts of the various sciences are reducible to a common basis (1930/1931, sec. 8, and sec. 9).

30 A categorical axiom system is such that any two of its models are isomorphic. Carnap called a categorical axiom system "monomorphic".

31 Here, an axiom system is decidable if no axiom can consistently be added to it.

> "The idea behind this odd-sounding term [is] that some axiom systems (e.g., Euclidean geometry without the parallel axiom) can be added to in ways that are incompatible with each other (e.g., the parallel axiom or Lobachevski's axiom). Such a system can be said to reach a 'fork' (in the road) at that point".[32]

The so-called *Gabelbarkeitssatz* asserts the equivalence of these three notions,[33] and its truth was supposed to show that a consistent and decidable axiom system has a unique model (up to isomorphism). This would allow us to distinguish the *good* cases in which the implicitly defined concepts have a unique interpretation from the *bad* cases in which they don't.[34]

Against this backdrop, Carnap summarized his program on September 5, 1930, at the famous congress in Königsberg, as follows:

> "My supposition [...] is that this logical analysis of the formalistic system will have the following result; if this supposition is true, then despite the formalist method of construction, logicism would be justified and the opposition between the two approaches would be overcome:

1. For every *mathematical sign* one or more *interpretations* are found; and in fact purely logical interpretations.
2. If the axiom system is consistent, then upon replacing each mathematical sign by its logical interpretation (or any one of its various interpretations), every *mathematical formula* becomes a *tautology* (a generally valid sentence).
3. If the axiom system is complete (in Hilbert's sense: no nonderivable formula can be added without contradiction), then the interpretation is unique ["eindeutig"]; every sign has exactly one interpretation; with that the formalist construction is transformed into a logicist one".
   (Gödel, 1931, pp. 143–144, my translation).

Carnap alludes to the *Gabelbarkeitssatz* in the third of these statements.

However, a little later in the same roundtable discussion, Gödel used the very first public statement of the first incompleteness theorem specifically to raise a problem for Carnap's view.[35] He says:

> "One can even (assuming the consistency of classical mathematics) give examples of sentences (of the same kind as Goldbach's and Fermat's) that are actually true [inhaltlich richtig], but not derivable within the

---

[32] 'Gabel' is German for 'fork', 'gabelbar' can be translated as 'forkable', and 'Gabelbarkeit' as 'forkability'.

[33] The first part of Carnap's posthumously published *Untersuchungen zur allgemeinen Axiomatik* (2000) is devoted to proving this sentence.

[34] An interpretation by a physical structure would require an assignment of suitable values to the implicitly defined concepts.

[35] According to Carnap's diaries (Carnap, forthcoming), he first heard from Gödel of the incompleteness theorems ten days earlier, on August 26, 1930. Awodey & Carus (2001, p. 155) say that Carnap learned of the incompleteness theorems "a month before the [Königsberg] conference", but this timeline is not even supported by the source Awodey & Carus (2001) cite (i.e., Köhler (1991, pp. 150–151)).

> formal system of classical mathematics. Adding the negation of such a sentence to the axioms of classical mathematics, one obtains a consistent system in which a sentence is derivable that is actually false [inhaltlich falsch]". (Gödel, 1931, p. 148).

Gödel here effectively points out that the axiom system of Peano arithmetic is forkable, whereas Carnap (2000) argues that this system is decidable on the grounds that it is categorical. However, care is required to identify the mistake in Carnap's view, since his notion of completeness is different from the one that is standard today. Specifically, Carnap (2000) does not distinguish between the axiom system that is under investigation and the system that is used to carry out the investigation (see Awodey & Carus (2001, sec. 4)). Given this shortcoming, Carnap's results do not establish what he wanted them to show. For instance, Carnap formalizes 'axiom system f is consistent' as follows: $\neg\exists g(f \rightarrow g \wedge f \rightarrow \neg g)$. The provability of this formula in the system that is used to carry out the investigation is not equivalent with the consistency of the axiom system under consideration. The main problem with Carnap's results hence is not that they are invalid (they aren't), but rather that they are unsound and do not establish what's intended (see Awodey & Carus (2001, p. 159)).[36]

§4. Carnap's 1931 defense of impredicative definitions. Having presented Carnap's pre-*Syntax* philosophy of mathematics, I will now go on to explain his 1931 critique of Ramsey, and his first alternative, non-Platonistic defense of impredicative definitions.

Carnap (1931, p. 102) criticizes Ramsey's view as follows: "It seems to me that this view is not far away from a belief in a Platonic realm of ideas, which exist in themselves, independently from whether and in which way finite people are able to conceive of them". In other words, Carnap finds Ramsey's defense of impredicative definitions unacceptable since it rests on a problematic form of Platonism. Against this backdrop, Carnap characterizes the "*most difficult* problem confronting contemporary studies in the foundations of mathematics" as follows:

> "How can we develop logic if, on the one hand, we are to avoid the danger of the meaninglessness of impredicative definitions and, on the other hand, are to reconstruct satisfactorily the theory of real numbers as classes (or properties) of fractions?" (Carnap, 1931, p. 101, my translation).

Carnap here describes a dilemma: either one accepts impredicative definitions that, according to some mathematicians, really are meaningless, or one runs into problems concerning the theory of real numbers. Simple type theorists take the first horn of the dilemma, and ramified type theorists the second. Given Carnap's preference for simple over ramified type theory (see p. 7), this dilemma turns into the following problem: "Is it possible to retain Ramsey's results without accepting his absolutist conception?" This, according to Carnap, is "*the decisive question*" concerning the foundations of mathematics (1931, p. 103, my translation, Carnap's emphasis). I will in what follows first explain what Carnap means by

---

36 According to Carnap's diaries, he was made aware of this point by Tarski already in February of 1930, as the following entry from February 24, 1930, professes: "Tarski visits me [...]. About my Axiomatics. It seems correct, but certain concepts don't capture what is intended; they must be defined metamathematically rather than mathematically".

"Platonism" or "absolutism", then go on to explain why he finds it problematic, and finally present his alternative defense of impredicative definitions.

Carnap does not offer a definition of absolutism in the 1931 article. He, however, does offer one in the *Untersuchungen zur Allgemeinen Axiomatik* (2000, sec. 1.10).[37] According to this definition, absolutism contrasts with constructivism, and the key difference between these views concerns the use of quantifiers. Absolutists regard an existentially quantified sentence of the form '$\exists x Fx$' as meaningful, whether or not an object b that is F can in fact be found. Constructivists, by contrast, regard an existentially quantified sentence of the form '$\exists x Fx$' as meaningful only if it has either been inferred from a sentence of the form 'Fb', or else an object b which is F can be found in finitely many steps. This condition is very strong, since it entails that all meaningful existence claims are true. A more plausible version of constructivism would hold that an existentially quantified sentence of the form '$\exists x Fx$' is meaningful only if it has either been inferred from a sentence of the form 'Fb', or an object b which is F can be found in finitely many steps, or else it can be ruled out in finitely many steps that any object b is F.

Carnap explicitly marks his constructivist views on quantification as being in agreement with intuitionism. Unlike intuitionists, Carnap nevertheless upholds the principle of excluded middle. He observes that there is a tripartite division between objects that have been shown to be F, objects that have been shown to be not-F, and objects that have neither been shown to be F nor been shown to be not-F. However, everything is such that we know that it can be shown to be F, or we don't know that it can be shown to be F, which is why Carnap accepts the principle of excluded middle.

The basic difference between "absolutists" and "constructivists" can be illustrated as follows.[38] "Absolutists" define the property of being the least upper bound x of a bounded class C of reals thus:

(Abs) x is the least upper bound of a bounded class C of reals if and only if for every $q \in x$ there is a P such that $P \in C$ and $q \in P$.

Knowability plays no role in this definition. However, "constructivists" define the least upper bound x of a bounded class C of reals thus:

(Con) x is the least upper bound of a bounded class C of reals if and only if it can be shown for every q that, if $q \in x$, then some P can be found in finitely many steps such that $P \in C$ and $q \in P$.

It is clear why Ramsey counts as an "absolutist" according to this conception: Ramsey regards an existentially quantified sentence '$\exists x(Fx)$' as meaningful whether or not an x which is F can be found, and hence "goes beyond the limits of the truly knowable and definable" (Carnap, 1931, p. 102, my translation).

Given Carnap's constructivism, one should expect him to restrict quantification over infinite domains. If the domain of an existential quantifier is infinite, then, for at least some predicates F, it is not guaranteed that the truth of '$\exists x Fx$' can be decided in finitely many steps, as Carnap requires for this sentence to be meaningful. However, Carnap grants

[37] Carnap worked on *Untersuchungen zur Allgemeinen Axiomatik* in the years around 1928, and (according to his diaries (Carnap, forthcoming)) presented this work in a series of talks to other members of the Vienna circle in 1928.

[38] Compare Carnap's (2000, sec. 1.9) discussion of "absolutist" vs. "constructivist" definitions of the algebraic numbers.

that the domains of interest in mathematics generally are infinite (p. 103). He resolves the apparent conflict with constructivism by distinguishing between two different ways of proving general statements, which he calls proofs of "numeric" and of "specific" generality, respectively.[39] Here is an example to illustrate the difference:

(2) Every whale is a mammal.

A proof of the numeric generality of (2) would proceed by considering every individual whale and showing that it is a mammal. A proof of the specific generality of (2), however, first assumes that some arbitrary x is a whale, and shows that, since x is a whale, x is a mammal. Such a proof of specific generality does not require to consider each element of the domain of quantification and is compatible with quantification over infinite domains.[40]

This distinction allows Carnap (1931, pp. 103–105) to defend impredicative definitions, as follows. Consider the definition of *being the least upper bound of a bounded class C of reals*:

(1') $\lambda x_{\langle i \rangle}.[\forall_i q_i(q_i \in x_{\langle i \rangle} \leftrightarrow \exists_{\langle i \rangle} P_{\langle i \rangle}(P_{\langle i \rangle} \in C \wedge q_i \in P_{\langle i \rangle}))]$.

Carnap's constructivism imposes certain constraints on when the use of the quantifiers $\forall_i$ and $\exists_{\langle i \rangle}$ in this definition is to be regarded as meaningful. According to these constraints, (1') is meaningful if it can be shown for each $q_i \in x_{\langle i \rangle}$ (in finitely many steps) that some $P_{\langle i \rangle} \in C$ can be found (in finitely many steps) of which q_i is a member. Showing that (1') is meaningful hence requires establishing the following proposition: it can be shown for each $q_i \in x_{\langle i \rangle}$ (in finitely many steps) that some $P_{\langle i \rangle} \in C$ can be found (in finitely many steps) of which q_i is a member. Since there are infinitely many $q_i \in x_{\langle i \rangle}$, the *numeric* generality of this proposition cannot be proved. Its *specific* generality, however, can be proved: the least upper bound x of a bounded class C of reals P_i just is the set of all q_i that are elements of some P_i. Being an element of the least upper bound x of a bounded class C of reals P_i hence entails being an element of some $P_{\langle i \rangle} \in C$.

Carnap then imposes the further condition that the definition of a property P is meaningful if it is possible to decide, for at least some x, whether x has P.[41] That means with respect to (1') that this definition is meaningful if it is possible to decide, for at least some real number x, whether x is the least upper bound of C. This condition is met, too. We just need to find a real number that is represented by a set of rationals which includes elements that are not shared with (the representations of) any of the real numbers that are elements of C.[42]

[39] Linnebo (unpublished) provides a useful discussion of this distinction. He introduces a more illuminating terminology and distinguishes between "instance-based" and "generic" generality. Carnap relates the distinction to work by Kaufmann (1978 [1930]).

[40] Thanks to Øystein Linneo for helpful discussions.

[41] In Carnap's words (p. 104): "the proof that the defined property is (or is not) present in a particular instance can be given; the definition is thereby shown to be sensible". He makes clear that decidability in all instances is not required.

[42] Carnap makes this point by reflecting on a different example, which is the impredicative definition of *being an inductive number*. A number is inductive, according to this definition, if it has every heritable property of the number 0. This definition is impredicative since *being inductive* is itself a heritable property of the number 0. Carnap shows that it is decidable whether the number 2 is inductive according to this definition, and concludes that the definition is "sensible".

This view provides a specific example of a non-Platonistic defense of impredicative definitions, even though impredicative mathematics is often thought to be acceptable only on Platonistic grounds. However, Carnap's views on the foundations of mathematics soon shifted in a way that required him to search for an alternative defense of impredicativity, as I will go on to discuss in the next section.

§5. Carnap's *syntax* program. Carnap's views on the foundations of mathematics changed dramatically after he learned of Gödel's incompleteness theorems. According to his intellectual biography, a completely new approach

> "came to me like a vision during a sleepless night in January 1931, when I was ill. On the following day, still in bed with a fever, I wrote down my ideas on forty-four pages under the title 'Attempt at a metalogic'. These shorthand notes were the first version of my book *Logical Syntax of Language*". (Carnap, 1963a, p. 53).

It is not clear how accurate this recollection is, published 32 years after the fact.[43] However, the *Logical Syntax of Language* (1937 [1934]) does present a radically new approach to the foundations of mathematics, as I will discuss in this section.[44] I will first present Carnap's new "syntactic method", and then turn to explaining his 1934 defense of impredicative definitions.

The key difference between Carnap's pre-*Syntax* philosophy of mathematics and the new syntactic approach is that he replaces his earlier definitional reductionism by a new metalinguistic approach.[45] As discussed earlier, one of Carnap's goals during the 1920s was to show that mathematics is analytic via a definitional reduction of all mathematical concepts to a small class of logical concepts (see §3). There is no trace of this reductionism after 1931. Instead, Carnap (1937 [1934]) clearly distinguishes between object- and meta-languages, and provides meta-linguistic definitions of 'analytic' as a term that applies to object-language sentences. On this new approach, there is not a single notion of analyticity anymore. The meaning of 'analytic' rather has to be formally defined, and can be defined variously in different formal systems. Carnap thus tries to achieve his old goal of showing that mathematics is analytic by radically new means. However, this is not his only goal. Carnap, more generally, wants to provide a new "syntactic method" for the analysis of statements and clarification of disputes. As he puts it in the foreword (p. xiii):

> "The aim of logical syntax is to provide a system of concepts, a language, by the help of which the results of logical analysis will be exactly formulable. *Philosophy is to be replaced by the logic of science*—that is to say,

---

[43] In his diaries, Carnap does not mention a sleepless night in January 1931 (though he does mention illness). Furthermore, the intellectual autobiography does not even mention Carnap's *Untersuchungen our Allgemeinen Axiomatik*, even though this work must have absorbed most of Carnap's intellectual energy for several years. So, at least parts of Carnap's recollection in the intellectual autobiography appear inaccurate.

[44] See Ebbs (2001), Goldfarb & Ricketts (1992), Goldfarb (1996), Ricketts (1994) and the articles in Wagner (2009) for helpful discussions of Carnap's *Syntax* program.

[45] See Awodey & Carus (2007) for a discussion of the excellent question of what Carnp's key insight in the "sleepless night" was, and more generally of Carnap's development from the *Untersuchungen zur Allgemeinen Axiomatik* (2000) to *The Logical Syntax of Language* (1937 [1934]).

> by the logical analysis of the concepts of the sciences [. . .]. The book itself makes an attempt to provide, in the form of an exact syntactical method, the necessary tools for working out the problems of the logic of science".

As is evident from this quotation, Carnap develops the new syntactic method in pursuit of much of the same goals as the ones that guided the development of the axiomatic method.

An application of the syntactical method consists in the definition of the logical syntax of a language. By a "language", Carnap means what we would today call a "formal system", i.e., a formal language together with a deductive proof system. Carnap's "languages" resemble formal systems since they are specified by means of two sets of rules: *formation* and *transformation* rules. The formation rules specify which strings of symbols are sentences in the system. The transformation rules may include inference rules such as *modus ponens* or a list of axiom schemata, and they settle, for every sentence s and every set R of sentences of the system, whether s is a consequence of R.[46]

The logical syntax of such a system specifies what would today be called its *syntax*, i.e., the signs that occur in the system and their possible combinations. Carnap was concerned, moreover, with providing definitions of concepts of formal deductive *logic*—including concepts of analyticity, provability, logical independence, and so on. Carnap constructed syntactic definitions of these logical concepts, so that their application conditions depend merely on the forms and not on the meanings of sentences. The logical syntax of a language hence is a *formal* theory that makes "no reference [. . .] either to the meaning of the symbols (for example, words), or to the sense of the expressions (e.g., the sentences), but simply and solely to the kinds and order of the symbols from which the expressions are constructed" (1937 [1934], p. 1).[47] Carnap's further discussion makes clear that it is possible to define the logical syntax of a language whose component expressions *are* meaningful, and whose sentences *do* possess "senses". The logical syntax of such a language is a theory that ignores these "meanings" and "senses", and refers to only syntactical properties of the language in question.[48]

A syntactic treatment of language was important to Carnap because, he thinks, it is impossible to "lay down sharply defined rules" (1937 [1934], p. 1) for linguistic meanings. That is, Carnap regards it as possible to lay down syntactic composition rules that define how complex sentences may be built up from simpler expressions, and syntactic derivation rules that define how a sentence may be derived from a set of sentences. He, however, regards it as impossible to similarly lay down semantic composition rules that define the meaning of a complex expression as a function of the meanings of its component parts, or semantic entailment relations. He learned of Tarski's semantic truth-definitions only

46 See Ebbs (2001, sec. 1.1) for a helpful discussion of Carnap's formation and transformation rules.

47 Carnap (1963a, p. 53) retrospectively offers a similar characterization of what is meant by the logical syntax of a language.

48 Carnap (1937 [1934], p. 5) explicitly says the following: "When we maintain that logical syntax treats language as a calculus, we do not mean by that statement that language is nothing more than a calculus. We only mean that syntax is concerned with that part of language which has the attributes of a calculus—that is, it is limited to the formal aspect of language. In addition, any language has, apart from that aspect, others which may be investigated by other methods. For instance, its word have meaning; this is the object of investigation and study for semantics".

in 1935, after the German edition of *The Logical Syntax of Language* had already been published (in 1934).[49]

Carnap illustrates the syntactical method with two examples: Language I and Language II.[50] Carnap specifies both Languages by means of *formation* rules that define sentencehood, and by means of *transformation* rules that define a relation of logical consequence. These rules are in both cases defined for an object language L in a metalanguage M. Language I corresponds to a constructivist viewpoint, while Language II corresponds to the viewpoint of simple type theory.[51] The beliefs of constructivists are reflected in Language I by, for instance, the fact that Language I contains only definite[52] predicates and only restricted quantifiers. Language II also contains unrestricted quantifiers and indefinite predicates.

From a formal standpoint, Language I is a version of primitive recursive arithmetic (PRA). The symbolic apparatus of Language I contains the usual Boolean connectives plus identity, negation and the so-called K-operator,[53] a restricted existential quantifier,[54] numerical variables and constants, first-order predicates, and mathematical functors (p. 16). The sentences of Language I are either identities between numerical expressions, or consist of the application of a predicate to some argument, or are formed using negation, Boolean operators, or restricted quantifiers (p. 26).

Carnap distinguishes between a relation of direct derivation (p. 32) and a direct consequence relation for his Language I (p. 38).[55] The relevant difference is that derivations have only finitely many premises, while consequences may be drawn from infinitely large sets of premises (using the infinitary omega rule DC2, p. 38). This duplication is needed since Gödel's first incompleteness theorem shows that a finitary inference rule cannot capture

---

49 According to Carnap's diaries (forthcoming), he first heard through personal communication with Tarski of his semantic truth-definitions on 06/30/1935, the year after the first publication of *The Logical Syntax of Language*. Carnap knew Tarski since February of 1930, when Tarski gave lectures on metamathematics in Vienna (see Carnap (1963a, p. 30)). Tarski's semantic theory of truth was first presented by Lesniewski a year later, in 1931, and published in Polish language in 1933 (see Tarski (1956 [1936], p. 154, footnote 1)). Carnap encouraged Tarski to present this theory at the International Congress for Scientific Philosophy in Paris in 1935, where it became known to a wider audience.

50 One often hears that Carnap's discussion of both Language I and II in *The Logical Syntax of Language* signals a change of mind or shift in Carnap's strategy. For instance, Coffa (1993, p. 187) says: "There is a certain lack of coherence in Carnap's treatment of analyticity and consequence in *LSL*, suggesting that there may have been two stages in the development of his ideas on the topic". I think this is a mistake. Language I and II are two applications of a general method, and developed in the service of a common, overarching goal.

51 Carnap's use of the term 'constructivist' in *The Logical Syntax of Language* is closer to its contemporary usage, as compared to Carnap's use in his (1931) article. For example, Carnap (1937 [1934]) specifies that a "constructivist" viewpoint requires using only quantifiers restricted to finite domains, whereas Carnap (1931) describes himself as a "constructivist" but nevertheless admits quantification over infinite domains.

52 Definite number-properties are ones whose "possession or nonpossession by any number whatsoever can be determined in a finite number of steps according to a fixed method" (Carnap, 1937 [1934], sec. 3).

53 If 'Gr(a, b)' means "a is greater than b", then '(Kx)9(Gr(x, 7))' means "the smallest number up to (and including) 9 which is greater than 7". I.e., '(Kx)9(Gr(x, 7))' designates the number 8.

54 Each instance of the quantifier ranges over only natural numbers up to a specific limit, e.g., over all the natural numbers up to 3.

55 He draws a similar distinction for Language II.

our pretheoretical notion of consequence for arithmetical languages.[56] Using contemporary notation, Carnap's infinitary omega rule DC2 says the following:

DC2 A sentence S_1 of the form '$\ldots\ x_1 \ldots$', with x_1 occurring as free[57] or bound variable in S_1, is a *direct consequence* (in Language I) of the infinite set of sentences $C_1 = \{S_1[x_1/1],\ S_1[x_1/2],\ S_1[x_1/3], \ldots\}$, where $S_1[x_1/n]$ is the result of replacing the variable x_1 by the numeral n. (see p. 38).

This rule differs in important ways from Tarski's (2002 [1936], p. 178) rule of infinite induction, which says the following:

Infinite induction A proof of all sentences in the infinite set $\{P(0), P(1), P(2), \ldots\}$ is a proof of the sentence $\forall x(Px)$.

The rule of infinite induction describes a relation between a set of proven sentences and a sentence, while the rule DC2 describes a relation between a set of sentences, for which we may or may not have a proof, and a sentence. So, DC2 is not a rule of infinite induction. Carnap therefore introduces an additional derivation rule, called 'Complete Induction' (p. 32), which (corresponding to Peano's fifth axiom) allows one to infer $\forall n(Pn)$ from the two premises $P(0)$ and $P(n) \supset P(n+1)$.[58] This point is important since, as I will explain next, Carnap uses the relation of direct consequence to define analyticity, and he is fully aware that some analytic sentences are undecidable (p. 28).[59]

The definition of 'analytic in Language I' is now straightforward: a sentence S is analytic in Language I iff it is a consequence of the nullset of sentences. That means that the proof of an analytic sentence requires no premises; analytic sentences are logical truths. A sentence is synthetic in Language I, by contrast, iff it is neither analytic nor contradictory in Language I (p. 40).

The chief difference between Language I and Language II is that the latter is a typed language that admits of higher-order quantification.[60] To this end, Language II makes use of an expanded symbolic apparatus, that also includes predicate, functor and sentential variables, as well as unrestricted quantifiers. The definition of 'analytic in Language II' superficially looks like a notational variant of a Tarskian truth-definition. Carnap proceeds in three steps: he first assigns "valuations" to all expressions. Carnap, second, assigns truth-values to unquantified sentences depending on the "valuations" of their component expressions. Specifically, 'Fa' is true iff the "valuation" of 'a' is a member of the valuation of 'F'; and an identity is true iff the two terms flanking the identity sign have the same "valuation". He then, in a third step, defines 'analytic in Language II'.

56 On p. 28 Carnap directly credits Gödel for showing that not all analytic sentences are "demonstrable" (i.e., decidable). This suggests that Carnap distinguishes between the derivation and the consequence relations precisely in order to deal with undecidable sentences.

57 Carnap uses free variables in order to express unlimited generality in Language I, see p. 21.

58 Thanks to Gary Ebbs for helpful discussions.

59 I am here disagreeing with Coffa (1993, p. 288), who says "Carnap's preferred 'indefinite' rule is infinite (transfinite) induction: From $F(n)$, for each natural number n, infer $(x)Fx$. Hilbert and Tarski had studied this rule, and Gödel's recent work had brought it to prominence by emphasizing the fact that even though the rule is intuitively sound, it is not a derived rule in standard systems of arithmetic". I argue that Carnap and Tarski use importantly different rules.

60 For Language I, Carnap first defines a consequences relation, and then uses it in order to define analyticity. For Language II, Carnap first defines analyticity, and then uses this definition to define the consequence relation.

In more detail, the definition of 'analytic in Language II' distinguishes between three main cases:[61] sentences that (i) contain no operator, (ii) are prefixed by an existential quantifier, or (iii) are prefixed by a universal quantifier. (i) Sentences containing no operator are analytic iff they have the form of a tautology. (ii) An existentially quantified sentence is analytic iff the embedded sentence is true on some valuation of the bound variable. (iii) A universally quantified sentence is analytic iff the embedded sentence is true on all possible valuations of the bound variable.

This definition of 'analytic in Language II' can in principle be given in a meta-language whose expressive powers outstrip those of Language II, and that contains predicate and functional variables of types that do not occur in Language II. However, an attempt to give a definition of 'analytic in Language II' *in Language II* induces a stratification into infinitely many sublanguages, or "regions" of Language II (see Carnap (1937 [1934], p. 113)).[62] This is because it is impossible to define 'analytic in L' in language L itself if L is consistent, as shown by Tarski's undefinability theorem.[63] Language II contains variables of infinitely many types, so that there is no highest type that could be used for defining 'analytic' for all sentences of a lower type. Carnap hence distinguishes between various "regions" in Language II (II_1, II_2, ...), where every region is contained in all subsequent regions and Language I is contained in II_1 (see Carnap (1937 [1934], p. 88)). For each of these regions II_n of Language II, 'analytic in II_n' is undefinable in II_n, but is always definable in a more extensive region II_{n+m} that functions as a metalanguage for II_n. Language II hence in fact constitutes an infinite hierarchy of languages.[64]

§6. Carnap's distinction between syntax and semantics. Commentators often remark that Carnap (1937 [1934]) makes use of semantic resources, despite his avowed intention to treat languages syntactically. For instance, Goldfarb (1997, p. 62) says that: "Already in [The Logical Syntax of Language] Carnap had given what amounts to Tarskian semantics for the mathematical part of the object language."[65,66] This point raises a question. Carnap stresses on p. 1 of *The Logical Syntax of Language* that he wants to treat languages purely syntactically, without reference to the meanings or senses of symbols. But Tarskian truth-definitions are commonly regarded as prototypical *semantic*

[61] I am omitting many details. For instance, Carnap introduces a separate clause, defining that a class of sentences is analytic iff each of its members is analytic. An open sentence is analytic iff the sentence that results by prefixing a universal quantifier is analytic. And an unreduced sentence in nonstandard form is analytic iff the sentence that results after applying certain reduction rules is analytic.

[62] Thanks to Gary Ebbs for helpful discussions related to this point.

[63] More precisely: Tarski's undefinability theorem concerns semantic truth-definitions, while Carnap gives a syntactic definition of 'analytic' that structurally resembles a Tarskian semantic truth-definition. Carnap (Theorem 60 C.1, p. 219) however proves a syntactic version of Tarski's undefinability theorem that applies to Carnap's syntactic definitions of 'analytic'.

[64] The undefinability theorem is commonly credited to Tarski. But Gödel had come upon a proof of the theorem already when he was working on the incompleteness proofs. Since Carnap was well acquainted with Gödel's work, it is no surprise that the undefinability theorem influenced Carnap's construction.

[65] See also Coffa (1993, p. 288) and Goldfarb & Ricketts (1992, p. 70).

[66] There also are important differences between Carnap's definition of 'analytic in Language II' and a Tarskian definition of 'true sentence'. For example, Carnap's clauses (ii) and (iii) of Carnap's definition are not inductive since he uses 'analytic' on the left-hand side and 'true' on the right-hand side. Thanks to Harold Hodes for this observation.

truth-definitions. Why should Carnap's definition of analyticity for Language II then be *syntactic*? More generally, it is unclear what the difference between syntax and semantics is supposed to be.[67] This unclarity is made worse by the fact that early versions of the work that was eventually published as *The Logical Syntax of Language* were called, first, 'Metalogic' and then 'Semantics'.[68] However, I will argue that there is a clear difference between a Tarskian truth-definition and Carnap's definition of analyticity for Language II, and that there is a sense in which, from Carnap's perspective at least, the definition of 'analytic in Language II' is syntactic.

To review: Tarski's (1956 [1936]) goal is to provide a "materially adequate and formally correct definition of the term 'true sentence"' (p. 152).[69] Tarski provides this definition in a meta-language M for a formal object-language L of set-theory. Such a meta-linguistic definition of truth is *formally correct* just in case it is of the following form:

For all x, $True(x)$ if and only if $\phi(x)$,

where 'true' does not appear in ϕ. Furthermore, the stated equivalence must be provable using axioms of the meta-language that do not contain 'true'. The definition is *materially adequate* just in case the objects satisfying ϕ are exactly the true sentences of L. For the kind of semantic truth-definitions that Tarski gives, the meta-language must satisfy several requirements. First, it needs to contain a copy, or translation, of the object-language, so that everything that can be said in M can also be said in L. Furthermore, L needs to contain names for all sentences in M. Tarski constructs these names using single quotation marks. He then proceeds in two steps.[70] He first defines a satisfaction relation between sentential functions and assignments (of values to sequences of variables). For example, a sequence of classes f satisfies the inclusion function $\iota_{k,l}$ if the k^{th} class is a subset of the l^{th} class in the sequence: $f_k \subseteq f_l$ (Definition 22, p. 193). Tarski then defines truth as follows: x is a true sentence if and only if x is a sentence and every infinite sequence of classes satisfies x (Definition 23, p. 195). This definition has each instance of the T-schema "'ϕ' is true if and only if ϕ" as a consequence.[71] This shows that the proposed definition is materially adequate.

---

67 Thanks to Lydia Patton for helpful questions and comments on this point.

68 In his diaries (forthcoming), Carnap first mentions "metalogic" on 12/19/1929, and then (after a long hiatus) frequently mentions working on this topic in 1931 and 1932 (beginning on 02/02/1931). On 03/26/1932, Carnap discussed the title with Neurath and Gödel. Neurath suggests "Universelle Syntax, logische Grundlagen der Einheitswissenschaft" ("Universal Syntax, Logical Foundations of Unified Science"). Gödel suggests "Semantik". Carnap henceforths (in 1932 and 1933) frequently mentions working on "semantics". On 05/25/1932, "Frank" suggests the eventual title "Die Logische Syntax der Sprache". 'Frank' presumably refers to the physicist Philipp Frank, who was the successor of Albert Einstein in Prague, and who supported the hiring of Carnap by the University of Prague in 1931 (see Limbeck-Lilienau & Stadler (2015, p. 62)). In 1933, Carnap frequently speaks of working on "syntax".

69 See Hodges (2014) for a very useful introduction.

70 The reason for this procedure has to do with the semantics of quantifiers and the wish to give a compositional theory, that defines the truth of a sentence as a function of the truth of its component parts. According to Tarski, '$\forall x Fx$' is true relative to an assignment a just in case 'Fx' is true relative to each assignment that differs from a at most with respect to the value assigned to x.

71 This is the disquotational version of the T-schema. The more general version of the T-schema requires only that in the place of the right-hand side occurrence of ϕ appears a formula with the same truth-conditions as the left-hand side occurrence of ϕ. I.e., "ϕ' is true of and only if ψ' is an instance of the more general T-schema if and only if the formula that 'ϕ' refers to has the

https://doi.org/10.1017/S1755020318000370 Published online by Cambridge University Press

I think that the main difference between Carnap's definition of analyticity for Language II and a Tarskian truth-definition is that Carnap does not require his meta-languages to contain a copy, or translation of the relevant object-languages. Two points of clarification are important.

First, Carnap's meta-languages do contain *names* for object-language expressions, as evidenced by Carnap's use of Gödel's method of arithmetization in the construction of Language I.[72] But a name of an object-language expression is not the same as a copy, since a name for an expression e typically does not have the same meaning as e (unless e is a self-referential expression). Carnap's meta-languages, just like Tarski's, do contain names for object-language expressions, but do not necessarily contain a translation of the relevant object-languages.

Second, Carnap does acknowledge that a meta-language *may* contain a translation of some object-language. He says (§62):

> "The interpretation of the expressions of a language S_1 is [...] given by means of a *translation* into a language S_2, the statement of the translation being effected in a syntax-language S_3; and it is possible for two of these languages, or even all three, to coincide". (p. 228).

In this example, there is one metalanguage S_3, which Carnap calls a "syntax-language", and two object-languages, S_1 and S_2. The acknowledgement that the three languages may coincide shows that Carnap thinks that a meta-language may contain a translation of some object-language. However, Carnap does not *require* that the syntax-languages used to define analyticity contain translations of the relevant object languages. For example, he identifies the various "regions" of Language II in purely syntactical terms, and there are no translations between them:

> "Not counting [predicates] and [functors], all the symbols already occur in II_1, and thus in every region. Operators with [sentential variables] occur for the first time in II_2. In II_1, [first-order predicates] and [first-order functors] occur both as constants and as free variables, but not as bound variables. Furthermore, in a region II_n ($n = 2, 3, \ldots$) [predicates] and [functors] occur as constants and as free variables up to the [order] n, but as bound variables up to the [order] $n - 1$". (p. 88).

Carnap here describes the subdivision of Language II into regions without using any semantic vocabulary, and does not require that higher regions contain copies or translations of regions that are lower in the hierarchy. He does describe the various regions as "concentric" (p. 88), but the mere repetition of symbols in the various regions does not guarantee the existence of a translation or sameness of meaning.

The circumstance that Carnap does not require his syntax-languages to contain a copy or translation of the object-language has important consequences. For one, Carnap's syntax-languages, unlike Tarski's meta-languages do not contain descriptive vocabulary. Informally, the difference between logical and descriptive expressions is that the logical

same truth-conditions as 'ψ'. Truth can be defined disquotationally using meta-languages that contain a *copy* of the relevant object-languages. A truth-definition in a meta-language that merely contains a *translation* of the object-language needs to make use of the more general version of the T-schema.

72 Thanks to an anonymous referee for raising this point.

https://doi.org/10.1017/S1755020318000370 Published online by Cambridge University Press

expressions have a purely logical or mathematical meaning, while descriptive expressions designate something nonlogical, such as physical objects (see Carnap (1937 [1934], p. 177)). More formally, sets of sentences that are composed of only logical vocabulary are determinate, in the sense that they are either valid or entail a contradiction.[73] This is not true of sets of sentences that are in part composed of descriptive sentences. Carnap's syntax-languages contain only logical vocabulary. But if an object-language that contains descriptive vocabulary is translated into a meta-language, then the meta-language inherits the descriptive vocabulary, as is the case with Tarski's meta-languages. Carnap remarks on this difference in the following passage form his *Intellectual Biography*:

> "In his treatise Tarski developed a general method for constructing exact definitions of truth for deductive language systems, that is, for stating rules which determine for every sentence of such a system a necessary and sufficient condition of its truth. *In order to formulate these rules, it is necessary to use a metalanguage which contains the sentences of the object language or translations of them and which, therefore, may contain descriptive constants, e.g., the word 'black' in the example mentioned.* In this respect, the semantical metalanguage goes beyond the limits of the syntactical metalanguage. This new metalanguage evoked my strongest interest. I recognized that it provided for the first time the means for precisely explicating many concepts used in our philosophical discussions". (Carnap, 1963a, pp. 60–61, my emphasis).

An important difference between Tarski's semantic meta-languages and Carnap's syntax languages hence is that only the former contain descriptive vocabulary.[74]

Furthermore, a translation of the object- into the meta-language is needed for stating instances of the T-schema: "'ϕ' is true if and only if ϕ". If the metalanguage does not contain either a copy or a translation of the object-language, then the instances of the T-schema can't be stated. As remarked earlier, Carnap's syntax-languages *may* contain translations of the relevant object-languages. However, the circumstance that he did not require the existence of such a translation shows that he was unaware of the important technical work that these translations may do for a theory of truth, and specifically was not aware of the importance of the T-schema. This point has important ramifications.

First, Tarski's truth-definitions and Carnap's definitions of analyticity have to be understood as serving somewhat different goals. Tarski assigns truth-conditions to object-language sentences, as expressed by means of instances of the T-schema. Carnap does not assign truth-conditions to sentences but rather identifies a certain class of sentences as the analytic ones.

Second, Carnap remarks (§51) that it is in principle possible to define rules that would classify all the accepted descriptive sentences as valid.[75] But he regards this project as impractical, given that which descriptive sentences are accepted constantly changes. Carnap

---

[73] For Carnap, the distinction between logical and descriptive expression is relative to a language system. A sentence is valid in a system if it is entailed by its axioms and transformation rules. An expression can be logical in one system and descriptive in another.

[74] This point is made by both Quine (1960 [1954], p. 367) and Goldfarb (1997, p. 61).

[75] To this end, one would need to define P-rules, which are transformation rules that apply to descriptive sentences. A descriptive sentence then is valid (or, better, P-valid), if it is entailed by the P-rules.

therefore chooses to treat descriptive and logical sentences asymmetrically. A descriptive sentence S_1 is defined to be analytic in Language II iff it is a substitution instance of a general analytic logical sentence. A logical sentence is analytic iff true under each of its possible valuations (see Carnap (1937 [1934], p. 112)). This definition contrasts with Tarski's semantic truth-definitions, which do not invoke a distinction between logical and descriptive sentences.

Third, the proof that a Tarskian truth-definition entails all instances of the T-schema is important for showing that the definition is *materially adequate*. That means that *without* the needed instances of the T-schema, there is no way of proving that a syntactic definition of a truth-predicate is materially adequate. This explains why Carnap regarded a syntactic truth-predicate that applies to descriptive sentence as impossible (see p. 216), at least assuming he means a materially adequate truth-definition. Since he was not aware of a method for proving the material adequacy of a truth-definition that applies to descriptive sentences, he regarded such a truth-definition as impossible.[76]

In sum, I have argued that Carnap, in contrast to Tarski, did not require his syntax-languages to contain a translation or copy of the relevant object-language. This point does not reveal a fundamental difference between the two approaches, since Carnap's syntax-languages *may* contain a translation of the relevant object-language. However, it does reveal that Carnap was not aware of the importance of the T-schema and its instances. Furthermore, the circumstance that Carnap did not require his syntax-languages to contain a translation of the object-language illuminates how he thought about the difference between syntax and semantics. Given that Carnap's syntax-languages do not contain a translation of the object-language, the object-language meaning of an expression may come apart from its meta-linguistically assigned valuation. Carnap, for this reason, distinguishes between "material interpretations" and "syntactical valuations" (see, e.g., sec. 50, p. 177). For example, the material interpretation (or object-language meaning) of the numeral 'five' is the number five, while its meta-linguistically assigned "syntactical valuation" is the numeral 'five' itself. This point is important. As explained in the last section, Carnap regards the syntactical method as "syntactic" because no reference is made to the meanings or senses of linguistic expressions. Given the distinction between material interpretations and syntactical valuations, this point must be disambiguated. I think Carnap means that an application of the syntactic method avoids all reference to the *material interpretations* of expressions, though reference to *syntactical valuations* is permitted. Carnap's definition of analyticity for Language II, on this interpretation, is syntactic in the following sense: the definition makes no reference to the material interpretations of object-language sentences, but only to their syntactical valuations. 'Analytic in Language II' is a syntactic predicate in the sense that it applies to a sentence solely in virtue of its syntactical valuation and independently of its material interpretation.

§7. Higher-order quantification. I have explained how, in *The Logical Syntax of Language* (1937 [1934]), Carnap turns the decision between simple and ramified type theory into a decision between alternative formal systems, that include alternative definitions of analyticity. Next, I will explain how this move supposedly freed the decision between

[76] However, as Coffa (1993, p. 303) correctly observes, since Carnap's syntax-languages can *in principle* contain a translation of the object-language, Carnap is *in principle* able to provide a truth-definition even for descriptive sentences.

https://doi.org/10.1017/S1755020318000370 Published online by Cambridge University Press

simple and ramified type theory from metaphysical implications. I will first explain why Carnap was under pressure to address this question, and then explain his answer.

Carnap's definition of 'analytic in Language II' raises a question. The definition of analyticity for any given region of Language II includes formulas that make use of higher-order quantifiers.[77] That is, valuations of predicates and of predicate variables are classes of numerical constants,[78] and a sentence of the form '$\forall F\varphi$' (binding a first-order predicate variable) is defined to be analytic iff the embedded sentence φ is true *for each* valuation of the variable F. (The relevant quantifier is in italics.) These higher-order quantified formulas require (what we would today call) an objectual interpretation of quantification. According to an *objectual* interpretation of quantification, '$\forall x(Fx)$' is true iff 'Fx' is true for every value of the variable 'x', whether or not this value can be named. This interpretation contrasts with a *substitutional* interpretation, according to which the sentence '$\forall x(Fx)$' is true iff '$F\overline{n}$' is true for every substitution instance '$\overline{n}$' of 'x'. Carnap needs an objectual interpretation since, for every region II_n of Language II, there are real numbers that are definable only at regions II_{n+m} that are higher up in the hierarchy.[79] The quantifiers that appear in the definitions of analyticity for regions of Language II hence cannot be restricted to definable classes. Or at least this is the case if a sentence that is classified as analytic at one step in the hierarchy should retain this status as one moves up in the hierarchy—since then the quantifiers must range over (unrestrictedly speaking) all real numbers. A quantifier that ranges over a domain with undefinable elements cannot be interpreted substitutionally but requires an objectual interpretation.

Higher-order quantifiers under an objectual interpretations bind variables that range over properties independently of whether these properties possess a name. Carnap's use of higher-order quantification under an objectual interpretation hence seems to commit Carnap to the view that properties exist independently of our means of referring to them, which amounts to taking a metaphysical position regarding the ontological status of properties. As Carnap puts it:

> "But do we not by this means arrive at a Platonic absolutism of ideas, that is at the conception that the totality of all properties, which is non-denumerable and therefore can never be exhausted by definitions, is something which subsists in itself, independent of all construction and definition?" (p. 114).

Carnap credits such a "metaphysical conception" to Ramsey, and rejects it, by saying that "from our point of view [it] is definitely excluded" (p. 114).

Carnap's remarks clearly echo his 1931 criticism of Ramsey's account of higher-order quantification. However, it is not easy to see what, given Carnap's 1934 perspective, the problem with Ramsey's view is supposed to be. In the 1931 article, Carnap contrasts his constructivist approach with Ramsey's absolutism, and criticizes the latter. By 1934 he

[77] See in particular the clauses DA1 Cb and DA2 Ca of the definition of 'analytic in Language II' (Carnap, 1937 [1934], pp. 111–112).

[78] Carnap (1937 [1934], p. 113) calls them "syntactical properties of accented expressions".

[79] Carnap (1937 [1934], p. 106) credits Gödel and not Cantor for this result, even though Carnap's corresponding Theorem 60 d.1 (p. 221) makes essential use of Cantor's diagonal method. The reason seems to be that Carnap's explication of Cantor's reasoning relies on Carnap's diagonal lemma, or fixed-point theorem, which in turn makes use of Gödel's method of arithmetizing the syntax of a language. Thanks to Gary Ebbs for helpful discussions of this point.

has given up constructivism. It is hence unclear what in 1934 the problem with Ramsey's conception should be. One might think that Carnap aims to avoid ontological commitments that Ramsey possesses. But at least in his later work Carnap accepts Quine's (1953 [1948]) criterion, according to which the proponent of a theory is committed to the existence of all and only those entities over which the values of the variables in the theory must range for it to be true (see Carnap (1956 [1950], p. 214, footnote 3)). Given this criterion, Carnap's use of higher-order quantifiers commits him to the existence of mathematical properties.

However, in *Meaning and Necessity* (1956 [1947], pp. 42–43), Carnap also clarifies that he finds the term 'ontological commitment' misleading:

> "I would prefer not to use the word 'ontology' for the recognition of entities by the admission of variables. This use seems to me to be at least misleading; it might be understood as implying that the decision to use certain kinds of variable must be based on ontological, metaphysical convictions".

An assessment of Carnap's (1937 [1934], p. 114) critique of Ramsey in the light of these later remarks suggests that he sees an important difference between the mere possession of an ontological commitment and the acceptance of a Platonic ontology. Only the latter involves the assumption of a metaphysical viewpoint. If that's right, then the key question Carnap needs to answer is this: what is the difference between the possession of an ontological commitment and the acceptance of a Platonic ontology?[80] (Carnap says that he does not like the term 'ontological commitment', but he does not offer an alternative one. I will in what follows continue to speak of ontological commitments, as the commitment to the existence of entities via the use of bound variables. It should be understood that an ontological commitment, according to Carnap, need not be based on metaphysical conviction. The main question is: why not?)

In my view, Carnap's answer to this question crucially brings in the notion of an analytic statement. It is here important that Carnap's notion of analyticity is very different from the contemporary notion of truth-in-virtue-of-meaning. As briefly remarked in §3, Carnap's notion of analyticity is inspired by Wittgenstein's notion of a tautology. Analytic statements, for Carnap, are distinguished by the lack of any kind of descriptive content.[81] They do not represent the world as being one way rather than another. This notion of analyticity helps to explain the difference between an ontological commitment and the acceptance of a Platonic ontology. An analytic statement, even if it is ontologically committing, nevertheless does not express a metaphysical viewpoint since it does not represent the world as being one way rather than another.

To apply this idea to the problem at hand: Carnap needs to explain why his use of higher-order quantifiers does not commit him to a metaphysical viewpoint. As I explain in more detail below, he solves this problem by arguing that it is analytic that predicate variables refer in Language II. That does not mean that the sentence 'predicate variables refer' satisfies 'analytic in Language II'. Language II does not even contain the predicate '(to) refer'. Rather, the rules that define Language II entail that predicate variables possess

80 This dialectic is mirrored in Carnap's (1956 [1950]) article "Empiricism, Semantics and Ontology", where Carnap, far from avoiding ontological commitments, argues that his acceptance of abstract entities is "compatible with empiricism and strictly scientific thinking" (p. 206).

81 I discuss what Carnap means by "descriptive content" in §6.

https://doi.org/10.1017/S1755020318000370 Published online by Cambridge University Press

a valuation, independently from what contingent features of reality may be like. To put the point another way, the use of predicate variables may appear to commit one to a metaphysical viewpoint given that, if properties do not exist, then predicate variables have no referent. Carnap argues, however, that predicate variables refer in Language II simply in virtue of how this system is defined. This is why Carnap's use of higher-order quantifiers, even though it is ontologically committing, does not imply the acceptance of a Platonic ontology.[82]

Carnap makes this point as follows. First, he raises the question: "can the phrase 'for all properties...' (interpreted as 'for all properties whatsoever' and not 'for all properties which are definable in *S*') be formulated in the symbolic syntax-language *S*?" Carnap is here concerned with the definition of analyticity for a particular region of Language II, i.e., region II_1.[83] He calls the metalanguage in which the definition for this region is formulated the "syntax-language *S*". Carnap raises the question of whether it is possible to formally interpret sentences involving a higher-order quantifier under an objectual interpretation in the meta-language *S* for region II_1. He answers this question in the affirmative (p. 114):

> "That [the higher-order quantified phrase '(F)(...)'] in the language *S* has the meaning intended is formally established by the fact that the definition of 'analytic in *S*' is formulated in the wider syntax-language S_2 not by substitutions [of the predicate variables] but with the help of syntactical valuations".

S_2 is a region of Language II that is higher up in the hierarchy as compared to *S*, and functions as a meta-language for *S*. Carnap's point here is that the statements he makes using higher-order quantified formulas in *S* can be formally interpreted, since they are assigned precise syntactical valuations using the meta-meta-language S_2.[84]

A natural objection to Carnap's view is that he cannot assume that the interpretation of object-language sentences by means of meta-linguistic sentences is successful. If terms in the meta-language do not refer, then it is of no use for the purposes of interpreting

[82] To avoid a possible misunderstanding, let me be clear that it is here inessential that "syntactical valuations" of numerals are themselves numerals and not numbers. The distinction between material interpretations and syntactical valuations which I have drawn in the last section is important for explaining how Carnap thought about the difference between syntax and semantics. It is inessential for explaining how Carnap tried to make use of higher-order quantifiers without committing to a metaphysical viewpoint.

[83] II_1 is the initial region in the hierarchy of regions which is Language II and is contained in all subsequent regions. II_1 contains Language I as a proper part, and contains predicate and functional constants as well as free predicate and functional variables, but no bound predicate and functional variables. I.e., II_1 contains no higher-order quantifiers. See Carnap (1937 [1934], p. 88).

[84] Carnap's claim that higher-order quantified formulas have "the meaning intended" may seem to conflict with Carnap's ambition of providing a purely syntactic treatment of language. However, as explained in the last section, Carnap distinguishes between material interpretations and syntactical valuations. The material interpretation of 'five' is the object-language meaning of this term, and identical to the number five, while the syntactical valuation of 'five' is the numeral 'five itself, and assigned to it by a meta-linguistic formula. Carnap's explanation for why higher-order quantified formulas have "the meaning intended" appeals to only their syntactical valuation and not their material interpretation, and is in this sense a syntactic explanation.

any object-language. However, Carnap continues the above quotation as follows: "This is correspondingly true for the valuations of higher types in the wider language regions" (p. 114). Carnap's point here is that expressions in the meta-language S are assigned valuations by sentences in some meta-meta-language S_2, where sentences of S_2 are in turn assigned valuations by some meta-meta-meta-language S_3, and so on.[85] So, Carnap effectively points out that meta-linguistic sentences can be formally interpreted by means of meta-metalinguistic sentences, and so on. The rules that define Language II hence entail that higher-order quantifiers possess a formal interpretation.

Carnap's argument may seem to beg the question. One may think that the appeal to stronger and stronger meta-languages is of no help for showing that Carnap is not committed to a metaphysical viewpoint. The appeal to stronger and stronger meta-languages only makes the problem worse, by incrementally increasing the number of entities to which one is committed. Lavers (2015, pp. 274–275) raises this objection as follows:

> "Carnap, of course, denies that the view he is defending is platonistic. The reason he gives is that we can define the set of valuations for some language S, in a stronger syntax language S_2. Of course for this to work properly S_2 must be interpreted in a standard way and so only pushes the problem back a step.[86] [...] I take it few philosophers today would view this 'but the definition can be given in another language' point to successfully eliminate the worry that too strong existence assumptions are being made. This is especially true since existential assumption at least as strong have to be made concerning the domain of quantification for S_2".

I think that this objection rests on a misunderstanding of Carnap's strategy. Carnap does not simply try to avoid ontological commitments (though he does not like the term 'ontological commitment' since, he thinks, it has the wrong connotations (Carnap 1956 [1947], pp. 42–42). He rather wants to show why the specific ontological commitments he incurs are different from the acceptance of a Platonic ontology. His appeal to stronger and stronger meta-languages then aims at showing that predicate variables possess a valuation simply by virtue of how Language II is defined. It is, for this reason, analytic that higher-order quantifiers are meaningful in Language II. The assignment of syntactic valuations is therefore successful no matter what contingent reality is like and not expressive of a metaphysical viewpoint. In order to challenge this view, one would need to argue against Carnap's conception of analyticity, according to which analytic statements lack descriptive content and (for this reason) do not incur metaphysical commitments. As is well known, this point has been the subject of debate between Carnap and Quine.

---

85 Carnap here skips over problems that may arise in ordinal set theory, since even in infinite hierarchies one may eventually run out of languages.

86 Lavers says (in the passage I am omitting here) that "[t]his was pointed out to Carnap, much later, [by Beth (1963)], to which Carnap [(1963b)] agreed". However, I see no evidence for the view that Beth and Carnap discuss the problem that Lavers raises. Beth and Carnap discuss the possibility of nonstandard interpretations of some object-language (see Beth (1963, sec. VI) and Carnap (1963b, p. 928)). This issue does not obviously have anything to do with Lavers' objection, according to which the hierarchy of languages which is Language II does not discharge ontological commitments, but at each step in the hierarchy only pushes them up a level. So, I see no evidence for the view that Carnap agrees that the issue Lavers raises is a problem.

§8. Gödel's influence. Carnap's discussion of higher-order quantification yields a genuinely Carnapian defense of impredicativity. Carnap and Ramsey discussed higher-order quantification for independent reasons. Ramsey was interested in arguing that impredicative definitions are properly meaningful, while Carnap needed to show that the definitions of analyticity for the various regions of Language II do not commit him to a metaphysical viewpoint. These definitions are predicative, since any definition of analyticity in some region of Language II applies to only sentences in regions that are further down in the hierarchy. Nevertheless, Carnap's discussion suggests a defense of impredicative definitions, which unlike (for instance) a Gödelian defense does not rest on ascribing any particular ontological status to impredicatively defined properties. Impredicative formulas are meaningful in Language II since they are interpreted by means of meta-linguistic formulas that assign valuations to the free and bound variables. These meta-linguistic formulas are in turn interpreted by means of meta-metalinguistic formulas, and so on. Impredicative definitions hence are meaningful in Language II no matter what the contingent features of reality are like.

An interesting historical fact is that Carnap got the idea for this view from Gödel.[87] In an early manuscript of *The Logical Syntax of Language*, entitled *Semantics*, Carnap used what we would today call a substitutional instead of an objectual account of quantification in the definition of analyticity. Gödel saw this manuscript in 1932 (see Goldfarb (2003, p. 338)). In a letter to Carnap from September 11, 1932, he pointed out that this definition of analyticity is problematic, and suggested as an alternative to let functional variables range over any sets and relations whatsoever. He remarked that one is otherwise lead "of necessity" to ramified type theory, and its corresponding problems (Gödel, 2003, p. 347).

Gödel provides this letter with an important footnote.[88] He says (2003, p. 347): "This doesn't necessarily involve a Platonistic standpoint, for I assert only that this definition (for 'analytic') be carried out within a definite language in which one already has the concepts 'set' and 'relation'". I take it that by a 'definite' language, Gödel simply means a specific language. It is not clear that Gödel and Carnap mean the same by 'Platonism'. However, Gödel's remark clarifies how Carnap can avoid Ramsey's "Platonism". Using the concepts 'set' and 'relation' one can assign valuations to predicate and functional variables, and thus establish that higher-order quantified formulas are meaningful without making any assumptions about the ontological status of properties and relations.

In response to Gödel, Carnap first expresses puzzlement (see Gödel (2003, p. 351)): "You say: [the universal quantifier] must range over 'all sets'; but what does that mean?" Carnap here effectively says that he does not know how to formally interpret quantifiers if not substitutionally. Carnap writes again just a day later (p. 355):

> "I found the solution: The locution 'for every valuation ...' that occurs in the definition [of analyticity] can still be expressed in a semantics formulated in a definite language, namely by '[F](...)', since a valuation is of course a semantic predicate".

By a "semantic predicate" Carnap means what he would later call a "syntactical property", i.e., a set of numerical constants.[89] Carnap here sketches the interpretation of higher-order

87 For discussions of the relationship between Carnap and Gödel, see Awodey & Carus (2001, 2007, 2010), Goldfarb (2005), and Lavers (unpublished).

88 Yes, Gödel did supply his letters with footnotes.

89 Thanks to Warren Goldfarb for pointing this out.

quantification that he would later give in *The Logical Syntax of Language*. Gödel responded (p. 355): "As I gather from your second letter, you have understood my suggestions about the definition of 'analytic' entirely as I meant them".

Why did Gödel (1984 [1944]) not go Carnap's way? Not even Gödel himself was quite sure of the correct answer to this question. Gödel tried to articulate his opposition to Carnap's *Syntax* program in a article which was originally to be included in the *Schilpp* volume on Carnap.[90] However, even though Gödel drafted six versions of this article, he did not finalize a version with which he was satisfied and the article was ultimately not included in the volume.[91] Perhaps the most important argument in this series of drafts of an article is an argument based on the second incompleteness theorem, which may be summarized as follows (see p. 357):

> There is no reason to suppose that arbitrary syntactical rules would correctly predict the truth or falsity of empirical propositions. Syntactical rules hence are admissible only if it is known that they do not entail the truth or falsity of any empirical propositions. This is known only if the rules are known to be consistent. However, syntactical rules alone cannot prove the rules of syntax to be consistent.

It seems to me that Goldfarb & Ricketts's (1992) response to this objection is essentially right. From Carnap's perspective, the distinction between analytic and synthetic (or empirical) propositions can only be drawn relative to a formal system, such as Language II. Relative to such a system, the empirical propositions just are the ones that are not entailed by its rules. Hence, no consistency proof is required for showing that the rules of syntax do not entail empirical propositions. Of course, this point only shows that, from Carnap's viewpoint, Gödel's objection has no force. It leaves open the more general question of whether Carnap's non-Platonistic defense of impredicativity is preferable over Gödel's Platonistic defense from the perspective of someone who is et undecided.[92]

§9. Conclusion. Carnap provided two defenses of impredicative definitions, one in (1931) and one in (1937 [1934]). Carnap's 1931 defense of impredicative definitions is shaped by his early constructivism, which he later gave up in the wake of Gödel's discovery of the incompleteness theorems. Carnap's 1937 [1934] defense of impredicative definitions is embedded in a larger project that reconceives the decision between ramified and simple type theory as a choice between two different syntactically defined formal systems, i.e., Language I and Language II. Carnap does not deny that the use of higher-order quantifiers in Language II incurs ontological commitments. He argues, however, that the relevant ontological commitments are unproblematic. To make this point, he argues that predicate variables refer in Language II simply by virtue of how this system is defined. In this system, certain meta-linguistic formulas formally interpret higher-order quantifiers, and are in turn interpreted by meta-meta-linguistic formulas, and so on. The interpretation hence is successful, no matter what contingent features of reality may be like. It is in this sense analytic that higher-order quantifiers are meaningful in Language II.

90 See Gödel's article "Is Mathematics Logical Syntax?", in his *Collected Works*, Vol. III (1995).

91 See Awodey & Carus (2004, 2010), Goldfarb & Ricketts (1992) and Lavers (2017), as well as the further references in footnote 1 of Awodey & Carus (2004), for discussions of Gödel's (1995 [1953/1959]) objections to Carnap's *Syntax* program.

92 Thanks to Patricia Blanchette for helpful discussions related to this point.

To conclude, I would like to point out a few connections between Carnap's 1937 [1934] defense of impredicativity and his famous article "Empiricism, Semantics and Ontology" (Carnap, 1956 [1950]). First, the problem that Carnap faces in the 1956 article is closely related to the problem he faces in relation to impredicative definitions. He wants to show that "using [. . .] a language [referring to abstract entities] does not imply embracing a Platonic ontology" (1956 [1950], p. 206). This problem is directly analogous to Carnap's (1937 [1934], p. 114) problem of showing that his use of higher-order quantifiers does not commit him to Ramsey's "Platonism". Furthermore, at least parts of Carnap's solutions in the two cases are analogous as well. Carnap (1956 [1950], p. 217) argues that rules of designation are analytic: "Generally speaking, any expression of the form "[X]' designates [X]' is an analytic statement provided the term '[X]' is a constant in an accepted framework". Carnap here effectively says that it is analytic in the number framework that 'five' designates five. This view is directly analogous to his 1934 conception, according to which it is analytic in Language II that higher-order quantifiers are meaningful.

These observations about the parallels between Carnap's earlier and his later view suggests that his famous internal/external distinction was deeply shaped by his reflections on pure logic and the foundations of mathematics. Understanding Carnap's views on the foundations of mathematics hence holds the key for the best understanding of his influential notion of a "linguistic framework" and his distinction between "internal" and "external" questions.

§10. Acknowledgments. I would like to thank several anonymous referees, Neil Barton, Patricia Blanchette, Tim Button, David Chalmers, Christian Damböck, Cian Dorr, Sean Ebels-Duggan, Hartry Field, Kit Fine, Warren Goldfarb, Harold Hodes, Thomas Hofweber, Nicholas K. Jones, Robert Kraut, Lydia Patton, Ian Rumfitt, Erich Reck, Sam Roberts and Chris Scambler for helpful comments, suggestions and discussions. I would like to thank especially Gary Ebbs for numerous helpful discussions of both big picture and detail questions, as well as Harold Hodes and Anja Jauernig for very careful and detailed comments on the penultimate draft. Further thanks to the members of the ABD seminar at NYU in Fall 2017, and to the audiences at my talks at the Existence Conference in Uppsala in 11/2016, at the Tenth Annual Cambridge Graduate Conference on the Philosophy of Mathematics and Logic in 01/2017, at the Sixth Annual Meeting of the Society for the Study of the History of Philosophy (SSHAP) in 05/17, at the Munich Center for Mathematical Philosophy (MCMP) in 06/2017, and at the 18th annual Midwest Workshop on the Philosophy of Mathematics (MWPMW 18) at the University of Notre Dame in 10/17.

BIBLIOGRAPHY

Awodey, S. & Carus, A. (2001). Carnap, Completeness, and Categoricity: The *Gabelbarkeitssatz* of 1928. *Erkenntnis*, **54**(2), 145–172.

Awodey, S. & Carus, A. (2004). How Carnap could have replied to Gödel. In Awodey, S. and Klein, C., editors. *Carnap Brought Home*: *The View from Jena*. Chicago: Open Court, pp. 203–224.

Awodey, S. & Carus, A. (2007). Carnap's Dream: Gödel, Wittgenstein and *Logical Syntax*. *Synthese*, **159**, 23–46.

Awodey, S. & Carus, A. (2010). Gödel and Carnap. In Feferman, S., Parsons, C., and Simpson, S. G., editors. *Kurt Gödel*: *Essays for his Centennial*. New York: Cambridge University Press, pp. 252–274.

Bernays, P. (1983 [1935]). On Platonism in Mathematics. In Benacerraf, P. and Putnam, H., editors. *Philosophy of Mathematics*: *Selected Readings*. Cambridge: Cambridge University Press, pp. 258–271.
Beth, E. W. (1963). Carnap's Views on the Advantages of Constructed Systems Over Natural Languages in the Philosophy of Science. In Schilpp, P. A., editor. *The Philosophy of Rudolf Carnap*. LaSalle, IL: Open Court and Cambridge University Press, pp. 469–502.
Cantor, G. (1895). Beiträge zur Begründung der transfiniten Mengenlehre. *Mathematische Annalen*, **46**, 481–512.
Carnap, R. (1927). Eigentliche und Uneigentliche Begriffe. *Symposion*: *Philosophische Zeitschrift für Forschung und Aussprache*, **1**, 355–374.
Carnap, R. (1929). *Abriss der Logistik. Mit besonderer Berücksichtigung der Relationstheorie und ihrer Anwendungen*. Wien: Springer.
Carnap, R. (1930). Die Mathematik als Zweig der Logik. *Blätter für deutsche Philosophie*, **4**, 298–310.
Carnap, R. (1930/1931). Bericht über Untersuchungen zur allgemeinen Axiomatik. *Erkenntnis*, **1**, 303–307.
Carnap, R. (1931). Die Logizistische Grundlegung der Mathematik. *Erkenntnis*, **2**, 91–105.
Carnap, R. (1937 [1934]). *The Logical Syntax of Language*. London: Routledge.
Carnap, R. (1956 [1947]). *Meaning and Necessity. A Study in Semantics and Modal Logic* (second edition). Chicago: The University of Chicago Press.
Carnap, R. (1956 [1950]). Empiricism, Semantics, and Ontology. In *Meaning and Necessity*. Chicago: The University of Chicago Press, pp. 205–221.
Carnap, R. (1963a). Carnap's Intellectual Autobiography. In Schilpp, P. A., editor. *The Philosophy of Rudolf Carnap*. La Salle, IL: Open Court and Cambridge University Press, pp. 3–84.
Carnap, R. (1963b). E. W. Beth on Constructed Language Systems. In Schilpp, P. A. editor. *The Philosophy of Rudolf Carnap*. La Salle, IL: Open Court and Cambridge University Press, pp. 927–933.
Carnap, R. (1998 [1928]). *Der Logische Aufbau der Welt*. Hamburg: Meiner.
Carnap, R. (2000). *Untersuchungen zur Allgemeinen Axiomatik*. Darmstadt: Wissenschaftliche Buchgesellschaft.
Carnap, R. (2004 [1928]). Scheinprobleme in der Philosophie. In Mormann, T., editor. *Scheinprobleme in der Philosophie und Andere Metaphysikkritische Schriften*. Hamburg: Felix Meiner, pp. 3–48.
Carnap, R. (2004 [1930]). Die Alte und die Neue Logik. In Mormann, T., editor. *Scheinprobleme in der Philosophie und Andere Metaphysikkritische Schriften*. Hamburg: Meiner, pp. 63–80.
Carnap, R. (forthcoming). *Tagebücher 1908–1935*. Hamburg: Meiner.
Cartwright, R. L. (1994). Speaking of Everything. *Noûs*, **28**, 1–20.
Chihara, C. S. (1973). *Ontology and the Vicious Circle Principle*. Ithaca, NY: Cornell University Press.
Church, A. (1940). A Formulation of the Simple Theory of Types. *The Journal of Symbolic Logic*, **5**, 56–68.
Chwistek, L. (2012 [1922]). The Principles of the Pure Type Theory. *History and Philosophy of Logic*, **33**, 329–352.
Coffa, A. (1993). *The Semantic Tradition from Kant to Carnap*. Cambridge: Cambridge University Press.
Curry, H., Feys, R., & Craig, W. (1958). *Combinatory Logic*. Amsterdam: North-Holland.

Dedekind, R. (1963 [1872]). Continuity and Irrational Numbers. In *Essays on the Theory of Numbers*. New York: Dover, pp. 1–30.

Ebbs, G. (2001). Carnap's Logical Syntax of Language. In Gaskin, R., editor. *Grammar in Early Twentieth-Century Philosophy*. London: Routledge, pp. 218–237.

Feferman, S. (2000). The Significance of Hermann Weyl's "Das Kontinuum". In Hendricks V. F., Pedersen, S. A., and Jørgensen, K. F., editors. *Proof Theory: History and Philosophical Significance.* Dordrecht: Springer, pp. 179–194.

Feferman, S. (2005). Predicativity. In Shapiro, S., editor. *The Oxford Handbook of the Philosophy of Mathematics and Logic*. New York: Oxford University Press, pp. 590–624.

Flocke, V. (forthcoming). Carnap's Noncognitivism about Ontology. *Noûs,* doi: 10.1111/nous.12267.

Gödel, K. (1931). Diskussion zur Grundlegung der Mathematik am Sonntag, dem 7. September 1930. *Erkenntnis*, **2**, 135–151.

Gödel, K. (1984 [1944]). Russell's Mathematical Logic. In Benacerraf, P. and Putnam, H. editors. *Philosophy of Mathematics*: *Selected Readings*. Cambridge: Cambridge Unversity Press, pp. 447–469.

Gödel, K. (1995). *Collected Works. Vol. III, Unpublished Essays and Lectures* (edited by S. Feferman, J. W. Dawson, W. Goldfarb, C. Parsons, and R. N. Solovay). New York: Oxford Unversity Press.

Gödel, K. (1995 [1953/1959]). Is Mathematics Syntax of Language? In Feferman, S., Dawson, J. W., Goldfarb, W., Parsons, C., and Solovay, R. N., editors. *Collected Works. Vol. III, Unpublished Essays and Lectures*. New York: Oxford Unversity Press, pp. 334–363.

Gödel, K. (2003). *Collected Works. Vol. IV, Correspondence A-G* (edited by S. Feferman, J. W. Dawson, W. Goldfarb, C. Parsons, and R. N. Solovay). Oxford: Clarendon Press.

Goldfarb, W. (1989). Russell's Reason for Ramification. In Savage, C. W. and Anderson, C. A., editors. *Rereading Russell*: *Essays on Bertrand Russell's Metaphysics and Epistemology*. Minneapolis, MN: Unversity of Minnesota Press, pp. 24–40.

Goldfarb, W. (1996). The Philosophy of Mathematics in Early Positivism. In Giere, R. N. and Richardson, A. W., editors. *Minnesota Studies in the Philosophy of Science*, Vol. XVI. Minneapolis, MN: University of Minnesota Press, pp. 213–230.

Goldfarb, W. (1997). Semantics in Carnap: A Rejoinder to Alberto Coffa. *Philosophical Topics*, **25**, 51–66.

Goldfarb, W. (2003). Rudolf Carnap: Introductory Note by Warren Goldfarb. In Feferman, S., Dawson, Jr., J. W., Goldfarb, W., Parsons, C., and Sieg, W. editors. *Kurt Gödel*: *Collected Works. Vol. IV, Correspondence A-G*. Oxford: Clarendon Press, pp. 335–341.

Goldfarb, W. (2005). On Gödel's Way In: The Influence of Rudolf Carnap. *The Bulletin of Symbolic Logic*, **11**, 185–193.

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

Hindley, J. R. (1997). *Basic Simple Type Theory*. New York: Cambridge University Press.

Hodes, H. T. (2013). A Report on Some Ramifield-Type Assignment Systems and their Model-Theoretic Semantics. In Griffin, N. and Linky, B., editors. *The Palgrave Centenary Companion to Principia Mathematica*. Basingstoke: Palgrave Macmillan, pp. 305–336.

Hodes, H. T. (2015). Why Ramify? *Notre Dame Journal of Formal Logic*, **56**, 379–415.

Hodges, W. (2014). Tarski's Truth Definitions. In Zalta, E. N., editor. *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition). Available at: http://plato.stanford.edu/archives/fall2014/entries/tarski-truth/.

Hylton, P. (2005). *Propositions, Functions and Analysis*: *Selected Essays in Russell's Philosophy*. New York: Oxford University Press.

Iemhoff, R. (2016). Intuitionism in the Philosophy of Mathematics. In Zalta, E. N., editor. *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition). Available at: https://plato.stanford.edu/entries/intuitionism/.

Kaufmann, F. (1978 [1930]). *The Infinite in Mathematics. Logico-Mathematical Writings*. Dordrecht: Reidel.

Köhler, E. (1991). Gödel und der Wiener Kreis. In Kruntorad, P., editor. *Jour Fixe der Venunft. Der Wiener Kreis und die Folgen*. Wien: Verlag Hölder-Pichler-Tempsky, pp. 127–158.

Lavers, G. (2015). Carnap, Quine, Quantification and Ontology. In Torza, A., editor. *Quantifiers, Quantifiers, and Quantifiers*: *Themes in Logic, Metaphysics and Language*. Cham: Springer, pp. 271–300.

Lavers, G. (unpublished). Hitting a Moving Target: Gödel, Carnap, and Mathematics as Logical Syntax. Unpublished Manuscript.

Limbeck-Lilienau, C. & Stadler, F. (2015). *Der Wiener Kreis. Texte und Bilder zum Logischen Empirismus*. Wien: LIT Verlag.

Linnebo, Ø. (2004). Predicative Fragments of Frege Arithmetic. *The Bulletin of Symbolic Logic*, **10**, 153–174.

Linnebo, Ø. (2017). Predicative and Impredicative Definitions. *The Internet Encyclopedia of Philosophy*. Available at: https://www.iep.utm.edu/predicat/.

Linnebo, Ø. (unpublished). Generality Explained. Unpublished Manuscript.

Linsky, B. (1999). *Russell's Metaphysical Logic*. Stanford, CA: CSLI Publications.

Myhill, J. (1979). A Refutation of an Unjustified Attack on the Axiom of Reducibility. In Russell, B. and Roberts, G. W., editors. *Bertrand Russell Memorial Volume*. New York: Humanities Press, pp. 81–90.

Quine, W. (1953 [1948]). On What There Is. In *From a Logical Point of View*. Cambridge, MA: Harvard University Press, pp. 1–19.

Quine, W. (1960 [1954]). Carnap and Logical Truth. *Synthese*, **12**, 350–374.

Ramsey, F. P. (1931 [1926]). The Foundations of Mathematics. In Braithwaite, R. B., editor. *The Foundations of Mathematics and other Logical Essays*. London: Routledge, pp. 1–61.

Reck, E. H. (2007). Carnap and Modern Logic. In Friedman, M. and Creath, R., editors. *The Cambridge Companion to Carnap*. Cambridge: Cambridge University Press, pp. 176–199.

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

Russell, B. (1908). Mathematical Logic as Based on the Theory of Types. *American Journal of Mathematics*, **30**, 222–262.

Russell, B. & Whitehead, A. N. (1927 [1910]). *Principia Mathematica,* Vol I (second edition). Cambridge: Cambridge University Press.

Schiemer, G., Zach, R., & Reck, E. (2017). Carnap's Early Metatheory: Scope and Limits. *Synthese*, **194**, 33–65.

Tarski, A. (1956 [1936]). The Concept of Truth in Formalized Languages. In Corcoran, J, editor. *Logic, Semantics, Metamathematics*: *Papers from 1923 to 1938*. London: Oxford at the Clarendon Press, pp. 152–278.

Tarski, A. (2002 [1936]). On the Concept of Following Logically. *History and Philosophy of Logic*, **23**, pp. 155–196.

Verein, E. M. (1929). *Wissenschaftliche Weltauffassung. Der Wiener Kreis*. Wien: Veröffentlichungen des Vereines Ernst Mach.

Wagner, P. (2009). *Carnap's Logical Syntax of Language*. Basingstoke: Palgrave Macmillan.

Weyl, H. (1918). *Das Kontinuum. Kritische Untersuchungen über die Grundlagen der Analysis*. Leipzig: Verlag von Veit.

Wittgenstein, L. (1921). Logisch-Philosophische Abhandlung. *Annalen der Naturphilosophie*, **14**, 185–262.

DEPARTMENT OF PHILOSOPHY
NEW YORK UNIVERSITY
5 WASHINGTON PLACE
NEW YORK, NY, 10003, USA
E-mail: vera.flocke@nyu.edu
URL: veraflocke.com