Carnap on the Foundations of Logic and Mathematics*

Peter Koellner

Throughout most of his philosophical career Carnap upheld and defended three distinctive philosophical positions: (1) The thesis that the truths of logic and mathematics are analytic and hence without content and purely formal. (2) The thesis that radical pluralism holds in pure mathematics in that any consistent system of postulates is equally legitimate and that there is no question of justification in mathematics but only the question of which system is most expedient for the purposes of empirical science. (3) A minimalist conception of philosophy in which most traditional questions are rejected as pseudo-questions and the task of philosophy is identified with the metatheoretic study of the sciences.

In this paper, I will undertake a detailed analysis of Carnap's defense of the first and second thesis. This will involve an examination of his most technical work *The Logical Syntax of Language* (1934), along with the monograph "Foundations of Logic and Mathematics" (1939). These are the main works in which Carnap defends his views concerning the nature of truth and radical pluralism in mathematics. The first work is quite subtle and sophisticated from a metamathematical point of view[1] and it is here that one finds Carnap's most sustained defense of the above theses. However, I will argue that the discussion is not quite subtle enough and that in the end, the defense of the first two theses is undermined by a series of metamathematical subtleties

*I would like to thank Michael Friedman, Warren Goldfarb, and Thomas Ricketts for helpful discussion.

[1]This work presents the state of the art in metamathematics in 1934. For example: In §§18–23 there is a detailed discussion of the arithmetization of syntax, in §§35–36 the incompleteness theorems are sketched, in §59 there is an accurate treatment of consistent but ω-inconsistent languages, in §60c the analogue of Tarski's theorem for analyticity is proved, §60d contains a discussion of the inexhaustibility of mathematics, and §71d contains an illuminating discussion of the Skolem paradox.

that Carnap did not fully appreciate. This will lead us to consider the second work—from Carnap's so-called semantic period—to see if these problems are remedied. We will see that many of them persist and that when one sifts through the details, what remains in the end is merely a proposal to adopt an instrumentalist conception of mathematics. I will close by suggesting that Carnap's conception of philosophy is too thin and that it is through a more meaningful engagement between philosophy and the exact sciences that we can most fruitfully approach the foundational disputes.

There have been many recent excellent studies that place Carnap's work on the foundations of logic in mathematics in their historical context.[2] The purpose of the present study, however, is quite different. My aim is to asses Carnap's work in the foundations of logic and mathematics from a contemporary perspective, in light of what we now know about the foundations of mathematics and (to a lesser degree) the foundations of physics. In addition to its intrinsic historical interest, a study of Carnap's views on the foundations of logic and mathematics is of contemporary relevance. For views quite similar to his are gaining wide currency today. By locating the problematic points in one of the most detailed accounts I believe that we can gain insight into the tenability of such views. We shall see that mathematics is rather resilient with respect to attempts to prove that it is fleeting.

1 Three Central Theses

In *Logical Syntax*,[3] Carnap defends three distinctive philosophical theses: (1) The thesis that the truths of logic and mathematics are analytic and hence without content and purely formal. (2) The thesis that radical pluralism holds in pure mathematics in the following sense: There is no question of justifying one system of pure mathematics over another; all consistent systems of pure mathematics are on an equal footing from the point of view of justification; the choice between conflicting systems is one of "mere expedience". (3) The thesis that most traditional philosophical questions are pseudo-questions and that the proper task of philosophy is the study of the logic of science.

[2]See, for example, the essays in Wagner (2009). The essays by Uebel and Awodey and Carus set the general historical context for *Logical Syntax*, while the remaining essays focus on more specific aspects.

[3]In what follows I will abbreviate '*The Logical Syntax of Language*' as '*Logical Syntax*'.

In this section I will elaborate these positions[4] by providing textual support for them.[5]

1.1 Nature of Truth in Mathematics

The fundamental difference, according to Carnap, between mathematical truths and empirical truths is that the latter are synthetic and have content while the former are analytic, without content, and purely formal.

> The latter [synthetic, empirical sentences], the so-called 'real' sentences, constitute the core of science; the mathematico-logical sentences are analytic, with no real content, and are merely formal auxiliaries. (xiv)

> An analytic sentence is not actually 'concerned with' anything, in the way that an empirical sentence is; for an analytic sentence is without content. (7)

> In material interpretation, an analytic sentence is absolutely true whatever the empirical facts may be. ... *Synthetic sentences* are the *genuine statements about reality.* (41)

One of the central aims of *Logical Syntax* is to establish a *formal* criterion of mathematical truth:

> One of the chief tasks of the logical foundations of mathematics is to set up a formal criterion of validity, that is, to state necessary and sufficient conditions which a sentence must fulfil in order to be valid (correct, true) in the sense understood in classical mathematics. (98)

[4]I have called these three positions 'theses' since at the outset Carnap speaks of them as theses that he aims to defend. However, as we proceed, we shall see an alternative interpretation emerge, according to which Carnap is not trying to defend theses but rather he is merely trying to persuade us to adopt proposals. Until that interpretation emerges, I will continue to call the above positions 'theses'.

[5]In my thinking about *Logical Syntax* I have benefited from Friedman (1999c), Gödel (1953/9) and Goldfarb & Ricketts (1992). After writing my account of *Logical Syntax* Warren Goldfarb drew my attention to Kleene's review (Kleene (1939)). I am in complete agreement with what Kleene has to say and there is some overlap between our discussions, though my discussion is much more comprehensive.

Carnap thought that he had succeeded in this task. For after describing his basic approach he writes:

> In this way a *complete criterion of validity for mathematics* is obtained. We shall define the term 'analytic' in such a way that it is applicable to all those sentences, and only those sentences, of Language II [a system to be described below] that are valid (true, correct) on the basis of logic and classical mathematics. (100-101)

It is important that the criterion be *formal* since this is what severs the tie linking mathematical truth to some supposed subject matter:

> We shall see that the logical characteristics of sentences (for instance, whether a sentence is analytic, synthetic, or contradictory...) ... are solely dependent upon the syntactical structure of the sentences. (1–2)

> But the decisive point is the following: *in order to determine whether or not one sentence is a consequence of another, no reference need be made to the meaning of the sentences. ... It is sufficient that the syntactical design of the sentences be given.* (258)

Thus the definition of analyticity plays a central role in Carnap's system. For it is through the definition of analyticity that one arrives at a complete criterion of mathematical truth; moreover, this criterion has the features that in contrast to empirical truths, logical and mathematical truths are free of content and their status as truths is secured not in virtue of any particular subject matter but in virtue of syntactic form alone.

1.2 Radical Pluralism in Mathematics

The thesis that logical and mathematical truths are free of content and that mathematical truth rests on syntactic form alone suggests that radical pluralism holds in mathematics. For by the second feature—that mathematical truth rests solely on syntactic form alone—it should be possible to give the formal criterion of truth for two apparently incompatible syntactical systems. And by the first feature—that mathematical truths are free of content—when one does this, one will have demonstrated that in both cases the truths of each system are free of content. Having thus shown that there is no contentual conflict between the two systems, one will have erased the apparent conflict and have legitimated each as standing on a par.

Here is what Carnap says about the second thesis—that radical pluralism holds in mathematics:

> [L]et any postulates and any rules of inference be chosen arbitrarily; then this choice, whatever it may be, will determine what meaning is to be assigned to the fundamental logical symbols. (xv)

In this way, the traditional foundational disputes are seen to be pseudo-disputes:

> By this method, also, the conflict between the divergent points of view on the problem of the foundations of mathematics disappears. For language, in its mathematical form, can be constructed according to the preferences of any one of the points of view represented; so that no question of justification arises at all, but only the question of the syntactical consequences to which one or other of the choices leads, including the question of non-contradiction. (xv)

> The first attempts to cast the ship off from the *terra firma* of the classical forms were certainly bold ones, considered from the historical point of view. But they were hampered by the striving after 'correctness'. Now, however, that impediment has been overcome, and before us lies the boundless ocean of unlimited possibilities. (xv)

This metaphor suggests that Carnap's pluralism is quite radical. And there is a good deal of evidence that it is quite radical indeed. For example, consider §34h where Carnap discusses the Principle of Complete Induction and the Principle of Selection (the Axiom of Choice (AC)). He shows that his definition of 'analyticity' is such that one can show in the metalangauge that these principles are analytic. The proof, of course, involves presupposing versions of these principles in the metalangauge. For this reason, Carnap notes:

> The proofs ... must not be interpreted as though by means of them it were proved that the Principle of Induction and the Principle of Selection were materially true. They only show that our definition of 'analytic' effects on this point what it is intended to effect, namely, the characterization of a sentence as analytic *if*, in material interpretation, *it is regarded as logically valid*. (124, my emphasis)

The last part of this quotation reveals that there is an element of freedom: We are free to regard these principles as logically valid and we are free to regard their negations as logically valid.

> The question as to whether the Principle of Selection should be admitted into the whole of the language of science ... as logically valid or not is not decided thereby. That is a matter of choice .... (124)

There is nothing in this reasoning that is tied to the Principle of Selection alone. It also applies to the Principle of Induction and indeed any statement of mathematics, including, for example, very simple statements of arithmetic such as Π^0_1-statements. In each case, the question of analyticity is passed on to the formal system underlying the metalanguage and, with regard to that system, one has complete reign. There is no question of truth or justification, only a question of "expedience". So it appears that Carnap is quite serious when he says that "*any* postulates and any rules of inference [may] be chosen arbitrarily" (xv, my emphasis);

On this conception of truth in mathematics, the traditional foundational disputes are dissolved. For on this conception any consistent system is equally legitimate and the questions of truth and justification do not arise when selecting between such systems; the question of truth only arises when addressing questions of the form "Does system S imply φ?". In this way, Carnap hopes to steer practitioners away from fruitless disputes concerning which system is the "correct" system of mathematics, and relocate their efforts at the level of metatheory.

1.3 Minimalist Conception of Philosophy

The above position concerning foundational disputes in mathematics is a special case of a much more general position. In general, Carnap upholds a minimalist conception of philosophy according to which most traditional philosophical questions are regarded as pseudo-questions and the task of philosophy is identified with the study of the logic of science.

> Apart from the questions of the individual sciences, only the questions of the logical analysis of science, of its sentences, terms, concepts, theories, etc., are left as genuine scientific questions. . . . *[T]he logic of science takes the place of the inextricable tangle of problems which is known as philosophy.* (279)

In §72 he gives an informal account of what he means by 'the logic of science' and in §73 he sharpens this by arguing that "the logic of science is the syntax of the language of science".

Thus, just as the fruitless foundational disputes in mathematics are to be dissolved and efforts are to be relocated within the study of the logical syntax of mathematical systems, so too the fruitless traditional disputes in philosophy are to be dissolved and efforts are to be relocated at the metalevel in the study of the logical syntax of science.

In our statement of Carnap's three theses we have employed numerous technical terms—'analytic', 'content', 'formal', 'syntactical consequence', 'logical syntax'. To sharpen our understanding of the content of these theses we must now examine Carnap's analysis of these terms.

2 Technical Results

The analysis of the notions of analyticity and content lies at the heart of Carnap's first two theses—the thesis concerning the nature of mathematical truth and the thesis concerning radical pluralism in mathematics. In this section we will give an exposition of Carnap's analysis of these and related notions. In the next section we will critique both his analysis and the philosophical claims he makes on its basis.

2.1 Preliminaries

It is necessary to begin with a few terminological remarks. This is because in some cases Carnap uses unfamiliar terms for familiar notions (for example, he uses the terms 'definite' and 'indefinite' for the notions of *recursive* and *non-recursive*) and in other cases he uses familiar terms in a way that is out of step with modern usage (for example, in some instances he uses 'language' and 'syntax' where today we would use 'formal system' and 'semantics').

2.1.1 D-notions and C-notions

For Carnap, a *definite* notion is one that is recursive,[6] (such as "is a formula" and "is a proof of φ") and an *indefinite* notion is one that is non-recursive (such as "is an ω-consequence of PA" and "is true in $V_{\omega+\omega}$"). This distinction is of central importance for Carnap. It leads to a distinction between the *method of derivation* (or *d-method*), which investigates the semi-definite (recursively enumerable) metamathematical notions, such as *demonstrable*, *derivable*, *refutable*, *resoluble*, and *irresoluble*, and the *method of consequence* (or *c-method*), which investigates the (typically) non-recursively enumerable metamathematical notions such as *consequence*, *analytic*, *contradictory*, *L-determinate*, and *synthetic*.

2.1.2 Language

For Carnap a *language* is "any sort of calculus, that is to say, a system of formation and transformation rules concerning what are called *expressions*" (167). In modern usage, this would, of course, be called a 'formal system'. This deviant usage of the term 'language' is quite entrenched in *Logical Syntax*. For example, the two central formal systems treated in that work (and to be described in detail below) are called 'Language I' and 'Language II'.

The rules of formation and transformation of most languages (including all of those that Carnap considers) are definite (recursive). However, the notion of a language (formal system) is understood in a broad sense in that it *may* include indefinite rules of formation and transformation (such as the ω-rule (172)). But it is important to note that although a language may

[6] Carnap did not know as much about the notion of being recursive as we do today, especially after the insightful analysis of Turing. Carnap actually gives two conflicting definitions of the notion of definiteness: on p. 11 he speaks of properties "of which the possession or non-possession by any number whatsoever can be determined in a finite number of steps according to a fixed method". Here he evidently has in mind the notion of being recursive. However, later on p. 45 when he gives an alternative formulation of the notion of definiteness, he describes the notion of being Δ^0_0. The two notions, of course, do not coincide—the syntactic notion corresponding to the notion of being recursive is Δ^0_1, not Δ^0_0. In what follows, I shall take 'definite' to mean recursive, as this seems to more closely capture Carnap's intentions.

involve indefinite d-notions the indefinite c-notions of a language are not part of the language; in fact, a language *cannot* include its own c-notions by Theorem 60c.1, an analogue for analyticity of Tarski's theorem on the undefinability of truth.[7] The c-notions for a given language (formal system) must be treated in an essentially richer metalanguage (metasystem).

2.1.3 Logical Syntax

We come finally to the term 'syntax', which is the most problematic term in Carnap. Carnap uses this term in both a *narrow* sense and a *wide* sense.

In the narrow sense of the term, the study of syntax is just the combinatorial study of the definite rules of formation and transformation in a formal system. There are points where Carnap uses the term in this narrow sense. This is true, for example, in the following two passages:

> The *syntax* of a language, or of any other calculus, is concerned, in general, with the *structures of possible serial orders* (of a definite kind) *of any elements whatsoever.* (6)
>
> Pure syntax is thus wholly analytic, and is nothing more than *combinatorial analysis*, or, in other words, the *geometry* of finite, discrete, serial structures of a particular kind. (7)[8]

Let us call this 'syntax in the narrow sense'. It coincides with what we mean today in metamathematics when we use the terms 'syntax' and 'proof-theoretic' .

Carnap is quite clear that syntax in general has much wider scope. For the study of syntax also includes the c-notions of a language, such as the notions of consequence and analyticity.

[7] Carnap is somewhat sloppy on this point. For at times he speaks as though the c-notions are part of the language (see, for example, p. 231) through he knows full well that because of Theorem 60c.1, this is impossible.

[8] Here *pure* syntax is to be contrasted with *descriptive* syntax. *Descriptive* syntax "is concerned with the syntactical properties and relations of empirically given expressions (for example, with the sentences of a particular book)". (7) *Pure* syntax is the more abstract study that is concerned with "the possible arrangements" without reference to "the question as to which of the possible arrangements of these elements are anywhere realized" (for example, whether "they exist on paper somewhere in the world") (7). In metamathematics one is concerned with pure syntax.

> In the treatment of Languages I and II we introduced the term 'consequence' only at a late stage. *From the systematic standpoint, however, it is the beginning of all syntax. If for any language the term 'consequence' is established, then everything that is said concerning the logical connections within this language is thereby determined.* (168)

It follows (from Theorem 60c.1) that the syntax (in this wide sense) of a language cannot be formulated within that language. And it is for this reason that Carnap distinguishes between the object language and what he calls the 'syntax language', which is just what today we would call the 'metalanguage' (or the 'metasystem') (4). Let us call this 'syntax in the wide sense'. It includes many notions that today in metamathematics we would call 'semantic' or 'model-theoretic'.

There are points where Carnap is clear about the distinction between syntax in the narrow sense and syntax in the wide sense. For example, in §18 he says that he shall formulate the syntax of Language I "as far as it is definite" in Language I itself.[9] This qualification is necessary since (as noted above), in the wide sense of the term, the syntax of a language cannot be formulated in that language. For a second example, consider Carnap's introduction of the term 'logical syntax'. He adds the qualifier 'logical' to emphasize his view that logic (by which he includes most of mathematics) is a part of syntax "provided the latter is conceived in a sufficiently wide sense" (2). In saying this he is once again showing sensitivity to the importance of the distinction between narrow and wide syntax.

There are also places where Carnap appears to slide from syntax in the narrow sense to syntax in the wide sense. For example, the first paragraph of the book reads:

> By the *logical syntax* of a language, we mean the formal theory of the linguistic forms of that language—the systematic statement of the formal rules which govern it together with the development of the consequences which follow from these rules. (1)

Here it appears that by 'the systematic statement of the formal rules which govern a language' he means the d-notions and by 'the consequences which

[9] He then gives a detailed account of Gödel's technique of the arithmetization of syntax. Here Language I is a version of Primitive Recursive Arithmetic (see below).

follow from these rules' he means the c-notions. However, the latter do not *follow from* the former.[10]

In what follows I will be careful to distinguish between syntax in the narrow sense and syntax in the wide sense. The distinction will play an important role in our critical discussion.

To summarize: 'Language' is often used to mean 'formal system' and 'syntax' is often used in a wide sense that includes semantic notions. In particular, the 'logical syntax' of a language (formal system) is the study of the c-notions for that language and this study takes place in an essentially richer metalanguage (metasystem).[11]

2.2 Nature of Mathematical Truth

With these preliminaries out of the way, let us now turn to Carnap's analysis of the notions of 'logico-mathematical', 'descriptive', 'analyticity', and 'content'. The analysis is supposed to have the outcome that the truths of logic and mathematics are analytic and hence without content, while the truths of the empirical sciences (descriptive truths) are synthetic and have content. Our aim is to determine whether there is any substance in this account.

Carnap actually has two approaches to this cluster of fundamental notions. The first approach occurs in his discussion of specific languages—Languages I and II. Here he starts with a division of primitive terms into 'logico-mathematical' and 'descriptive',[12] and upon this basis he defines the notions of analytic and synthetic. The second approach occurs in the discussion of general syntax. Here Carnap reverses the procedure—he starts with an arbitrary relation of direct consequence and uses it to define the other c-notions, along with the division of primitive terms into 'logico-mathematical' and 'descriptive'.

[10] As noted above there are places where Carnap is clear about this.

[11] Later, during his semantic period, Carnap would take care to distinguish between 'language' and 'formal system' and between 'syntax' (syntax in the narrow sense) and 'semantics' (syntax in the wide sense). Cf. Section 4 below.

[12] More precisely, he presents open-ended languages, specifies the logico-mathematical terms, and deems all additional terms one might add descriptive. Carnap uses the term 'logical' for what I am calling 'logico-mathematical'. The latter term is more apt since many of the expressions he classifies under this label are what we would normally call mathematical (for example, the primitive symbols of arithmetic and set theory). As will become apparent in §3.1 Carnap's choice of the term 'logical' is doing some mischief.

2.2.1 The First Approach

In the first approach, Carnap introduces two languages—Language I and Language II. It is important to note (as we have above) that here by 'language' Carnap means what we would call a 'formal system'. The background languages (in the modern sense) of Language I and Language II are perfectly general—they may include expressions that we would informally label 'descriptive', expressions such as 'red', 'neutron', etc. Carnap begins by classifying the primitive terms into 'logico-mathematical' and 'descriptive'. The expressions that he classifies 'logico-mathematical' are exactly those which, in informal discourse, we would label as logical or mathematical; the remaining expressions are classified as 'descriptive'. On the side of formal systems, Language I is a version of PRA and Language II is a version of finite type theory built over PA.[13] The d-notions for these languages are the standard proof-theoretic ones. So let us concentrate on the c-notions.

For Language I, Carnap starts with a consequence relation based on two rules—(i) the rule that allows one to infer φ if $T \vdash \varphi$ (where T is some fixed Σ^0_1-complete formal system) and (ii) the ω-rule. It is then easily seen that one has a complete theory for the logico-mathematical fragment, that is, for any logico-mathematical sentence φ, either φ is a consequence of the null set or $\neg\varphi$ is a consequence of the null set..[14] The other c-notions are then defined in the standard fashion. For example, a sentence is *analytic* if it is a consequence of the null set; *contradictory* if its negation is analytic; etc.

For Language II, Carnap starts by defining analyticity.[15] From the mod-

[13]The intended model is essentially $V_{\omega+\omega}$ (except that V_ω is replaced by $\mathbb{N}$) and the axiom system that Carnap gives is easily seen to be mutually interpretable with ZC, that is, ZFC − Replacement.

[14]Quine (1963) says that for Language I, "Carnap's formulation of logical truth is narrowly syntactical in the manner of familiar formulations of logical systems by axioms and rules of inference", but that "Gödel's proof of the incompletability of elementary number theory shows that no such approach can be adequate to mathematics in general" (Section VII). This is incorrect. Carnap's approach to Language I is not narrowly syntactical. He invokes the ω-rule and he even proves the completeness theorem for ω-logic that we mention in the text (Theorem 14.3).

[15]The definitions that Carnap gives has an interesting history. While he was writing *Logical Syntax*, Carnap discussed the subject with Gödel. Gödel pointed out a problem with Carnap's original attempt to define analyticity (an approach that followed the pattern of the approach used for Language I but ran into problems because of impredicativty), and he suggested his own approach: In a letter of Sept. 11, 1932 Gödel says that "in the second part of my work I will give a definition of 'truth'" and in a letter of Nov. 28, 1932

ern vantage point, it is seen to be a notational variant of the Tarskian truth definition but with *one important difference*—namely, it involves an asymmetric treatment of the logico-mathematical and descriptive expressions.[16] For the logico-mathematical expressions his definition really just is a notational variant of the Tarskian truth definition. But descriptive expressions must pass a more stringent test to count as analytic—they must be such that if one replaces all descriptive expressions in them by variables of the appropriate type, then the resulting logico-mathematical expression is analytic, i.e. true.[17] In other words, to count as analytic, a descriptive expression must be a substitution-instance of a general logico-mathematical truth. With this definition in place, the other c-notions are defined in the standard fashion.

The *content* of a sentence is defined to be the set of its non-analytic consequences. It then follows immediately from the definitions that logico-mathematical sentences (of both Language I and Language II) are analytic or contradictory and (assuming consistency) that analytic sentences are without content.

2.2.2 The Second Approach

In the second approach, for a given language, Carnap starts with an *arbitrary* notion of direct consequence[18] and from this notion he defines the other c-notions in the standard fashion. More importantly, in addition to defining the other c-notions, Carnap also uses the primitive notion of direct consequence (along with the derived c-notions) to effect the classification of terms into 'logico-mathematical' and 'descriptive'. The guiding idea is that "the formally expressible distinguishing peculiarity of logical symbols and expressions [consists] in the fact that each sentence constructed solely from

he continues, saying that "with its help one can show that undecidable sentences become decidable in systems which ascend farther in the sequence of types." Carnap worked out the details for himself and hit upon a definition that came close to that which Tarski was soon to give.

[16]There is also a second difference: Instead of the the basic T-sentences for atomic sentences (something he had not yet learned from Tarski) he has rules of evaluation for deciding the atomic sentences.

[17]The relevant clause is DA I.C.b. on p. 111.

[18]There are no conditions placed on this notion—for example, it could be a definite notion (such as 'provable in T').

them is determinate" (177).[19] He then gives a formal definition that aims to capture this idea. His actual definition is problematic for various technical reasons[20] and so we shall leave it aside. What is important for our purposes (as shall become apparent in §3.1) is the fact that (however the guiding idea is implemented) the actual division between 'logico-mathematical' and 'descriptive' expressions that one obtains as output is sensitive to the scope of the direct consequence relation with which one starts.

With this basic division in place, Carnap can now draw various derivative divisions, most notably, the division between analytic and synthetic statements: Suppose φ is a consequence of Γ. Then φ is said to be an *L-consequence* of Γ if either (i) φ and the sentences in Γ are logico-mathematical, or (ii) letting φ' and Γ' be the result of unpacking all descriptive symbols, then for every result φ'' and Γ'' of replacing every (primitive) descriptive symbol by an expression of the same genus (a notion that is defined on p. 170), maintaining equal expressions for equal symbols, we have that φ'' is a consequence of Γ''. Otherwise φ is a *P-consequence* of Γ. This division of the notion of consequence into *L-consequence* and *P-consequence* induces a division of the notion of demonstrable into *L-demonstrable* and *P-demonstrable* and the notion of valid into *L-valid* and *P-valid* and likewise for all of the other d-notions and c-notions. The terms '*analytic*', '*contradictory*' and '*synthetic*' are used for 'L-valid', 'L-contravalid' and 'L-indeterminate'.

The *content* of a sentence is defined to be the set of its non-analytic consequences. It follows immediately from the definitions (just as in the case of the first approach) that logico-mathematical sentences are analytic or contradictory and that analytic sentences are without content. This is what Carnap says in defense of the first of his three basic theses.

3 Critique

We are now in a position to determine (in Section 3.1) whether this account of mathematical truth is tenable and (in Section 3.2) whether Carnap has made the case for radical pluralism in pure mathematics.

[19] A sentence is *determinate* if either it or its negation is valid, that is, a consequence of the null set.

[20] For some of these problems see Quine (1963), though note that Quine's discussion appears at points to mistakenly assume that the notion of direct consequence that Carnap uses is a d-notion.

3.1 Nature of Mathematical Truth

There are three problems with Carnap's account of the nature of mathematical truth: the problem of free parameters, the problem with the conception of formality, and the problem of assessment sensitivity. These problems will be discussed in the next three subsections. Collectively they undermine Carnap's claims concerning the nature of mathematical truth. More precisely: §3.1.1 shows that Carnap has merely provided us with a piece of technical machinery with various parameters that can be set so as to yield a number of outcomes; he has set the parameters to yield his desired outcome and this amounts to a triviality; but he has provided no argument for setting the parameters one way rather than another. §3.1.2 shows that the claim that mathematical truth is purely formal rests on a confusion. §3.1.3 shows that the claim that in mathematics "no question of justification arises at all, but only the question of the syntactical consequences to which one or other of the choices leads" is problematic (xv).

3.1.1 Problem of Free Parameters

Both the first and second approach to the fundamental notions suffer from sensitivity to free parameters. The trouble with the first approach is that the definitions of analyticity that Carnap gives for Languages I and II are highly sensitive to the original classification of terms into 'logico-mathematical' and 'descriptive'. And the trouble with the second approach is that the division between 'logico-mathematical' and 'descriptive' expressions (and hence division between 'analytic' and 'synthetic' truths) is sensitive to the scope of the direct consequence relation with which one starts. This threatens to undermine Carnap's thesis that logico-mathematical truths are analytic and hence without content. Let us discuss this in more detail.

In the first approach, the original division of terms into 'logico-mathematical' and 'descriptive' is made by stipulation and if one alters this division one thereby alters the derivative division between analytic and synthetic sentences. For example, consider the case of Language II. If one calls only the primitive terms of first-order logic 'logico-mathematical' and then extends the language by adding the machinery of arithmetic and set theory, then, upon running the definition of 'analytic', one will have the result that true statements of first-order logic are without content while (the distinctive) statements of arithmetic and set theory have content.[21] For another example, if one takes the language of arithmetic, calls the primitive terms 'logico-mathematical' and then extends the language by adding the machinery of finite type theory, calling the basic terms 'descriptive', then, upon running the definition of 'analytic', the result will be that statements of first-order arithmetic are analytic or contradictory while (the distinctive) statements of second- and higher-order arithmetic are synthetic and hence have content. In general, by altering the input, one alters the output, and Carnap adjusts the input to achieve his desired output.

In the second approach, there are no constraints on the scope of the direct consequence relation with which one starts, and if one alters it one thereby alters the derivative division between 'logico-mathematical' and 'descriptive' expressions. Recall that the guiding idea is that logical symbols and expressions have the feature that sentences composed solely of them are determinate. The trouble is that (however one implements this idea) the resulting division of terms into 'logico-mathematical' and 'descriptive' will be highly sensitive to the scope of the direct consequence relation with which one starts.[22] For example, let S be first-order PA and for the direct consequence relation take "provable in PA" . Under this assignment, Fermat's Last Theorem will be deemed descriptive, synthetic and to have non-trivial content.[23] For an example at the other extreme, let S be an extension of PA that contains a physical theory and let the notion of direct consequence be given by a Tarskian truth definition for the language. Since in the metalanguage one can prove that every sentence is true or false, every sentence will be either analytic (and so have null content) or contradictory (and so have total content). To overcome such counter-examples and get the classification that Carnap desires, one must ensure that the consequence relation is (i) complete for the sublanguage consisting of expressions that one wants to come out as 'logico-mathematical' and (ii) not complete for the sublanguage

[21] This is more in keeping with the standard use of the term 'content'. For, in a straightforward sense, the truths of first-order logic do not pertain to a special subject matter (they are perfectly general) while those of arithmetic and set theory do.

[22] Carnap was fully aware of this sensitivity. See, for instance, the example involving $g_{\mu\nu}$ that Carnap gives (on p. 178) right after he draws the division.

[23] Carnap is fully aware of such counter-examples. See p. 231 of §62 where he notes that his definitions have the consequence that the universal numerical quantifier in Whitehead and Russell's *Principia Mathematica* is really a *descriptive* symbol, the reason being that the system involves only d-rules and hence (by Gödel's incompleteness theorem) it will leave some Π^0_1-sentences undecided.

consisting of expressions that one wants to come out as ‘descriptive’. Once again, by altering the input, one alters the output, and Carnap adjusts the input to achieve his desired output.

To summarize: What we have (in either approach) is not a principled distinction. Instead, Carnap has merely provided us with a flexible piece of technical machinery involving free parameters that can be adjusted to yield a variety of outcomes concerning the classifications of analytic/synthetic, contentful/non-contentful and logico-mathematical/descriptive. In his own case he has adjusted the parameters in such a way that the output is a formal articulation of his logicist view of mathematics that the truths of mathematics are analytic and without content. And one can adjust them differently to articulate a number of other views, for example, the view that the truths of first-order logic are without content while the truths of arithmetic and set theory have content. The possibilities are endless. The point, however, is that we have been given no reason for fixing the parameters one way rather than another. The distinctions are thus not principled distinctions. It is trivial to prove that mathematics is trivial if one trivializes the claim.[24]

[24]There are other points in the framework where the output rests on an artifact of the setup and if one alters this artifact one alters the output. The first such feature is the asymmetrical treatment of ‘logico-mathematical’ and ‘descriptive’ terms in the definition of ‘analyticity’. Carnap has not substantiated this feature of the setup. Moreover, if one alters this feature, say by giving a uniform treatment, treating descriptive terms on a par with logico-mathematical terms then the result just is a (notational variant of a) Tarskian truth definition (provided one replaces Carnap’s reduction method for atomic sentence by the T-conditionals)—all expressions are then analytic or contradictory and all analytic (true) expressions have null content. The second such feature is the definition of analyticity itself. Carnap has not argued that his particular technical definition captures the informal concept and indeed when he moves to a different setting he actually *changes* his technical definition to yield the desired output. This occurs in §25 when Carnap discusses physical geometry. Here he employs an entirely different technical definition of analyticity by taking the analytic sentences to be such that “[f]or proofs of these the *definitions* belonging to the axiomatic system may be used, but not the *axioms* themselves.” (79) Thus, the analytic statements will be those that are logical consequences (in our modern, first-order sense) of the definitions. This has the outcome that the statements “every point is a point” and “if each of three straight lines intersects the other two at different points, then the segments between the points of intersection form a triangle” are analytic, while “the sum of the angles of a triangle is equal to 2R” is synthetic. This is in accord with Carnap’s view (defended in his dissertation) that one brings to the world only a minimal geometric framework that is constitutive of the basic notions (and hence analytic) and then (with the help of coordinative definitions) one consults the world to narrow down the geometry (for example, to determine whether it is Euclidean, has constant curvature, etc.).

3.1.2 Problem with Conception of Formality

Let us now turn to the claim that Carnap has provided a purely *formal* criterion of mathematical truth, one that rests on syntactic form alone. In §2.1.3 we had occasion to note that there are places where Carnap does not carefully distinguish between syntax in the narrow sense (the combinatorial study of signs) and syntax in the wide sense (the study of c-notions). A similar confusion surrounds his discussion of the notion of formality.

After introducing the notion of logical syntax and telling us that it is the "*formal* theory of linguistic forms", (my emphasis) he elaborates on the notion of formality as follows:

> A theory, a rule, a definition, or the like is to be called *formal* when no reference is made in it either to the meaning of the symbols (for example, the 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. (1)

This is the modern sense of 'formal' as it occurs in metamathematical studies, most notably in proof theory and formalist conception of mathematics.

Now it is clear that syntax in the narrow sense is purely formal in this sense. But in giving a criterion of mathematical truth Carnap has outstripped syntax in the narrow sense and invoked syntax in the wide sense through the c-notions. So to secure the claim that mathematical truth is purely formal, he would have to establish that this holds of syntax in the wide sense. Evidently he thinks it does:

> We shall see that the logical characteristics of sentences (for instance, whether a sentence is analytic, synthetic, or contradictory...) ... are solely dependent upon the syntactical structure of the sentences. (1–2)

This, however, appears to rest on a confusion (regardless of whether one considers the first approach or the second approach to the c-notions).

The first approach involves the strong logic ω-logic. It is true that in this manner Carnap gives a complete criterion for arithmetical truth. This is simply because he has incorporated a Σ^0_1-complete theory T and it is easy to see that for any arithmetical sentence φ, φ holds iff $T \vdash_\omega \varphi$, where here we have used '$\vdash_\omega$' to indicate provability in ω-logic. So the issue is whether the answer to questions of the form "Does T $\vdash_\omega \varphi$?" is one that is "solely dependent on the syntactical structure of the sentences". Now there are two ways of viewing ω-logic: "proof-theoretically" or model-theoretically. A "proof" in ω-logic is an infinitary structure. The reason that Σ^0_1-complete theories T are complete for the statements of arithmetic in ω-logic is simply because these infinitary proofs can trace out the complete diagram of the standard model of arithmetic. On the model-theoretic side, the completeness theorem for ω-logic shows that $T \vdash_\omega \varphi$ if and only if φ is true in all ω-models of T, where an ω-model is simply a model that is correct in its computation of the natural numbers. In either case, whether one views the matter "proof-theoretically" or model-theoretically, the entanglement of the logic with the standard model of arithmetic is clear. Given this, it is hard to maintain that results in ω-logic depend on syntactic form alone.

Even if we pass over this objection and grant that ω-logic depends only upon the syntactic form, it will be very difficult to maintain this for much stronger logics such as β-logic or second-order logic, which Carnap will have to invoke when he passes beyond first-order arithmetic.

The problem with the second approach is even clearer. For on this approach, the definition of analyticity is simply a modified version of the definition of truth. In that definition, although one merely refers to syntactic items on the left-hand-side, on the right-hand-side one deals with mathematical objects. Indeed, the definition is essentially a translation of the object language into the mathematical metalanguage.

In summary, although strong logical relations such as those of ω-logic and β-logic involve relata that are syntactic items, and although such notions as "true in $V_{\omega+\omega}$" pertain to syntactic items, these notions themselves have substantial mathematical content (as can be revealed by looking at their definitions).[25] It is thus misleading at best to say that whether a sentence φ is analytic depends only on the "syntactic structure" or "syntactical design" (258) of the sentence.

3.1.3 Problem of Assessment Sensitivity

Carnap is perfectly aware that to define c-notions like analyticity one must ascend to a stronger metalanguage. However, there is a distinction that he often overlooks,[26] namely, the distinction between (i) having a stronger system S' that can *define* "analytic in S" and (ii) having a stronger system

[25]Cf. Section VII of Quine (1963).

[26]See e.g. pp. 1–2.

S' that can, in addition, *evaluate* a given statement of the form "φ is analytic in S".[27] It is an elementary fact that two systems S_1 and S_2 can employ the same definition (from an intensional point of view) of "analytic in S" (using either the definition given for Language I or Language II) but *differ* on their evaluation of "φ is analytic in S" (that is, differ on the extension of "analytic in S"). Thus, to determine whether "φ is analytic in S", one needs to access much more than the "syntactic design" of φ—in addition to ascending to an essentially richer metalanguage one must move to a sufficiently strong system to evaluate "φ is analytic in S". The first step need not be a big one.[28] But for certain φ the second step must be huge.[29]

In fact, it is easy to see that to answer "Is φ analytic in Language I?" is just to answer φ and, in the more general setting, to answer all questions of the form "Is φ analytic in S?" (for various mathematical φ and S), where here "analytic" is defined as Carnap defines it for Language II) *just is to answer all questions of mathematics.*[30] The same, of course, applies to the c-notion of consequence. So, when in first stating the Principle of Tolerance, Carnap tells us that we can choose our system S arbitrarily and that "no question of justification arises at all, but *only* the question of the syntactical consequences to which one or other of the choices leads" (p. xv, my emphasis)—where here, as elsewhere, he means the c-notion of consequence—*he has given us no assurance, no reduction at all.*[31]

[27]There are times when Carnap appears to be sensitive to this distinction. See, for example, the quotations at the end of §1.2 above.

[28]For example, taking S to be PA one can simply extend the language by adding a truth predicate and extend the axioms by adding the Tarskian truth axioms and allow the truth predicate to figure in the induction scheme. The resulting system S' is only minimally stronger than S. It proves Con(PA) but not much more.

[29]To continue the example in the previous footnote, suppose one wishes to show that "Con(ZF + AD) is analytic in S" (which, as I shall argue below, it is). To do this one must move to a system that has consistency strength beyond that of "ZFC + there are ω-many Woodin cardinals".

[30]For the first, note that $T \vdash_\omega \varphi$ if and only if φ (where T is the fixed Σ^0_1-complete theory) and for the second, note that the Tarskian truth definition has the feature that $T(\ulcorner\varphi\urcorner)$ if and only if φ (where T is the truth predicate).

[31]In his early papers on *Logical Syntax*, Michael Friedman maintained that Carnap was attempting to establish a neutral setting in which all disputants could discuss their disagreements. This neutral setting was embodied in Language I, a version of Primitive Recursive Arithmetic, wherein one studies syntax in the narrow sense. It is then easy to see that Gödel's incompleteness theorems undermine the attempt. (Friedman subsequently changed his view of Carnap's aims.) What the above discussion of assessment sensitivity

To summarize: In §3.1.1 we showed that Carnap has merely provided us with a piece of technical machinery with various parameters that can be set so as to yield a number of outcomes; he has set the parameters to yield his desired outcome and this amounts to a triviality; but he has provided no argument for setting the parameters one way rather than another. In §3.1.2, we showed that the claim that mathematical truth is purely formal rests on a confusion. In §3.1.3 we showed that the claim that in mathematics "no question of justification arises at all, but only the question of the syntactical consequences to which one or other of the choices leads" is problematic (xv).

3.2 Radical Pluralism in Mathematics

We turn now to the second thesis—that radical pluralism holds in mathematics. We noted that Carnap's pluralism is quite radical. Carnap maintains that "*any* postulates and *any* rules of inference can be chosen arbitrarily" (xv, my emphasis) and that "no question of justification arises at all, but only the question of the syntactical consequences to which one or other of the choices leads, including the question of non-contradiction" (xv). The view thus appears to be that any consistent system is equally legitimate and that the questions of truth and justification do not arise when selecting between such systems; truth and justification only pertain to questions of the form "Does system S imply φ?"; the remaining questions are questions of "mere expedience".

3.2.1 Initial Difficulty

Now there is a immediate problem with such a view, a problem brought out by Gödel's incompleteness theorems. To begin with, through the arithmetization of syntax, the metamathematical notions that Carnap takes to fall within the provenance of critical investigation were themselves seen to be a part of arithmetic. Thus, one cannot, on pain of inconsistency, say that there is a question of truth and falsehood with regard to the former but not the latter. More importantly, the incompleteness theorems buttressed the view that truth outstrips consistency. This is most clearly seen using

shows is that not only is it impossible to encompass syntax in the wide sense within syntax in the narrow sense but, even more dramatically, to answer all "syntactic" questions of the form "Is φ analytic in S?" (for various mathematical φ and S) just is to answer all of the questions of mathematics.

Rosser's strengthening of the first incompleteness theorem: Let T be an axiom system of arithmetic that (a) falls within the provenance of "critical investigation" and (b) is sufficiently strong to prove the incompleteness theorem.[32] Then, assuming that T is consistent (something which falls within the provenance of "critical investigation"), by Rosser's strengthening of the first incompleteness theorem, there is a Π^0_1-sentence φ such that (provably within $T + \mathrm{Con}(T)$) both $T + \varphi$ and $T + \neg\varphi$ are consistent. However, not both systems are equally legitimate. For it is easily seen that if a Π^0_1-sentence φ is independent from such a theory then it must be true.[33] So, although $T + \neg\varphi$ is consistent it proves a false arithmetical statement.[34] In short, one cannot, on pain of inconsistency, think that statements about consistency are not "mere matters of expedience" without thinking that Π^0_1-statements generally are not mere "matters of expedience".

The default view is that the question of whether a given Π^0_1-sentence holds is *not* a mere matter of expedience; rather, such questions fall within the provenance of theoretical reason.[35] What does Carnap have to say that

[32] A natural choice for such an axiom system is Primitive Recursive Arithmetic (PRA) but much weaker systems suffice, for example, $\mathrm{I}\Delta_0 + \exp$. Either of these systems can be taken as T in the argument that follows.

[33] The point being that T is Σ^0_1-complete (provably so in T).

[34] For the reader concerned that this argument involves the notion of truth in a problematic way, notice (as we have indicated in the parenthetical remarks) that it can be implemented in $T + \mathrm{Con}(T)$ (which is taken to fall within the provenance of "critical investigation"); that is, $T + \mathrm{Con}(T)$ proves that $T + \varphi$ and $T + \neg\varphi$ are consistent and it *also* proves φ.

[35] Some readers might be tempted to interpret me as saying that there is a "fact of the matter" concerning Π^0_1-sentences. I want to resist such a formulation since I am not sure that I understand the phrase "fact of the matter" as it is often employed. I have some sense of this phrase when it is used merely as a point of contrast with "matter of mere expedience". On this reading it means no more than that the issue is one of theoretical reason, one concerning something more than mere utility, one having something to do with the truth of one theory over another (not in some robust metaphysical sense of the 'truth' but in the ordinary sense). This distinction is not as sharp as one would like but one can point to clear cases (as we have seen above and as we shall see below) and the distinction strikes me as significant. In contrast, the phrase "fact of the matter" is often used in a way that strives for something more—"thick truth", "Truth with a capital 'T' ", the idea of "carving reality at the joints", etc. I cannot think of examples where I could go along with such talk with any confidence. Moreover, it seems to me that such talk buys into the myth that there is some Archimedean vantage point from which we can survey the array of theories and compare them with "reality as it is in and of itself"—in short, a "sideways-on view" (in McDowell's apt phrase). This is something that I think Kant

will sway us from the default view, and lead us to embrace his radical form of pluralism?

3.2.2 Two Interpretations

In approaching this question it is important to bear in mind that there are two general interpretations of Carnap. According to the first interpretation—the *substantive*—Carnap is really trying to *argue* for the pluralist conception. According to the second interpretation—the *non-substantive*—he is merely trying to *persuade* us of it, that is, to show that of the options it is most "expedient".[36]

There is evidence for each interpretation. There are certainly times when Carnap speaks as though he is taking a substantive approach. For example, he says, rather clearly, in the foreword, that the upshot of his work is that "the conflict between the divergent points of view on the problem of foundations of mathematics disappears" (xv). And, in his *Intellectual Autobiography*, he says that he wanted to "*show* that everyone is free to choose the rules of his language and thereby his logic in any way he chooses" (Carnap (1963), p. 17, my emphasis).[37] But there is also evidence for the non-substantive interpretation. One piece of evidence is the fact (alluded to earlier) that Carnap seems happy with the free parameters in his definition of analyticity and it seems that his real aim is to formally articulate his informal philosophical views and not defend a substantive claim. Another piece of evidence is that Carnap says repeatedly that the d-method is the fundamental method.[38] This view manifests itself in §§43–45 where Carnap discusses the admissibility of various indefinite notions (such as Π^0_1-properties, impredicative definitions, and indefinite notions in syntax). In each case he sidelines the foundational question by saying that it really amounts to the question of whether we should move to a system in which, say, one can prove the

and Carnap were right to reject and, once we follow them in doing so, we are left with the former, thinner distinction—the one I employ in the text. (See Tait (1986) for a critical discussion of the "myth of the model in the sky". The reader might press me with the concern that the thinner distinction that I invoke (following Carnap) rests on a similar myth. I do not think that it does. But it would take us too far afield to explore the issue here.)

[36]See Friedman (1999c) for the substantive interpretation and see Goldfarb & Ricketts (1992) for the non-substantive interpretation.

[37]Though perhaps 'show' in intended in the Wittgensteinian sense.

[38]See, for example, pp. 39, 100, 175, and 182.

relevant Π^0_1-statement (or existence of an impredicative set, or satisfaction of indefinite property). Since "φ is analytic in S" is likewise indefinite, he seems to be happy with a similar move to refer the question to a particular system and ask for the verdict there. But the whole point is that the answer will be dependent on which system one chooses. So, to be happy with these manoeuvres is to shun the foundational questions and advocate a non-substantive approach.[39]

In what follows I shall not try to settle this interpretive issue (though I am inclined to think that the evidence is in favour of a non-substantive interpretation). Instead I will criticize both versions of Carnap.

3.2.3 Substantive Interpretation

The most obvious approach to securing pluralism is to appeal to the work on analyticity and content. For if mathematical truths are without content and, moreover, this claim can be maintained with respect to an arbitrary mathematical system, then one could argue that even apparently incompatible systems have null-content and hence are really compatible (since there is no contentual conflict).

Now, in order for this to secure radical pluralism, Carnap would have to first secure his claim that mathematical truths are without content. But, as we have argued above, he has not done so. Instead he has merely provided us with a piece of technical machinery that can be used to articulate any one of a number of views concerning mathematical content and he has adjusted the parameters so as to articulate his particular view. So he has not *secured* the thesis of radical pluralism. Thus, on the substantive interpretation, Carnap has failed to achieve his end.

3.2.4 Non-Substantive Interpretation

This leaves us with the non-substantive interpretation. There are a number of problems that arise for this version of Carnap. To begin with, Carnap's technical machinery is not even suitable for *articulating* his thesis of radical pluralism since (using either the definition of analyticity for Language I or Language II) there is no metalanguage in which one can say that two apparently incompatible systems S_1 and S_2 both have null content and hence are

[39] For further evidence that Carnap is not advancing a substantive approach see the eyewitness account in Stein (1992), pp. 278–9.

really contentually compatible. To fix ideas, consider a paradigm case of an apparent conflict that we should like to dissolve by saying that there is no contentual conflict: Let $S_1 = \mathrm{PA} + \varphi$ and $S_2 = \mathrm{PA} + \neg\varphi$, where φ is any arithmetical sentence, and let the metatheory be MA = ZFC. The trouble is that on the approach to Language I, although in MT we can prove that each system is ω-complete (which is a start since we wish to say that each system has null content) we can also prove that one has null content, while the other has *total* content (that is, in ω-logic, *every* sentence of arithmetic is a consequence). So, we cannot within MT articulate the idea that there is no contentual conflict.[40] The approach to Language II involves a complementary problem. To see this, note that while a strong logic like ω-logic is something that one can apply to a *formal system*, a truth definition is something that applies to a *language* (in our modern sense). Thus, on this approach, in MT the definition of analyticity given for S_1 and S_2 is the same (since the two systems are couched in the same language). So, although in MT we can say that S_1 and S_2 do not have a contentual conflict this is only because we have given a deviant definition of analyticity, one that is blind to the fact that in a very straightforward sense, φ is analytic in S_1 while $\neg\varphi$ is analytic in S_2.

Now, although Carnap's machinery is not adequate to articulate the thesis of radical pluralism (for a given collection of systems) in a given metatheory, under certain circumstances he can attempt to articulate the thesis by changing the metatheory. For example, let $S_1 = \mathrm{PA} + \mathrm{Con}(\mathrm{ZF} + \mathrm{AD})$ and $S_2 = \mathrm{PA} + \neg\mathrm{Con}(\mathrm{ZF} + \mathrm{AD})$ and suppose we wish to articulate both the idea that the two systems have null content and the idea that '$\mathrm{Con}(\mathrm{ZF} + \mathrm{AD})$ is analytic in S_1' while '$\neg\mathrm{Con}(\mathrm{ZF} + \mathrm{AD})$ is analytic in S_2'. As we have seen, no single metatheory (on either of Carnap's approaches) can do this. But it turns out that because of the kind of assessment sensitivity that we discussed above, there are two metatheories, MT_1 and MT_2, such that in MT_1 we can say both that S_1 has null content and that '$\mathrm{Con}(\mathrm{ZF} + \mathrm{AD})$ is analytic in S_1', while in MT_2 we can say both that S_2 has null content and that '$\neg\mathrm{Con}(\mathrm{ZF} + \mathrm{AD})$ is analytic in S_2'. But, of course, this is simply because (any such metatheory) MT_1 proves $\mathrm{Con}(\mathrm{ZF} + \mathrm{AD})$ and (any such metatheory) MT_2 proves $\neg\mathrm{Con}(\mathrm{ZF} + \mathrm{AD})$. So we have done no more than (as we must) reflect the difference between the systems in the metatheories. Thus, although Carnap does not have a way of articulating his radical pluralism (in

[40]This is related to a point made by Michael Friedman on p. 226 of Friedman (1999c).

a given metalanguage), he certainly has a way of *manifesting* it (by making corresponding changes in his metatheories).

As a final retreat, Carnap might say that he is not trying to persuade us of a *thesis* that (concerning a collection of systems) can be articulated in a given system but rather is trying to persuade us to adopt a thorough radical pluralism as a "*way of life*". He has certainly shown us how we can make the requisite adjustments in our metatheory so as to consistently manifest radical pluralism. But does this amount to more than an algorithm for begging the question?[41] Has Carnap shown us that there is no question to beg? I do not think that he has said anything persuasive in favour of embracing a thorough radical pluralism as the "most expedient" of the options. The trouble with Carnap's entire approach (as I see it) is that the question of pluralism has been detached from actual developments in mathematics. To be swayed from the default position, something of greater substance is required.[42] I will return to this in the final section. But before doing so let us examine his position in later work to see if it changes in any substantial way.

4 "Foundations of Logic and Mathematics"

Shortly after writing *Logical Syntax*, Carnap learned of Tarski's definition of truth and he modified his system by incorporating it.[43] The central texts from this period are "Foundations of Logic and Mathematics" (1939), *In-*

[41]There are many places where Carnap quite obviously begs the question in the metatheory. See, for example, §§43 and 44 where Carnap discusses intuitionism and predicativism; to people like Brouwer and Poincaré these sections would be maddening.

[42]In Koellner (2009b) I map out a scenario in which one could plausibly be swayed from the default position at a given level of the hierarchy of mathematical systems, for example, with regard to CH.

[43]Recall that the definition of "analyticity" that Carnap gives in *Logical Syntax* is really a variant of the Tarskian truth definition. More precisely, the version Carnap gives deviates from the standard Tarskian truth definition in two key respects: First, instead of the the basic T-sentences for atomic sentences (something he had not yet learned from Tarski) he has rules of evaluation for deciding the atomic sentences. Second, there is an asymmetry in the definition in that logico-mathematical sentences are treated differently than descriptive sentences; to count as analytic, a descriptive sentence must meet a higher demand—it must be such that all of its substitution instances are analytic. For the logic-mathematical fragment (where the second change is no longer relevant), the definition he gives is *logically equivalent* to the standard Tarskian definition; though, of course, the two definitions agree in extension, they do not agree in intension.

troduction to Semantics (1942), and *Formalization of Logic* (1943). In this period, Carnap is careful to divide semiotics (the general theory of signs) into pragmatics, semantics, and syntactics. The semantic notions now take center stage and for this reason this period is often referred to as Carnap's "semantic period".

The paper "Foundations of Logic and Mathematics" contains the most detailed discussion of Carnap's views on the foundations of logic and mathematics in this new, semantic setting. During this period Carnap continued to uphold his three central theses.

Let us begin by documenting this: Concerning the thesis that logical and mathematical theorems are without factual content he writes:

> It will become clear that [logical and mathematical theorems] do not possess any factual content. (2)

> If S_1 is an L-true [analytic] sentence, then the truth of S_1 can be established without regard to the facts, e.g., to the properties of those things whose names occur in S_1. (15)

Moreover, he continues to hold that mathematical truth is purely formal:

> A definition of a term in the metalanguage is called *formal* if it refers only to the expressions of the object-language (or more exactly, to the kinds of signs and the order in which they occur in the expressions) but not to any extralinguistic objects and especially not to the designata of the descriptive signs of the object-language. . . . Now our question is whether it is possible to define within syntax, i.e., in a formal way, terms which correspond more or less to these semantical terms ['true', 'L-true', 'L-implicate']. The development of syntax—chiefly in modern symbolic logic—has led to an affirmative answer to that question. (16–17)

Concerning the thesis that radical pluralism holds in pure mathematics, he writes:

> Now, if we regard interpreted mathematics as an instrument of deduction within the field of empirical knowledge rather than as a system of information, then many of the controversial problems are recognized as being questions not of truth but of technical expedience. (50)

We shall continue to focus on these two theses, leaving the third thesis for the final section.

In this section we shall examine his modified views and his defense of them in some detail in order to see whether his views change in any substantial way and whether our above criticisms continue to hold in the new setting. We shall see that much of the framework remains intact and our earlier criticisms carry over to the new setting. However, in the end there is an interesting twist. While tracing out the main components of his earlier view—now with the new Tarskian definition in place of his original definition of "analyticity"—and trying to dissolve the foundational disputes, his course is modified almost as if he were aware of the above criticisms and in the end he is reduced to making a mere *proposal*, namely, that we regard mathematics as a mere instrument for the empirical sciences.

4.1 The New Setting

Let us begin by discussing the terminology of the new setting and comparing it with the old.

4.1.1 Pragmatics, Semantics, Syntax

Carnap begins his discussion at a very general level by discussing languages in use (such as German). A *language in use* is a "system of activities, or rather habits, i.e. dispositions to certain activities, serving mainly for the purposes of communication and of co-ordination of activities among members of a group" (3). Since his interest is in *theoretical languages* (such as the language of physics) he restricts attention to languages that contain only declarative sentences. There are tree sub-disciplines in the study of theoretical language: The subject of *pragmatics* is concerned with "the action, state, and environment of a man who speaks or hears, say, the German word 'blau'" (4). The subject of *semantics* is concerned with "the expressions of the language and their relation to their designata" (4). The subject of *syntactics* (or *logical syntax*) involves abstracting even further and considering not the designata of expressions but rather only their formal properties and relationships.

4.1.2 Semantical Systems

For a given language in use one can construct a *semantical system* for the language, that is, a system describing the relationship between the signs of the language and their designata. In anticipation of Quine's indeterminacy thesis, Carnap notes that "[t]hese rules are not unambiguously determined by the facts" (6). Therefore, the question of whether a semantical system is faithful to a given language is one that admits of a certain degree of indeterminacy. Although this presents a challenge for the field linguist, it is not of relevance for the subsequent discussion since Carnap leaves languages in use by the wayside and concentrates on precisely articulated semantical systems.

A semantical system is constructed as follows: First the signs are divided into *descriptive* and *logical* signs:

> As descriptive signs we take those which designate things or properties of things. ... The other signs are taken as logical signs: they serve chiefly for connecting descriptive signs in the construction of sentences but do not themselves designate things, properties of things, etc. Logical signs are, e.g., those corresponding to English words like 'is', 'are', 'not', 'and', 'or', 'if', 'any', 'some', 'every', 'all'. (7)[44]

Carnap illustrates these ideas by introducing a simple language and giving a semantical system for it. The semantical rules give the *designata of descriptive signs* and the *truth conditions* for the sentences. For the former we have, for example, "'titisee' designates a certain lake" and "'kalt' designates the property of being cold" (9). For the latter we have the basic clause for atomic sentences: "A sentence of the form '$\ldots_n$ ist - - -$_p$' is true if and only if the things designated by '$\ldots_n$' has the property designated by '- - -$_p$'" (9); and for the logical connectives we have the standard Tarskian clauses.[45] The

[44]Notice that from this representative example it is unclear how one is to classify the signs of areas of mathematics such as arithmetic. For example, on Carnap's view, does '7' designate an object that has various properties (such as *being prime*) or is it a logical sign? This issue will resurface and occupy center stage in what follows.

[45]Notice once again that in this representative example of a semantical system, Carnap ignores the problematic case of expressions that we would call mathematical but non-logical, such as the distinctive expressions of arithmetic and set theory. One could just as easily include descriptive clauses for such expressions; for example, "'7' designates 7" and "'prime' designates the property of primality". Once again, this issue will gain prominence in what follows.

semantical rules yield an implicit definition of *truth* in the language system and they thereby provide an *interpretation* of the expressions of the language system.

4.1.3 Syntactical Systems

A *syntactical system* or *calculus* concerns the formal properties of expressions in the object-language.

> A definition of a term in the metalanguage is called *formal* if it refers only to the expressions of the object-language (or, more exactly, to the kinds of signs and the order in which they occur in the expressions) but not to any extralinguistic objects and especially not to the designata of the descriptive signs of the object-language. (16)

The notion of a syntactical system coincides with the modern notion of an axiom system except that Carnap includes both systems with *finite rules* and systems with *transfinite rules*.[46] The statements that are provable in the calculus C are said to be C-true. Likewise for C-false, C-implicate, etc. A *language system* is a semantical system S conjoined with a syntactical system C.

4.1.4 Comparison with Earlier Terminology

The two aspects of the earlier notion of language (underlying system of expressions and formal system) are now clearly separated and the terms 'language' and 'syntactical system' are now used to track these separate components. In addition, the two aspects of the earlier notion of syntax (syntax in the narrow sense and syntax in the wide sense) are now clearly separated and the terms 'syntax' and 'semantics' now replace what was (roughly) the old distinction.

[46]The systems with finite rules coincide exactly with the modern notion of an axiom system. The systems with transfinite rules are much stronger. However, there appears to be a limitation to their strength. Carnap tells us that a rule is called *transfinite* "if it refers to an infinite number of premises" (23). Since he appears to envisage only countable languages it would appear that the strongest transfinite system he envisages is ω-logic.

4.2 Nature of Mathematical Truth

With these terminological preliminaries out of the way, we can now turn to Carnap's new account of mathematical truth. As before, the analytic-synthetic distinction is central, although here it is labeled the 'logical-factual' distinction.

4.2.1 Distinction between Factual and Logical

The informal distinction between *factual* and *logical* sentences is that for the former we have to "look to the world" to determine the truth-value while for the latter it suffices to "understand the language". As examples, Carnap gives 'Australia is large' (S_1) and 'Australia is large or Australia is not large' (S_2).

> Then, for S_1 we have to know, in addition [to understanding the language to which it belongs], some facts about the thing whose name occurs in it, i.e., Australia. Such is not the case for S_2. Whether Australia is large or small does not matter here; just by understanding S_2 we become aware that it must be right. If we agree to use the same term 'true' in both cases, we may express their difference by saying that S_1 is factually (or empirically) true while S_2 is logically true. (12)

Now this distinction is hardly precise. The illusion of clarity arises from restriction to simple and extreme cases. To see this, consider two more troublesome examples. First, the equivalence principle in general relativity; second, the Π^0_1-statement Con(ZF + AD) of arithmetic. Carnap would presumably want to say that the first is factually true (empirically true) while the latter is logically true. But it is far from the case that to establish the truth value of the former one merely looks to the empirical world to see whether certain objects have certain properties. And to establish the latter it is far from the case that one merely reflects on language. The trouble lies in the reference to the notion of understanding.

Carnap is aware of the lack of precision engendered by reference to the vague notion of "understanding the language". But he assures us that this references is eliminable in a way that secures clarity: "These unprecise explanations can easily be transformed into precise definitions by replacing the former reference to understanding by a reference to semantics" (12–13). This is done as follows:

> We call a sentence of a semantical system S (logically true or) *L-true* if it is true in such a way that the semantical rules of S suffice for establishing its truth. We call a sentence (logically false or) *L-false* if it is false in such a way that the semantical rules suffice for finding that it is false. The two terms just defined and all other terms defined on their basis we call *L-semantical terms*. If a sentence is either L-true of L-false, it is called *L-determinate*, otherwise (L-indeterminate or) *factual*. (The terms 'L-true', 'L-false', and 'factual' correspond to the terms 'analytic', 'contradictory', and 'synthetic', as they are used in traditional terminology, usually without exact definitions.) (13)

In this way, the notion of a truth-value being determined by understanding is replaced by the notion of a truth-value being determined by semantical rules.

4.2.2 Criterion of Mathematical Truth

The relationship between a semantical system and a syntactical system now takes center stage. Let S be a semantical system and let E be the underlying language of expressions. Thus S provides a truth definition for E. Let C be a calculus formulated in the language E. An interpretation S of C is said to be a *true interpretation* if it is sound (in the modern sense), that is, if whenever φ is C-true then φ is true in S.

It is straightforward to arrive at sound interpretations of various syntactic systems. Indeed, if the syntactic system is consistent then the completeness theorem guarantees that there is a sound interpretation. The key question is whether given a semantical system S one can arrive at a calculus that is *complete* with respect to S.

> [I]t is an important problem whether it is possible to construct for a given system S a calculus C such that C is not only in accordance with S, in the sense explained [that is, sound with respect to S], but that the extensions of 'C-implicate', 'C-true', and (if defined at all) 'C-false' coincide with those of 'L-implicate', 'L-true', and possible 'L-false', respectively. If this is the case, we call C an *L-exhaustive calculus* with respect to S. (22)

Carnap maintains that this can be achieved:

> Now our question is whether it is possible to define within syntax, i.e., in a formal way, terms which correspond more or less to those semantical terms [namely, 'true', 'L-true', and 'L-implicate'], i.e., whose extensions coincide partly or completely with theirs. The development of syntax—chiefly in modern symbolic logic—has led to an affirmative answer to that question. Especially is the possibility of defining in a formal way terms which completely correspond to 'L-true' and 'L-implicate' of fundamental importance. This shows that logical deduction can be completely formalized. (16–17)

In this way, in the semantic setting, one arrives at the formal criterion of mathematical truth.

4.2.3 Critique

We know by the incompleteness theorems that this cannot be achieved if S is sufficiently rich to interpret a minimal amount of arithmetic (e.g. $I\Delta_0 + \exp$ suffices) and by a syntactical system or calculus one means formal system (in the modern sense).

Carnap is quick to point out that Gödel has shown that "a calculus of the ordinary kind (in our terminology, a finite calculus) cannot be constructed for the whole of arithmetic" (23), that is, there can be no L-exhaustive finite calculus for the language of arithmetic under its standard interpretation. However, as he has observed in *Logical Syntax*, if one allows the transfinite ω-rule then one can construct a calculus (in this extended sense) that is complete for first-order arithmetic, that is, one can construct a rich calculus C such that C-true coincides with L-true, that is, true in arithmetic. Thus, in the above passage Carnap is using the notion of a syntactical system (or calculus) in a *wide* sense.

But our criticisms in Section 3.1.2 apply in the present setting. First, it is tendentious to say one can define in "a formal way" the notion of L-truth for arithmetic by invoking ω-logic. Second, it becomes even more tendentious when one moves to stronger logics such as β-logic and second-order logic. Third, in the end Carnap has to move from a strong logic to simply giving a Tarskian truth definition (as he did in the case of Language II)[47] and

[47] Carnap addresses the question of "whether it is possible to construct for a given semantical system S an L-exhaustive calculus C". He says that "[t]he answer depends

yet to say that one has defined truth in set theory in a purely formal way simply by giving the Tarskian definition is misleading at best since on the right-hand-side of the definition one refers to non-syntactic items; indeed the definition is essentially a translation of the object language into the set theoretic metalanguage.

The distinction between logical and factual truth does not fare any better in the present setting. The informal classification depended on the notion of a truth-value being determined by understanding. But this was vague. So Carnap revised the account by replacing this notion with the notion of a truth-value being determined by semantical rules. However, this is hardly an improvement. The semantical rules are just a Tarskian truth definition which is, in effect, simply a translation of the object language into the metalanguage. It is quite unclear what it means to say that these rules determine the truth-value of a given statement. For example, do they determine the truth-value of Con(ZF + AD)? Is it analytic? This question doesn't even make sense unless one specifies some derivational system and there is great freedom in doing so. We have here the problem of assessment sensitivity discussed in Section 3.1.3.

The present account of the nature of mathematical truth is thus even less a success than the account in *Logical Syntax*. The account in *Logical Syntax* suffered from the problem of parameters, the problem with the conception of formality, and the problem of assessment sensitivity. But at least it was a precise account. The present account suffers from these problems but has the additional defect of resting on the problematic notion of "being determined by the semantical rules".

4.3 Pluralism: First Pass

We now turn to the second thesis—that radical pluralism holds in pure mathematics. The first pass that Carnap takes at the subject is a high-level

upon whether there are in S a sentence S_2 and an infinite class of sentences C_1 such that S_2 is an L-implicate of C_1 but not an L-implicate of any finite class of C_1". In other words, the answer depends upon whether there are statements that are ω-provable which are not first-order provable. He continues: "If this is not the case, then there is a finite L-exhaustive calculus C. If, however, it is the case, an L-exhaustive calculus C can be constructed if and only if transfinite rules are admitted" (23). The 'if' part of this last statement is mistaken if by transfinite rules Carnap means rules like the ω-rule. For example, such rules will never lead to an L-exhaustive calculus for full second-order arithmetic.

discussion of the controversy between pluralists (conventionalists) and non-conventionalists. He argues that there is a fundamental confusion in the debate and that once this confusion is cleared away—something that can be done by carefully distinguishing between syntax and semantics—one sees that the two parties are largely talking past each other and that each has glimpsed a facet of the truth. In this way he hopes to reconcile opposing parties. However, as we shall see, residual controversies remain.

4.3.1 *S*-First Approach versus *C*-First Approach

The fundamental confusion arises from not distinguishing between two very different approaches to setting up a language system (which recall is the conjunction of a semantical system and a syntactical system (calculus)). On the *first approach* (the *S-first approach*) one starts by first constructing a semantical system S (classifying signs into 'descriptive' and 'logical', laying down rules of designation and the clauses of the truth definition) and then constructing a calculus C under the constraint that C be sound with respect to S. On the *second approach* (the *C-first approach*) one starts by first constructing a calculus C and then constructing a semantical system under the constraint that C be sound with respect to S.

With this distinction in place, Carnap turns to the question of whether logic and mathematics are conventional.

> There has been much controversial discussion recently on the question whether or not logic is conventional. Are the rules on which logical deduction is based to be chosen at will and, hence, to be judged only with respect to convenience but not correctness? Or is there a distinction between objectively right and objectively wrong systems so that in constructing a system of rules we are free only in relatively minor respects (as, e.g., the way of formulation) but bound in all essential respects? (26–7)

Carnap is quite clear that for this question to be non-trivial it must pertain to *interpreted* languages:

> Obviously, the question discussed refers to the rules of an interpreted language, applicable for purposes of communication; nobody doubts that the rules of a pure calculus, without regard to any interpretation, can be chosen arbitrarily. (27)

His answer to this (non-trivial) version of the dispute between the conventionalist and non-conventionalist is that both sides can be seen to be correct once one attends to the distinction between the C-first approach and the S-first approach: On the C-first approach one has total freedom in erecting the calculus C.

> [T]here is no question of a calculus being right or wrong, true or false. A true [sound] interpretation is possible for any given consistency calculus (and hence for any calculus of the usual kind, not containing rules for 'C-false'), however the rules may be chosen. (27)

Here Carnap is evidently invoking the completeness theorem. So on this approach, the proponents of the conventional nature of logic and mathematics are correct. In contrast, on the S-first approach "the "meanings" of the logical signs are given before the rules of deduction are formulated" and hence we are "indeed bound in the choice of the rules in all essential respects" (28). In short, on the S-first approach

> those who deny the conventional character of logic, i.e., the possibility of a free choice of the logical rules of deduction, are equally right in what they mean if not in what they say. (27)

In this way, by distinguishing between the C-first and the S-first approach, "we come to a reconciliation of the opposing views" (28). In summary:

> The result of our discussion is the following: logic or the rules of deduction (in our terminology, the syntactical rules of transformation) can be chosen arbitrarily and hence are conventional if they are taken as the basis of the construction of the language system and if the interpretation of the system is later superimposed. On the other hand, a system of logic is not a matter of choice, but either right or wrong, if an interpretation of the logical signs is given in advance. (28)

4.3.2 Critique: Residual Controversy

Now, one can hardly dispute that on the C-first approach one has great freedom in the construction of a language system and hence, in this sense the conventionalists are correct. However, the participants in the foundational disputes would hardly be assured, for they would maintain, first, that all along they had been concerned with the S-first approach and, second, that even on this approach the controversies remain. Take, for example, the dispute concerning the axiom of choice (AC) and other non-constructive principles. The parties to the dispute use the same language of arithmetic and set theory, they can give the same Tarskian truth definition and yet this is of little aid in resolving the disputes about whether the principles in question hold. So, even on the S-first approach controversies remain.

Carnap seems to be aware that such residual controversies remain even on the S-first approach. For consider his discussion of higher-order calculi. In §14, Carnap gives an informal description of the language of higher-order logic (for all finite orders). He does not specify the corresponding calculus but he does note that some of the proposed axioms are controversial:

> Some new rules of transformation for these new kinds of variables have to be added. We shall not give them here. Some of them are still controversial. (33)

Of course, there is only controversy if one is taking the S-first approach. For further evidence that Carnap is aware that controversy can arise even on the S-first approach, consider his discussion in §16, which concerns controversies that might arise in the course of a derivation in the physical sciences. He notes that one virtue of the axiomatic method is that it enables disputants to place deductions under a "logical microscope" and thereby "expand the critical part of the controversial deduction to the degree required by the situation". When one does this, once again, one will typically locate controversies concerning the axioms of higher order logic (if one locates them at all):

> The critical point will usually not be within the elementary part of the logical calculus ... but to a more complex calculus, e.g., the higher, mathematical part of the logical calculus, or a specific mathematical calculus, or a physical calculus. This will be discussed later. (37)

Thus Carnap appears to be aware that even on the S-approach there are residual controversies.

There is a large amount of controversy indeed. For example, there is controversy over whether AC holds. Indeed the entire apparatus of full-second order logic is highly entangled with mathematical notions. One way to make this point is to note that the consequence relation for full second-order logic is intimately entangled with Π_2 over the entire universe of sets. For example, the problem of determining what follows from what in full second-order logic is (Turing-) equivalent to the problem of determining which Π_2 statements hold in the universe of sets V. In short, the controversy is as deep as the controversy over Π_2 truth in V. This controversy includes CH and *much* more. Higher-order logic is thus highly entangled with mathematical truth and the residual controversies are as broad as those that arise in set theory.[48]

4.4 Pluralism: Second Pass

We have seen that residual controversies arise even on the S-approach. Carnap appears to be aware of this and he tells us that "[t]his will be discussed later" (37). He returns to the subject in his discussion of the traditional controversy between logicism, formalism, and intuitionism. We will describe this in the present section. But let us first start with a brief account of Carnap's logicism. This is of interest in itself since in the present work his logicism takes on new significance.

4.4.1 Carnap's Logicism

In *Logical Syntax*, systems that would normally be classified as mathematical and thought to have special subject matter (such as Language I (PRA)) are classified from the start as 'logico-mathematical'. The task is to vindicate logicality by showing that the logico-mathematical truths are free of content. This is achieved by the technical account of such terms as 'analyticity' and 'content'. In the present work, the terms 'logical' and 'mathematical' are separated from the start and logical and mathematical systems are treated separately. Now logicism (in something close to the traditional sense) plays a key role in reducing the mathematical systems to the logical systems. The logicist reduction now plays a key role in vindicating the claim that mathematical truths are free of content.

In §§16–17 Carnap discusses nonlogical calculi, which he also calls 'axiom systems'.

[48]See Väänänen (2001) and Koellner (2009a) for more on the entanglement of strong logics with set theory.

> Each of the nonlogical calculi to be explained later consists, strictly speaking, of two parts: a logical *basic calculus* and a *specific calculus* added to it. (37)

On the standard interpretation, the logical calculus is fully general while the specific calculus pertains to a special subject matter. Carnap gives what he calls a *mathematical* example:

> We take here as mathematical calculi those whose customary interpretation is mathematical, i.e., in terms of numbers and functions of numbers. As an example, we shall give the classical axiom system of Peano for (elementary) arithmetic. It is usually called an axiom system of arithmetic because in its customary interpretation it is interpreted as a theory of natural numbers, as we shall see. (38–39)

He then gives the standard axioms of PA, formulated against the backdrop of second-order logic.

Now, as Carnap has said, the customary interpretation of this system is mathematical, that is, it involves mathematical objects such as numbers and functions. According to Carnap this interpretation is problematic:

> [I]n this interpretation the system concerns the progression of finite cardinal numbers, ordered according to magnitude. Against the given semantical rule "'b' designates the cardinal number 0" perhaps the objection will be raised that the cardinal number 0 is not an object to which we could point, as to my desk. This remark is right; but it does not follow that the rule is incorrect. (40)

The reason (according to Carnap) that it does not follow that the rule is incorrect is that the semantic rules can be vindicated by translating the calculus into another calculus that admits a *properly logical* interpretation. To demonstrate this Carnap proceeds to give a translation[49] of (second-order) PA into the finite theory of types (that is, the higher-order logic he discussed

[49]The notion of translation that Carnap gestures at is more sharply articulated in the modern notion of an *interpretation* in mathematical logic. Roughly speaking one translates the expressions of one system into those of another in such a way that theorems in one system are mapped into theorems in the other system. For further details see p. 40 and compare what Carnap says there with the contemporary account in Lindström (2003).

earlier). One can then use the translation to transfer the customary interpretation of the finite theory of types (which is deemed logical) to the syntactical system of PA, thereby vindicating the logicality of the latter. He concludes:

> If we assume that the normal interpretation of the logical calculus is true, the given interpretation for the Peano system is shown to be true by showing that the correlates of the axioms are C-true. And it can indeed be shown that the sentences [of PA] are provable in the higher functional calculus, provided suitable rules of transformation are established. As the normal interpretation of the logical calculus is logical and L-true, the given interpretation of the Peano system is also logical and L-true. (41)

Thus, the customary *mathematical* interpretation of PA can be replace by a *properly logical* interpretation.[50] In a similar fashion, one can translate analysis, the higher-orders of arithmetic, and all of the mathematics that is used in the physical sciences into the functional theory of types and thereby give these systems of mathematics a properly logical interpretation. In this manner logicism now plays a key role in showing that mathematics is logical and not factual.[51]

The claim that the interpretation that Carnap gives is *logical* is problematic for three familiar reasons: First, as we have already noted above, higher-order logic is entangled with Π_2-truth in the universe of sets to a degree that arguably undermines the claim that it is logical and not properly mathematical. Second, Carnap neglects to mention that for his translation to yield PA one must assume that the axiom of infinity holds for the first-order domain and this assumption is hardly logical. Finally, Carnap is mistaken when he writes: "In a customary interpretation of a mathematical calculus [that is, an interpretation into a logical calculus, as discussed above] every sign in it is interpreted as a logical sign, and hence every sentence consists only of logical signs and is therefore L-determinate (see §7)." (43) For, as noted above, this claim is not exact since it involves the notion of L-determinateness which in turn rests on the problematic notion of "being determined by the semantical rules" and, moreover, when one attempts to make the claim exact by referencing a particular formal system the claim is false. For these reasons

[50]Carnap refers to the latter as a "reformulation" of the former. It would be more appropriate to call it a "regimentation" or simply a "translation".

[51]This likely explains why in his representative examples (cf. footnotes 44 and 45 in §4.1.2 above) Carnap discusses only terms we would now classify as logical.

Carnap's logicism is problematic. He has not given an interpretation which can be said to be purely formal. He has not established L-determinateness and he has not established logicality.

4.4.2 Logicism, Formalism, Intuitionism

We turn now to Carnap's final attempt to dissolve the traditional foundational controversies such as the controversy between logicism, formalism, and intuitionism.

Here his how Carnap understands the three foundational frameworks: The chief thesis of *logicism* is that "mathematics is a branch of logic" and (in Carnap's view) "this thesis was demonstrated" by providing an interpretation of all of mathematics within the higher-order functional calculus. (48) In contrast *formalism* involves starting with specific axiom systems, such as the axiom systems for the natural numbers and the real numbers but "without regard to interpretation". (49) Finally, in contrast to both of these approaches, *intuitionism* "rejects both the purely formal construction of mathematics as a calculus and the interpretation of mathematics as consisting of L-true sentences without regard to factual content" but instead regards mathematics "as a field of mental activities based upon "pure intuition" "; on this view, "reference to a mere possibility is not allowed unless we know a method of actualizing it". (50)

Carnap attempts to dissolve the conflict between formalism and logicism as follows:

> The controversy between the two doctrines concerning the question whether first to construct logic and then mathematics within logic without new primitive signs, or both simultaneously, has at present lost much of its former appearance of importance. We see today that the logico-mathematical calculus can be constructed in either way and that it does not make much of a difference which one we choose. (49)

Concerning the conflict between logicism and intuitionism he writes:

> If we compare, e.g., the systems of classical mathematics and of intuitionistic mathematics, we find that the first is much simpler and technically more efficient, while the second is more safe from surprising occurrences, e.g., contradictions. (51)

This account of the foundational disputes is problematic. First, Carnap claims that logicism has been demonstrated, and this claim is problematic for the reasons given above. Second, he claims that the formalist seems to regard it as impossible to give an interpretation for the mathematical calculus. It is hard to see why he would think this. For surely the formalist could give an interpretation for any mathematical system if by 'interpretation' we mean 'interpretation in Carnap's sense'; for to do so one simply has to run the truth definition. S formalist would find little value in this because it would amount to no more than a translation of the object language into the metalanguage, and would provide no insight into the question of consistency, which is the main thrust of the formalist program. Finally, Carnap's statement comparing the advantages and disadvantages of classical mathematics and intuitionistic mathematics is also problematic; for example, intuitionistic arithmetic (Heyting arithmetic (HA)) has the same consistency strength as classical arithmetic (PA).

Let us set these difficulties aside and turn to Carnap's attempted resolution of the residual debates.

4.4.3 Carnap's Proposal: Instrumentalism

After his discussion of the conflicting foundational viewpoints, Carnap makes his final attempt to dissolve the residual controversies. He begins as follows:

> Concerning mathematics as a pure calculus there are no sharp controversies. These arise as soon as mathematics is dealt with as a system of "knowledge"; in our terminology, as an interpreted system.

This is beyond dispute—all parties would agree. He continues:

> Now, if we regard interpreted mathematics as an instrument of deduction within the field of empirical knowledge rather than as a system of information, then many of the controversial problems are recognized as being questions not of truth but of technical expedience.

This is the final attempt at dissolution. Carnap is suggesting that we treat mathematics as a mere instrument for deduction in the empirical sciences. Once we do we shall see that the residual controversies disappear. Once one adopts this perspective on mathematics, one sees that there are no substantive issues in the foundations of mathematics; rather, there are only issues of "mere expedience". The key question is: *Should* we regard interpreted mathematics as a mere instrument for deduction in the empirical sciences?

It is unclear whether Carnap is *defending a thesis* to the effect that instrumentalism is true and hence that there are no real controversial problems in mathematics, or whether here is merely *suggesting a perspective* ("if we regard…") to the effect that we regard mathematics as a mere instrument for the physical sciences. At times, he speaks as though he defending a thesis: For example, when he announces his instrumentalist conception early in the monograph, he writes:

> It is one of the chief tasks of this essay to make clear the role of logic and mathematics as applied in empirical science. We shall see that they furnish instruments for deduction, that is, for the transformation of formulations of factual, contingent knowledge. However, logic and mathematics not only supply rules for transformation of factual sentences but they themselves contain sentences of a different, non-factual kind. Therefore, we shall have to deal with questions of the nature of logical and mathematical theorems. It will become clear that they do not possess any factual content. If we call them true, then another kind of truth is meant, one not dependent on facts. A theorem of mathematics is not tested like a theorem of physics, by deriving more and more predictions with its help and then comparing them with the results of observations. But what else is the basis of their validity? We shall try to answer these questions by examining how theorems of logic and mathematics are used in the context of empirical science. (2)

Later he elaborates as follows:

> The application of mathematical calculi in empirical science is not essentially different from that of logical calculi. Since mathematical sentences are, in the customary interpretation [that is, an interpretation into a logical calculus, as discussed above], L-determinate, they cannot have factual content; they do not convey information about facts which would have to be taken into consideration besides those described in empirical science. The

> function of mathematics for empirical science consists in providing, first, forms of expression shorter and more efficient than non-mathematical linguistic forms and, second, modes of logical deduction shorter and more efficient than those of elementary logic. (44)

The language in these passages is quite strong: He says that we shall *see* that mathematics is merely an instrument for deduction in the empirical sciences and that upon the analysis of mathematical truth it will *become clear* that mathematical statements do not possess any factual content. His aim is to discern the basis of the validity of mathematical theorems by examining their use in the empirical sciences. His ultimate claim has two parts: First, the various mathematical systems are conservative with respect to the domain of empirical sentences. Second, this is because his analysis reveals that they do not have factual content. His proposal that we regard mathematics as a mere instrument would appear to rest upon these two claims. It is because we *can* regard mathematics as a mere instrument for the empirical sciences that the proposal is a live option.

We have discussed in detail why the second part of the claim is problematic. Carnap has not shown that mathematical statements are without content. But the first part of the claim is problematic as well. To begin with, it is difficult to even make precise the claim that mathematical systems are conservative with respect to the empirical language since much of that language is laden with mathematics. Furthermore, allowing a minimal amount of expressive resources within the empirical language, one can show that in many physical systems mathematics is not conservative with respect to the empirical language. For example, in classical mechanics, it is possible to set up a system of perfect elastic reflectors and set a point mass bouncing among them in such a way that one can encode the halting problem. More precisely, for any Turing machine M, one can encode the question "Will M halt on null input?" by the question "Will the particle pass through point p of the output screen?"[52] To be sure, this form of the question involves a minimal amount of mathematics. But it seems clear that it is an (idealized) empirical question. Let M be a Turing machine such that M halts on null input if and only if ZF + AD is inconsistent. Now take the arrangement of perfect reflectors to encode this version of the halting problem. Question:

[52] This result is due to Christopher Moore. See Moore (1990) and Moore (1991). See Pitowski (1996) for an account and a discussion of the philosophical implications.

Does the particle ever pass through the halting point p of the output screen? Mathematical systems are *not* conservative with respect to this empirical statement. For example, it is undecided by PA (assuming PA is consistent) and it is settled negatively by strong systems of set theory.[53]

So both parts of Carnap's claim are problematic. He has given us little reason to adopt his instrumentalist conception. Perhaps then he is merely advancing a proposal.

4.5 Conclusion

Many of the problems raised concerning the earlier account persist in the later account—the problem of free parameters, the problem with the conception of formality, the problem of assessment sensitivity. But more importantly, the later account of mathematical truth is even more problematic than the earlier account since (i) it rests on the problematic notion of "being determined by the semantic rules" and (ii) it rests on a version of logicism that cannot be defended. The route to radical pluralism is equally problematic. In the first pass, Carnap used the distinction between the S-first approach and the C-first approach to setting up semantical systems in attempt to reconcile the pluralists and non-pluralists. The pluralists were correct with regard to the C-first approach, while the non-pluralists were correct with regard to the S-first approach. At least, this is how matters looked at first. But there were residual controversies that arose on the S-first approach. In fact, this is where all of the true mathematical controversies lay. So Carnap took a second pass at the topic. In his attempt to dissolve the residual controversies he suggested that we adopt an instrumentalist conception of mathematics. But we showed that he has not given a defense of this conception and, moreover, that the conception is problematic. In the end, he is in a position to not to defend a thesis, but merely make a proposal. One cannot argue with a proposal. One can merely explain one's reasons for not following it. In the remainder of this paper I will do this and make a counter-proposal. The counter-proposal will turn on a conception of the relationship between philosophy and the exact

[53] Indeed most set theorist who have studied the connection between inner model theory and descriptive set theory are quite confident (for set theoretic reasons) that the answer is 'no'. One system that yields this answer is "ZFC + There is a proper class of Woodin cardinals". If one is bothered by the intrusion of mathematics in the element of idealization, one can simply replace the question with "Will anyone ever write down a correct proof of a contradiction from the axioms ZF + AD?"

science. So let us turn to Carnap's understanding of this relationship.

5 Conception of Philosophy

This brings us finally to Carnap's third philosophical thesis—the thesis that most of the traditional questions in philosophy are pseudo-questions and philosophy is to be regarded as the logical syntax of the language of science. This conception embodies a pluralism more wide-ranging than mathematical pluralism. For just as foundational disputes in mathematics are to be dissolved by replacing the theoretical question of the justification of a system with the practical question of the expedience of adopting the system, so too many philosophical and physical disputes are to be dissolved in a similar fashion:

> It is especially to be noted that the statement of a philosophical thesis sometimes ... represents not an *assertion* but a *suggestion*. Any dispute about the truth or falsehood of such a thesis is quite mistaken, a mere battle of words; we can at most discuss the utility of the proposal, or investigate its consequences. (299–300)

Once this shift is made one sees that

> the question of truth or falsehood cannot be discussed, but only the question whether this or that form of language is the more appropriate for certain purposes. (300)

Philosophy has a role to play in this. For, on Carnap's conception, "as soon as claims to scientific qualifications are made" (280), philosophy just *is* the study of the syntactical consequences of various scientific systems. It would take us too far afield to discuss Carnap's full defense of this minimalist conception of philosophy. Suffice it to say that he is rejecting the pretensions of *first philosophy*, a conception of philosophy that places philosophy before the exact sciences, and embracing a conception of philosophy that largely places philosophy after the sciences.

I want to focus on this distinction between "the question of truth or falsehood" and "the question of whether this or that form of language is the more appropriate for certain purposes" as Carnap employs it in his discussion of metaphysics, physics and mathematics.

But first a preliminary remark: As many Carnap scholars have stressed, Carnap is a Duhemian holist—he believes that when one tests a physical hypothesis experimentally, the entire physico-mathematical system confronts experience as a whole and that when there is a conflict it can be resolved not only by changing a P-rule but also by changing a protocol sentence or even an L-rule.[54] This is in tension with the sharpness of the distinction between purely pragmatic ("external") questions about which system to adopt and substantive theoretical ("internal") questions that have their footing within a system. For in considering a question within a system, we always have the option (by Duhemian holism) of changing the system under pragmatic pressure. Thus, his view has the consequence that there are no *purely non-pragmatic* questions. Pragmatic questions permeate the systems and play a role in any decision.[55] But let us put this aside and turn to examples where we can see the distinction in action.

Let us start with metaphysics. The first example that Carnap gives (in a long list of examples) concerns the apparent conflict between the statements "*Numbers* are classes of classes of things" and "*Numbers* belong to a special primitive kind of objects". On Carnap's view these material formulations are really disguised versions of the proper formal or syntactic formulations "Numerical expressions are class-expressions of the second-level" and "Numerical expressions are expressions of the zero-level" (300). And when one makes this shift (from the "material mode" to the "formal mode") one sees that "the question of truth or falsehood cannot be discussed, but only the question whether this or that form of language is the more appropriate for certain purposes" (300). There is a sense in which this is hard to disagree with. Take, for example, the systems PA and ZFC − Infinity. These systems are mutually interpretable (in the logician's sense). It seems that there is no theoretical question of choosing one over the other—as though one but not the other were "getting things right", "carving mathematical reality at the joints"—but only a practical question of expedience relative to a given task. Likewise with apparent conflicts between systems that construct lines

[54] See pp. 317–319. In particular, see the place where Carnap references Duhem and Poincaré.

[55] This is not to say that the distinction between a change *within* a system and a change *of* system is not clear. (Whether *this* distinction is clear depends on whether one individuates systems in terms of the d-notions or the c-notions. On the former approach, the distinction is relatively clear But on the latter it is unclear since the question of whether "φ is an L-consequence" of a given system is itself sensitive to the system in which one evaluates the statement, and thus the question of whether one has changed the system is itself assessment sensitive.)

from points versus points from lines, or use sets versus well-founded trees, etc. So, I am inclined to agree with Carnap on such *metaphysical* disputes. I would also agree concerning theories that are mere *notational* variants of one another (such as, for example, the conflict between a system in a typed language and the mutually interpretable system obtained by "flattening" (or "amalgamating domains")). But I think that Carnap goes too far in his discussion of physics and mathematics. I have already discussed the case of mathematics (and I will have much more to say about it below). Let us turn to the case of physics.

Carnap's conventionalism extends quite far into physics. Concerning physical hypotheses he writes:

> The construction of the physical system is not affected in accordance with fixed rules, but by means of conventions. These conventions, namely, the rules of formation, the L-rules and the P-rules (hypothesis), are, however, not arbitrary. The choice of them is influenced, in the first place, by certain practical methodological considerations (for instance, whether they make for simplicity, expedience, and fruitfulness in certain tasks). ... But in addition the hypotheses can and must be tested by experience, that is to say, by the protocol-sentences—both those that are already stated and the new ones that are constantly being added. Every hypothesis must be compatible with the total system of hypotheses to which the already recognized protocol-sentences also belong. That hypotheses, in spite of their subordination to empirical control by means of the protocol-sentences, nevertheless contain a conventional element is due to the fact that the system of hypotheses is never univocally determined by empirical material, however rich it may be. (320)

Thus in contrast to the mathematical case where full freedom reigns modulo the constraint of consistency, in physics full freedom reigns modulo the constraint of (accepted) protocol-sentences. This freedom is guided by matters of simplicity, expedience, and fruitfulness, but ultimately the choice of one such framework over another (both reasonably thought to be consistent in the mathematical case and both equally empirically adequate in the physical case) is something that falls within the provenance of persuasion, not reason. The choice of one such framework over another is a pseudo-question. Thus, while in pure mathematics it is convention and question of expedience all the way (modulo consistency), in physics it is convention and question of expedience modulo empirical data.

I think that this conception of theoretical reason in physics is too narrow. Consider two examples: first, the historical situation of the conflict between the Ptolemaic and the Copernican accounts of the motion of the planets, and second, the conflict between the Lorentz's mature theory of 1905 (with Newtonian spacetime) and Einstein's special theory of relativity (with Minkowski spacetime). In the first case, the theories are empirically equivalent to within $3'$ of arc and this discrepancy was beyond the powers of observation at the time. In the second case, there is outright empirical equivalence.[56] Yet I think that the reasons given at the time in favour of the Copernican theory were not of the practical variety; they were not considerations of expedience, they were solid theoretical reasons pertaining to truth.[57] Likewise with the reasons for Einstein's theory over Lorentz's (empirically equivalent) mature theory of 1905.[58]

Likewise I think that his conception of theoretical reason in mathematics is too narrow. I think that there are good mathematical reasons for believing in certain arithmetical sentences that cannot be settled in the standard system of arithmetic PA and there are good mathematical reasons for believing in certain set theoretic statements that cannot be settled in the standard system of set theory ZFC. In particular, I think that the reasons given in Gödel's program for new axioms are theoretical reasons and not mere matters of expedience.[59]

There is something right in Carnap's motivation. His motivation comes largely from his rejection of the myth of the model in the sky. One problem with this myth is that it involves an alienation of truth (to borrow an apt phrase of Tait).[60] However, the non-pluralist can agree in rejecting such

[56]See Janssen (2002) for a discussion of the empirical equivalence of Lorentz's mature theory of 1905 and special relativity.

[57]See Evans (1998) p. 412 and Wilson (1970), p. 109 for six solid reasons.

[58]See DiSalle (2006), Friedman (1999a) and Janssen (2002). To be sure, the full vindication of the Copernican theory over the Ptolemaic theory (and Tychonic theory) came with Galileo (and Newton) and the full vindication of special relativity over Lorentz's mature theory came with general relativity. For my purposes, it is sufficient that the reasons given before the discovery of the telescope (and Newtonian gravitation theory), in the first case, and the discovery of general relativity, in the second case, have *some* force.

[59]For a survey of Gödel's program for new axioms see Koellner (2006) and for further details see the references therein.

[60]See §72.

a myth. Once we reject the myth and the pretensions of first philosophy we are left with the distinction between substantive theoretical questions and matters of mere expedience. I agree with Carnap in thinking that the choice between certain metaphysical frameworks (e.g. whether we construct lines from points or points from lines) and between certain notational variants (e.g. whether we use a typed language or amalgamate types) are not substantive theoretical choices but rather matters of mere expedience. But I think that he goes too far in saying that the choice between empirically equivalent theories in physics and the choice between arbitrary consistent systems in mathematics are likewise matters of mere expedience. In many such cases one can provide convincing theoretical reasons for the adoption of one system over another. To deny the significance of such reasons appears to me to reveal a remnant of the kind of first philosophy that Carnap rightfully rejected.

Just as Kant was right to take Newtonian physics as exemplary of theoretical reason and retreat from a maximalist conception of philosophy—one that placed philosophy before science—I think that we should take Einsteinian physics as exemplary of theoretical reason and retreat from a minimalist conception of philosophy—one that places philosophy after science. Similarly, in the case of mathematics, I think we are right to take the developments in the search for new axioms seriously. But we must ensure that the pendulum does not return to its starting point. The proper balance, I think, lies in between, with a more meaningful engagement between philosophy and science.[61]

[61] I attempt to map out such a conception in Koellner (2009b).

References

Carnap, R. (1934). *Logische Syntax der Sprache*, Springer, Wien. Translated by A. Smeaton as *The Logical Syntax of Language*, London: Kegan Paul, 1937.

Carnap, R. (1939). Foundations of Logic and Mathematics, *International Encyclopedia, vol. I, no. 3*, University of Chicago Press, Chicago.

Carnap, R. (1942). *Introduction to Semantics*, Harvard University Press, Cambridge, Mass.

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

Carnap, R. (1963). Intellectual autobiography, *in* P. A. Schilpp (ed.), *The Philosophy of Rudolph Carnap*, Vol. 11 of *The Library of Living Philosophers*, Open Court, pp. 3–84.

DiSalle, R. (2006). *Understanding Spacetime: The Philosophical Development of Physics from Newton to Einstein*, Cambridge University Press, Cambridge.

Evans, J. (1998). *The History and Practice of Ancient Astronomy*, Oxford University Press.

Friedman, M. (1992). *Kant and the Exact Sciences*, Harvard University Press, Cambridge, Massachusetts.

Friedman, M. (1999a). *Dynamics of Reason: The 1999 Kant Lectures at Stanford University*, CSLI Publications.

Friedman, M. (1999b). *Reconsidering Logical Positivism*, Cambridge University Press.

Friedman, M. (1999c). Tolerance and analyticity in carnap's philosophy of mathematics, *in (Friedman 1999b)*, Cambridge University Press.

Gödel, K. (1953/9). Is mathematics syntax of language?, *in (Gödel 1995)*, Oxford University Press, pp. 334–362.

Gödel, K. (1995). *Collected Works, Volume III: Unpublished Essays and Lectures*, Oxford University Press, New York and Oxford.

Goldfarb, W. & Ricketts, T. (1992). Carnap and the philosophy of mathematics, *in* D. Bell & W. Vossenkuhl (eds), *Science and Subjectivity*, Akadamie, Berlin, pp. 61–78.

Janssen, M. (2002). Reconsidering a scientific revolution: The case of Einstein *versus* Lorentz, *Physics in Perspective* (2): 421–446.

Kleene, S. C. (1939). Review of *The Logical Syntax of Language*, *Journal of Symbolic Logic* **4**(2): 82–87.

Koellner, P. (2006). On the question of absolute undecidability, *Philosophia Mathematica* **14**(2): 153–188.

Koellner, P. (2009a). Strong logics of first and second order, *Bulletin of Symbolic Logic* . Forthcoming.

Koellner, P. (2009b). Truth in mathematics: The question of pluralism, *in* O. Bueno & Ø. Linnebo (eds), *New Waves in Philosophy of Mathematics*, New Waves in Philosophy, Palgrave Macmillan. Forthcoming.

Lindström, P. (2003). *Aspects of Incompleteness*, Vol. 10 of *Lecture Notes in Logic*, second edn, Association of Symbolic Logic.

Moore, C. (1990). Unpredictability and undecidability in dyntamical systems, *Physical Review Letters* **64**(20): 2354–2357.

Moore, C. (1991). Generalized shifts: unpredicability and undecidability in dynamical systems, *Nonlinearity* **4**: 199–230.

Pitowski, I. (1996). Laplace's demon consults and oracle: the computational complexity of prediction, *Studies in the History and Philosophy of Modern Physics* **27**(2): 161–180.

Quine, W. V. (1963). Carnap and logical truth, *in* P. A. Schilpp (ed.), *The Philosophy of Rudoph Carnap*, Vol. 11 of *The Library of Living Philosophers*, Open Court, pp. 385–406.

Stein, H. (1992). Was Carnap entirely wrong, after all?, *Synthese* **93**(1/2): 275–295.

Tait, W. W. (1986). Truth and proof: The Platonism of mathematics, *Synthese* (69): 341–370. Reprinted in (Tait 2005).

Tait, W. W. (2005). *The Provenance of Pure Reason: Essays in the Philosophy of Mathematics and Its History*, Logic and Computation in Philosophy, Oxford University Press.

Väänänen, J. (2001). Second-order logic and foundations of mathematics, *Bulletin Symbolic Logic* **7**(4): 504–520.

Wagner, P. (ed.) (2009). *Carnap's Logical syntax of language*, History of analytic philosophy, Palgrave Macmillan.

Wilson, C. (1970). From Kepler's laws, so-called, to universal gravitation: Empirical factors, *Archive for History of Exact Sciences* **6**(2): 89–170.