Pergamon
www.elsevier.com/locate/shpsa

Stud. Hist. Phil. Sci., Vol. 31, No. 1, pp. 151–172, 2000

Printed in Great Britain
0039–3681/00 $ - see front matter

Rudolf Carnap's 'Theoretical Concepts in Science'

*Stathis Psillos**

1. Editor's Introduction

Rudolf Carnap delivered the hitherto unpublished lecture 'Theoretical Concepts in Science' at the meeting of the American Philosophical Association, Pacific Division, at Santa Barbara, California, on 29 December 1959. It was part of a symposium on 'Carnap's views on Theoretical Concepts in Science'. In the bibliography that appears in the end of the volume, 'The Philosophy of Rudolf Carnap', edited by Paul Arthur Schilpp, a revised version of this address appears to be among Carnap's forthcoming papers. But although Carnap started to revise it, he never finished the revision,[1] and never published the unrevised transcript. Perhaps this is because variants of the approach to theoretical concepts presented for the first time in the Santa Barbara lecture have appeared in other papers of his (cf. the editorial footnotes in Carnap's lecture). Still, I think, the Santa Barbara address is a little philosophical gem that needs to see the light of day. The document that follows is the unrevised transcript of Carnap's lecture.[2] Its style, then, is that of an oral presentation. I decided to leave it as it is, making only very minor stylistic changes—which, except those related to punctuation, are indicated by curly brackets.[3] I think that reading this lecture is a rewarding experience, punctuated as the lecture is with odd remarks and autobiographical points. One can almost envisage

* Department of Philosophy and History of Science, University of Athens, Athens, Greece
Received 28 April 1998; in revised form 1 September 1998.

[1]Document 089–53–06 of the Carnap Archive, University of Pittsburgh, consists of the first three and a half pages of a revised version.

[2]The unrevised transcript is document 089–53–08 in the Carnap Archive; hereafter SB. I have however replaced its first one and a half pages with the three and a half pages of the revised version; cf. note 1.

[3]Document 089–53–08 includes an almost six-page transcript of the question–answer period that followed Carnap's address. I decided to leave this part out, since it adds very little to what there is in the body of the lecture.

PII: S0039-3681(99)00031-X

Carnap standing up and delivering it. I inserted in the text the relevant figures that Carnap drew on the blackboard (doc. 089–54–01), and added a few editorial footnotes with references and other points of elucidation.

The Santa Barbara paper brings together, and presents very clearly, Carnap's final views on theoretical concepts. More importantly, it contains the definitive statement of Carnap's attempt to distinguish between the analytic and the factual content of scientific theories as well as of his attempt to explicitly define theoretical terms, by means of David Hilbert's ε-operator. After a decade of intensive work on semantics and inductive logic, Carnap focused, in the early 1950s, on the status of theoretical concepts in science. His starting point was that an analysis of the language of science, and an account of the meaning of theoretical terms, require a distinction between analytic truths—truths in virtue of meaning—and synthetic truths—truths in virtue of fact. Despite Hempel's and Quine's attacks on the concept of analyticity, Carnap thought that an explication of this concept is imperative for the methodology of science. As he once put it, the analytic/synthetic distinction reflects the difference between 'pure mathematics on the one side and physics, which contains mathematics but in applied form, on the other side' (Carnap Archive, doc. 198–53–08, p. 4). But Hempel had almost persuaded him that such a distinction cannot be drawn within the language of theories, the reason being that the theoretical postulates and the correspondence postulates which constitute a theory play a dual role: they contribute to the meaning of theoretical terms, but they also delineate the empirical content of the theory. Hence, according to the standard empiricist understanding of theories, the view that Carnap himself defended in his 'The Methodological Character of Theoretical Concepts'[4] (MCTC), the meaning-fixing function of the theory cannot be separated from its fact-reporting function. It may not be surprising then that in MCTC Carnap made no attempt to characterise 'analyticity' for a theoretical language. In fact, he was driven towards the conclusion that in a language which contains theoretical terms, the concept of analyticity coincided with the narrower concept of logical truth. In an unpublished precursor of MCTC, he reluctantly noted that in a theory in which the primitive T-terms were quantitative concepts, expressing physical magnitudes, 'analyticity coincides with logical truth' ('Remarks on the Theoretical Language', Carnap Archive, doc. 089–34–06, p. 5).[5]

So, one can imagine Carnap's delight when he at last managed to re-formulate a scientific theory in such a way that it could be separated into two components, one analytic, the other synthetic. Hempel knew of the new idea because of his

[4]Carnap (1956).

[5]In a rather autobiographical note, Carnap stressed: 'Earlier, although I did not share the pessimism of Quine and Hempel, I always admitted that [defining analyticity for the theoretical language] was a serious problem and that I could not see a satisfactory solution. For a while I thought we would perhaps have to resign ourselves to taking a sentence that contained theoretical terms and no observation terms as analytic only under the most narrow and almost trivial condition that it is L-true' (Carnap, 1974, p. 273).

extended correspondence with Carnap on this matter in the 1950s. And Carnap published a short piece explaining his new view in 1958, in German, of which at least Herbert Feigl was aware (cf. the editorial footnotes in the main text). But the Santa Barbara address was the first public announcement in English of all of Carnap's new views.

Carnap's new view on analyticity utilised the so-called Ramsey-sentence, which was first proposed by Frank Ramsey in his paper 'Theories'.[6] In fact, Carnap re-invented what came to be called the 'Ramsey-sentence approach', where all theoretical terms that feature in a theory are replaced by variables, bound by existential quantifiers. He called it 'the existentialised form of a theory', and first presented it at a conference at Los Angeles in 1955. It was only after a belated reading of Hempel's 'The Theoretician's Dilemma' that Carnap realised that his idea had already been suggested by Ramsey.[7] What the Santa Barbara paper shows is how Ramsey's idea can be used to distinguish between the analytic and the synthetic element of a scientific theory. This is entirely Carnap's own contribution. The details are explained in the lecture, but the main point is that a theory TC can be

[6]Ramsey ([1929]1978).

[7]This is a rather fascinating episode in the history of logical empiricism which needs to be highlighted. Here is how Carnap states it (from a letter to Hempel; Carnap Archive, 102–13–53).

February 12, 1958.
Dear Lante:
In the last week I have thought much about you, your ideas, and writings, because I was working at the Reply to your essay for the Schilpp volume. On the basis of your article 'Dilemma' I reworked a good deal of it and some new ideas came in. I think this article of yours is a very valuable work which helps greatly in clarifying the whole problem situation. Originally I read only §§ 6 and 7 because you had commented that they refer to my article on theoretical concepts. Unfortunately I postponed reading the remainder (and thus the last two sections) because I was too busy with other replies for the Schilpp volume.
The case of the Ramsey-sentence is a very instructive example how easily one deceives oneself with respect to the originality of ideas. At Feigl's conference here in 1955 [this is the Los Angeles Conference—S.P.], where Pap, Bohnert and others were present, I represented the existentialized form of a theory as an original recent idea of my own. Sometime after the Conference Bohnert said that he had now remembered having found this idea some years ago and having explained it to me in a letter to Chicago. Although I could not find that letter in the files, I had no doubt that Bohnert was correct, so I ceded the priority to him. He thought more about it and became more and more enthusiastic about this form and he even gave up his old thesis project (on dispositions) and developed new ideas how to use the existentialized form of the theory in order to clarify a lot of methodological problems in science; this he intended to work out as his thesis. Then, I believe it was last summer, when I read the rest of your 'Dilemma', I was struck by your reference to Ramsey. I looked it up at the place you referred to in Ramsey's book, and there it was, neatly underlined by myself. Thus there was no doubt that I had read it before in Ramsey's book. I guess that was in the Vienna time or the Prague time (do you remember whether we talked about it in Prague?). At any rate, I had completely forgotten both the idea and its origin . . .

What exactly is the existentialized form of a theory that Carnap refers to? In the protocol of the Los Angeles conference, Carnap is reported to have extended Craig's theorem to 'type theory (involving introducing theoretical terms as auxiliary constants standing for existentially generalised functional variables in "long" sentence containing only observational terms as true constants)' (Feigl Archive, 04–172–02, p. 14). He is also reported to have shown that '(a)n observational theory can be formed which will have the same deductive observational content as any given theory using non-observational terms; namely, by existentially generalising non-observation terms' (cf. *ibid*., p. 19). There should be no doubt that, inspired by Craig's theorem, Carnap literally re-invented the Ramsey-sentence approach.

written in the following logically equivalent form: ^{R}TC & (^{R}TC ⊃ TC), where ^{R}TC is the Ramsey-sentence of the theory, while the conditional (^{R}TC ⊃ TC) says that *if* there is a class of entities that satisfy the Ramsey-sentence, *then* the t-terms of the theory denote the members of this class. Carnap suggested that the Ramsey-sentence of the theory captured its factual content, and that the conditional (^{R}TC ⊃ TC) captured its analytic content. This is so, Carnap noted, because the conditional (^{R}TC ⊃ TC) has no factual content: its own Ramsey-sentence, which would express its factual content, if it had any, is logically true. He thereby thought that he had solved the problem of 'how to define A-truth [analytic truth] in the sense of analyticity or truth based on meaning also for the theoretical language' (SB p. 12).

A common criticism against analyticity, made by both Quine and Hempel, is that there is no point in distinguishing between analytic and synthetic statements, because all statements in empirical science are revisable: any statement can be abandoned for the sake of resolving a conflict between the theory and the evidence. Analytic statements are perceived as 'inviolable truths'.[8] But since, Hempel said, there are no such truths—'with the possible exception of the formal truths of logic and mathematics'—there is no point in characterising analyticity. However, such criticisms have always misfired against Carnap. Carnap never thought that analyticity was about inviolable truth, 'sacrosanct statements', unrevisability or the like. Instead, he thought that 'the difference between analytic and synthetic is a difference internal to two kinds of statements inside a given language structure; it has nothing to do with the transition from one language to another'.[9] Already in 1937, Carnap noted that no statements (not even mathematical ones) were unrevisable. Anything can go in view of recalcitrant evidence: 'No rule of the physical language is definite; all rules are laid down with the reservation that they may be altered as soon as it seems expedient to do so. This applies not only to the P-rules [theoretical postulates] but also to the L-rules [logical rules], including those of mathematics. In this respect, there are only differences in degree; certain rules are more difficult to renounce than others'.[10] Given his view that the distinction between analytic and synthetic statements can only be drawn within a language, his inability to explicate 'analyticity' *within* the language of science was all the more hurtful. But then, when he managed to draw this line, he was, understandably, very pleased. There might still be independent reasons to jettison analyticity. But the fact that theoretical terms are introduced via theoretical postulates and correspondence rules cannot be one of them.

How did Hempel and Quine react to Carnap's new account? In a note to Carnap, Hempel stressed that 'I find [the explication of analyticity for a theoretical language] very ingenious. Somehow, the use of [^{R}TC ⊃ TC] as a meaning postulate

[8]Hempel (1963), p. 705.
[9]Carnap ([1950]1990), p.431; cf. also Carnap (1963), p. 921.
[10]Carnap (1937), p. 318.

seems intuitively plausible; it is as if you were saying: Granted that the Ramsey-sentence for TC holds true, we want to use the terms of the theory to express somewhat more conveniently just what the Ramsey-sentence tells us' (Carnap Archive, doc. 091–37–05). Quine, too, notes that Carnap proposed 'an ingenious way of separating the factual content . . . from the linguistic, or quasi-definitional component'.[11] But, curiously, neither Hempel nor Quine further discussed the implications of Carnap's new approach for the notion of analyticity. Rather, they re-asserted that there was no point in distinguishing meaning postulates from empirical postulates of a theory, because they were both revisable, and either could go in the light of recalcitrant evidence.[12] Once again, such reaction could not possibly hit its target. Carnap stressed that not all potential changes or re-adjustments of a theory are on a par. When a theory is replaced by a radically different one, then this change amounts to a change of language as a whole. Analyticity is not supposed to be invariant under language-change. In radical theory-change, in a 'revolution' perhaps,[13] the distinction between the analytic and the synthetic has to be re-drawn within the successor theory. Under Carnap's new account of analyticity, there is nothing to stop us from doing just that. But there are also less radical, or minor, re-adjustments and changes which amount to discovering new facts, or to determining the truth-values of several statements entailed by the theory. Such changes present no threat to the analytic–synthetic distinction as drawn within the existing theory. They are made holding fixed the analytic statements. But—and this is the crucial point—when minor changes are due, the basic synthetic postulates are also held fixed: the adjustments are directed to less basic elements of the theory. So, 'being held true, come what may' holds for both the analytic and the synthetic statements, when minor changes are required. And similarly, both analytic statements and basic synthetic postulates may be revised, when radical changes are in order. As Carnap put it: '[T]he concept of an analytic statement which I take as an explicandum is not adequately characterised as "held true come what may"'.[14]

In the Santa Barbara paper, Carnap took one further step. He showed how theoretical terms could be explicitly defined—yes, you heard well—in an extended observational language, which contained the whole of logic and mathematics, plus Hilbert's ε-operator. Carnap presented this idea with caution—and good humour. He then went on to publish it in a paper entitled 'On the Use of Hilbert's ε-operator in Scientific Theories' (cf. the editorial footnotes in the lecture). Given the fact that this paper is little known and difficult to get hold of, the Santa Barbara paper will cast new light on Carnap's views. Carnap wanted to improve on Ramsey's views. He took Ramsey to have shown how we could get rid of the 'bothersome' theoretical terms (SB p. 13). But he thought that this move created certain 'incon-

[11]Quine (1985), p. 330.
[12]Cf. Hempel (1963), p. 705; Quine (1985), p. 331.
[13]In his MCTC, Carnap called such radical changes in theory 'revolutions' (Carnap, 1956, p. 46).
[14]Cf. Carnap (1963), p. 921.

veniences' (SB p. 17). Put simply, Carnap's point was that if we tried to characterise even elementary sentences such as 'The temperature yesterday in Santa Barbara was so-and-so degrees Fahrenheit' à la Ramsey, we would have to cite the Ramsey-sentence of the *whole* theory in which the term 'temperature' featured. According to Ramsey's suggestion, we would have to replace the term 'temperature' with an existentially bound variable, say *u*. But, despite the fact that the expression 'There is a *u* such that its value yesterday in Santa Barbara was so and so degrees Fahrenheit' is a sentence, it fails, as it stands, to capture anything about *temperature*. As Carnap notes, the essential characteristics of the magnitude called 'temperature' are captured by the 'the combinations and connections with other theoretical terms {which} are expressed in the T-postulates, and the combinations and connections with the observation terms, {which are} expressed in the C-postulates. So you must give *them* in order to give the full sentence' (SB pp. 17–18). So, the Ramsey-sentence characterises theoretical terms in a *holistic* way: it can only specify the meaning of a set of theoretical terms as a whole, by stating their mutual connections as well as their links with observational terms. In the Santa Barbara lecture, Carnap tried precisely to devise a way in which empiricists can avoid the holism implicit in the Ramsey-sentence approach, without using resources other than logic, mathematics and an observation language. The details, which rely on Hilbert's selection operator ε, can be found in the text.[15] But the basic idea is the following. The ε-operator is defined by one axiom: $\exists x Fx \supset F(\varepsilon_x Fx)$. This simply means that *if* anything has the property F, *then* the entity $\varepsilon_x Fx$ has this property. $\varepsilon_x Fx$ may be thought of as the ε-representative of the elements of a non-empty class F, without further specifying which element it stands for.[16] Let, now, the t-terms of the theory TC form an n-tuple $t = \langle t_1, \ldots, t_n \rangle$. Hilbert's ε-operator allows us to select an arbitrary class among the classes of entities which satisfy the theory such that the n-tuple t of t-terms designate this class. That is, the n-tuple t of t-terms designate the ε-representative of the classes of entities which satisfy the theory. Then, each and every t-term of the n-tuple is explicitly defined as the ε-representative of the i-th member of the n-tuple. The theory can still be split up into two parts, one analytic, the other synthetic. The synthetic part is still the Ramsey-sentence $^R TC$ of the theory. But the holistic—and cumbersome—meaning postulate ($^R TC \supset TC$) is now replaced by n + 1 explicit definitions of each and every one of the n t-terms of the theory.[17] Carnap showed that this new way to characterise the analytic component of the theory logically implies the meaning postulate ($^R TC \supset TC$). Its

[15]Hilbert's ε-calculus is developed in Hilbert and Bernays ([1939]1970), section 1. An illuminating discussion is given in Fraenkel and Bar-Hillel (1958), pp. 183–185.

[16]Fraenkel and Bar-Hillel (1958) stress that there is a close connection between the ε-operator and the axiom of choice. However, they add that the ε-operator is not a generalisation of the axiom of choice, since 'the ε-formula allows for a single selection only, while the axiom of choice allows for a simultaneous selection from each member of an (infinite) set of sets and guarantees the existence of the set comprising the selected entities' (p. 184).

[17]There are n definitions for each of the n t-terms and one for the n-tuple.

sole advantage lies in the fact that it provides an explicit definition of each and every t-term of the theory. My own view is that Carnap preferred the logically stronger version of the theory—the one based on the ε-operator—because he thought that, being not holistic, it could be used to restore some criterion of *atomistic significance* for theoretical terms. To be sure, each and every theoretical term is explicitly defined *relative* to the n-tuple t of the theoretical terms of the theory. Still, however, relative to this n-tuple, the meaning of each and every theoretical term of the theory can be disentangled from the meanings of the rest.

Isn't there an outright contradiction in Carnap's views? The claim that theoretical terms are introduced by means of theoretical postulates and correspondence postulates is based on the fact that t-terms cannot be explicitly defined in an observational language: they always have 'excess content'. How can this be reconciled with the view currently expressed? Carnap tried to deal with this objection in the Santa Barbara paper. He pointed out that the ε-operator has the peculiar feature of being an 'indeterminate constant'. It can *fully* specify the designata of t-terms only to the extent that there is a non-empty class of entities which satisfies the theory and that this class has *exactly one* representative; that is, only to the extent that uniqueness is assumed. Carnap did not want to build into his approach the requirement of unique realisation.[18] He thought that the indeterminacy associated with the ε-operator was good for the methodology of science, since it allowed for a future narrower (and better) specification of the designata of t-terms, by the addition of new theoretical postulates and correspondence rules. For if a set of t-terms are defined such that they refer to the unique set of entities that satisfy the postulates of the theory, then when new postulates are added to the theory, the reference of the t-terms will have to change.[19] So, Carnap noted, his ε-definition gives 'just so much specification as we can give, and not more. We do not want to give more because the meaning should be left unspecified in some respect, because otherwise the physicist could not—as he wants to—add tomorrow more and more postulates,

[18]It should be noted that in the ε-calculus, the so-called uniqueness (or ι-)operator (the equivalent of the definite article) can be easily defined: if there exists only one entity satisfying Fx, then 'ε_xFx' is to be read as 'the entity having the property F' (cf. Fraenkel and Bar-Hillel, 1958, p. 184). So, in a sense, the ε-operator characterises an *indefinite description*, whereas the ι-operator characterises a definite one. In his 'How to Define Theoretical Terms', David Lewis has modified Carnap's approach based on the ι-operator. He therefore insists on the uniqueness requirement. He suggests that if there is no unique realisation of the theory, the t-terms should be considered denotationless. His motivation for this claim is that this is the lesser of two evils. In case of multiple realisation, he notes, there is no non-arbitrary way to pick one realisation. So we are forced to either accept that t-terms do not name anything, or that they name the elements of one arbitrarily chosen realisation. Lewis thinks that 'either of these alternatives concedes too much to the instrumentalist view of a theory as a mere formal abacus' (Lewis, 1970, p. 432). For a recent modification of Lewis's views which brings together the Ramsey-sentence approach with the thought that there is some vagueness associated with the meaning of theoretical concepts, see Papineau (1996).

[19]This point had also been anticipated by Ramsey. He rejected a theory of meaning of t-terms based on explicit definitions because it did not do justice to the fact that theoretical concepts in science were taken to be open-ended: they were capable of application to new empirical situations. As he noted: '[I]f we proceed by explicit definition we cannot add to our theory without changing the definitions, and so the meaning of the whole' (Ramsey, [1929]1978, p. 119).

and even more and more correspondence postulates, and thereby make the meaning of the same term more specific than {it is} today'. And he concluded: 'it seems to me that the ε-operator is just exactly the tailor-made tool that we needed, in order to give an explicit definition, that, in spite of being explicit, does not determine the meaning completely, but just to the extent that is needed' (SB pp. 21–22). The readers will certainly make up their own minds as to whether Carnap's attempt is successful. But it seems to me that Carnap was after a theory of *reference* which would allow for referential continuity in theory-change (at least in non-revolutionary theory-change). He thought he found the elements of this theory in his ε-operator approach: Hilbert's device fixes the designata of t-terms as the entities which realise the theory; but what exactly these entities are, what it is true of them, and so on, are issues that are left open for further scientific investigation.

2. 'Theoretical Concepts in Science' by Rudolf Carnap

For many years it has been found useful in the analysis of the language of science to divide the terms of the language into three main kinds: logical terms (including those of pure mathematics), observational terms or O-terms, and theoretical terms or T-terms (sometimes called 'constructs'). It is true that it is hardly possible to draw a clear-cut boundary line between O-terms and T-terms. The choice of an exact line is somewhat arbitrary. Still, from a practical point of view, the distinction is clear enough between terms like 'blue', 'red', 'hard', 'soft', 'cold', etc. on the one hand (here understood not as terms for sensory qualities, but for properties of observable things and for relations among things, e.g., 'x is warmer than y'), and, on the other hand, terms like 'electro-magnetic field', 'electric charge', 'protons', 'neutrons', and so on—terms which occur in theoretical science and for which we cannot claim that we have knowledge by direct observation. With respect to the sentences of the language we make a three-fold division:

(1) the logical sentences, which contain no descriptive terms,
(2) the observational sentences or O-sentences, which contains O-terms but no T-terms,
(3) the theoretical sentences or T-sentences, which contains T-terms; these are subdivided into:
  (3a) mixed sentences, containing both O- and T-terms, and
  (3b) pure T-sentences, containing T-terms, but no O-terms.

The whole language L is divided into two parts. Each part contains the whole of logic (including mathematics); they differ with respect to the descriptive (i.e., non-logical) sentences:

(1) the *observational language* or O-language (L_O), containing (besides logical sentences) only O-sentences; hence, no T-terms
(2) the *theoretical language* or T-language (L_T), containing (besides logical sentences) only T-sentences (with T-terms, with or without O-terms).

I have sometimes made a distinction between the restricted observation language which contains only a first-order logic with observable objects as individuals and the extended observation language which contains a comprehensive logic including the whole of mathematics, either in set-theoretic form or in type-theoretic form. This distinction is of interest for certain problems, e.g., those of finitism and constructivism. For our present discussion, however, it does not seem necessary. [For this reason I use here the terms 'observational language' and 'L_O' not for the restricted language (as I did previously) but for the extended language.][20]

Descriptive *primitive terms* in L:

Observational	Theoretical
$O_1, \ldots, O_m$	$T_1, \ldots, T_n$

The T-terms are introduced by a theory based on postulates of two kinds. The *theoretical postulates* or T-postulates, e.g., laws of theoretical physics, are pure T-sentences. The *correspondence postulates* or C-postulates are mixed sentences, because they combine T-terms and O-terms. They constitute what Campbell called the dictionary between the two languages, what Reichenbach called coordinative definitions of terms occurring in axiom systems of theoretical science, and {what} in Bridgman's terminology might be called operational postulates or operational rules.

Descriptive sentences in L:

Observational	Mixed	Theoretical
	C−postulates	T−postulates
L_O		L_T

This is the distinction which we make between terms and between sentences. One of the most important characteristics of the T-terms, and therefore of all sentences containing T-terms—at least if they occur not in a vacuous way—is that their interpretation is not a complete one, because we cannot specify in an explicit way by just using observational terms what we mean by the 'electromagnetic field'. We can say: if there is a distribution of the electromagnetic field in such and such a way, then we will see a light-blue, and if so and so, then we will see or feel or hear this and that. But we cannot give a sufficient and necessary condition entirely in the observational language for there being an electromagnetic field, having such and such a distribution. Because, in addition to observational consequences, the content is too rich; it contains much more than we can exhaust as an observational consequence.

So this is the original set-up, and on the basis of {it} we want to make a distinction between logical truth and factual truth. I believe that such a distinction is very

[20]Carnap refers to sections II and IX of his 'The Methodological Character of Theoretical Concepts' (Carnap, 1956, pp. 38–76).

important for the methodology of science. I believe that { . . . }[21] the distinction between pure mathematics on the one side and physics, which contains mathematics but in applied form, on the other side, can only be understood if we have a clear explication of the distinction that in traditional philosophy is known under the terms analytic and synthetic, or necessary truth and contingent truth, or however you may express it.

Quine has pointed out that here a new distinction should be made: logical truth in the narrower sense comprises those sentences whose truth is established in deductive logic plus substitution-instances of them, which may contain descriptive terms. {But} there are other sentences which we may regard also as analytic: his example, as you may remember, is: ‘No bachelor is married’—that is certainly true, but its truth is not a matter of the contingent facts of the world; it is a matter merely of the meaning of the terms. {This sentence} is true in virtue of the meanings of the terms, but in distinction to logical truth in the narrower sense, here also in virtue of the meaning of the descriptive terms. If we allow the meaning of ‘bachelor’ and the meaning of ‘married’—or at least {if} we know, or are told, by somebody who understands this language that these two terms are incompatible—then we know that the sentence ‘No bachelor is married’ is true in virtue of meanings alone. So it is analytic, as Quine would propose to distinguish: logical truth in a narrower sense or logical truth in a wider sense, or analytic. For the first I will take the term L-true as the term for the explicatum; for the second, A-true. My main purpose here is to indicate how we can make the distinction between A-true and other-true, namely factual-true sentences, not only in the observation language but also in the theoretical language. In the observation language we know a way of doing it—I explained that years ago in a paper called ‘Meaning Postulates’.[22] I would now call them A-postulates. They say that every term that is a logical consequence of the A-postulates is then A-true. This can easily be done in a language like the observation language, where we presuppose that we are in the possession of a complete interpretation of the terms. That need not be done in an explicit way by semantical rules. You just ask somebody: ‘Is this part of the English language completely understood by you, do you know what you mean by the words which you use there?’ {A}s I said before, the terms of the theoretical language are not completely interpreted. The interpretation which they have, is not learned in the same way as the interpretation of terms like ‘red’ and ‘blue’ which we learn, let’s say, as we learn our mother tongue, by hearing how they are applied and then imitating these applications and making an unconscious general inductive inference—and so we know now what we mean by ‘red’, ‘hot’, and so on. But with ‘electromagnetic field’ it is different. There we cannot simply point and say: by the ‘electromagnetic field’ we mean this and that; or an electromagnetic field having

[21]I have crossed off the words: ‘in order to understand’.
[22]Carnap (1952).

an intensity of so and so much or a vector so and so. We cannot simply point and thereby learn it. We learn it by the postulates. These terms are introduced by the postulates; namely the T-postulates, general laws of physics, which connect these terms among each other—which obviously is not sufficient to give any meaning to them—and then the second kind of postulates, the C-postulates, which connect these terms with those of the observation language.

For instance, the term 'temperature' is connected by a C-postulate with observational terms, namely a C-postulate which describes how you proceed in constructing a thermometer and constructing its scale, then putting the thermometer into a certain liquid, and then reading a certain number. Then you are told: Take that number as the value of the temperature of that liquid. So, here we are given rules which connect certain observables with a theoretical term like 'temperature'. {T}his term thereby obtains a partial interpretation. Partial, because not in all occurrences of the term 'temperature' can we use this operational definition. An ordinary thermometer works only in a rather limited interval of the scale; for too low temperatures, for too high temperatures, we must use entirely different methods. So each C-postulate applies only to certain cases and all of them together would not help us to determine temperatures or electric fields and so on unless we had also the T-postulates. So it is then the T-postulates together with the C-postulates which give interpretation, all of {the} interpretation that the T-terms have, which is not a full interpretation. That we have to keep in mind. But what interpretation they have, they get by the postulates and by the postulates of these two kinds together.

When the question is raised how to distinguish between sentences whose truth is due to meaning and other sentences, then—as Hempel has especially clearly pointed out—there is a great difficulty in the theoretical language. Hempel was—with some hesitation, I believe—willing to accept the distinction with respect to the observation language. On the one hand, he is influenced by my way of thinking—we are old friends from the days in Germany.[23] {O}n the other hand, he is also influenced by Quine's scepticism with respect to making a clear distinction between factual and logical truth, or meaning truth. But he pointed out that he can hardly imagine how a distinction could be made also with respect to the theoretical language, {in} any sentences containing either theoretical terms or theoretical and observation terms, for the following reason. The interpretation of the T-terms is given by these postulates, not by explicit definitions on the basis of the observation language. But these postulates have a dual role. They have two different functions {which} each of them—or their totality—fulfils simultaneously, namely they give some meaning to the term and they give some factual information to us. That they give factual information is seen from the fact that if the physicist gives his whole theory to us, then we are in a much better position to predict the weather of tomorrow than if we rely only on a few generalisations which can be formulated in the

[23]Two question marks appear above the word 'Germany' in the typescript.

observation language: If you see clouds of such a shape, then tomorrow it will probably rain, or something of that kind. So the theoretical system of physics certainly gives factual information. But it gives factual information and specification of meaning simultaneously. And Hempel said: ‘That makes the concept of A-truth entirely elusive, because {it} is hardly to be imagined that we could split up these two functions of the T- and C-postulates, so that we could say: this part of them contributes to meaning, therefore the sentences which rely on that part are then, if they are true, true due to meaning only, the other are factual sentences.’[24] So this is the big problem for which I want to present a solution, namely, how to define A-truth in the sense of analyticity or truth based on meaning also for the theoretical language.

In order to do that, I first will speak of a device that has already been introduced a long time ago by Ramsey.[25] It is the so-called Ramsey-sentence, as we call it today, corresponding to any given theory in the theoretical language or mixed sentences. Given the theory, {which} I write TC for short, or in a slightly more explicit form if the observational terms occur in it—of course, they occur only in the C-postulates—but let T be the conjunction of the T-postulates and C of the C-postulates, then TC of both together, then, in order to indicate the descriptive terms, let’s write {the theory} in the following form: $TC(t_1, \ldots, t_n; o_1, \ldots, o_m)$ (the theoretical terms t_1 and so on, t_n, and the observational terms o_1, and so on, o_m). We form from this the Ramsey-sentence in the following way. We keep the observation terms unchanged, but we replace the theoretical terms by variables. Let’s say, for the constant t_1 we put the variable u_1, for t_2, u_2, and so on—if it is in a type system, then they must be variables of the corresponding types—and then we prefix the whole by existential quantifiers, one for each of these n variables, corresponding to the n theoretical terms. So we presuppose that there is a finite number of theoretical terms in that language.

$$TC(t_1, \ldots, t_n; o_1, \ldots, o_m)$$
$$^{R}TC: (\exists u_1) \ldots (\exists u_n) TC(u_1, \ldots, u_n; o_1, \ldots, o_m)$$

Ramsey showed that this sentence is what I would propose to call for short O-equivalent, or observationally equivalent, to the original theory, namely the total theory of the T- and C-postulates, in the following sense. We will say that the sentence S is O-equivalent to a sentence S′ if all the observational sentences, that

[24]I have been unable to locate the exact reference of this quote. Perhaps it comes from a letter from Hempel. But Hempel has made essentially the same point in print elsewhere. ‘[I]t even appears doubtful whether the distinction between analytic and synthetic can be effectively maintained in reference to the language of empirical science’ (Hempel, 1954). Elsewhere he notes: ‘Thus, the only sense in which the concept of analyticity remains applicable to the sentences of a scientific theory is the narrow one of truth by virtue of being an instance of a logically valid schema’ (Hempel, 1963, pp. 704–5). A draft of Hempel’s piece for the Schilpp volume (Hempel, 1963) was available to Carnap as early as 1954.

[25]Frank Ramsey gives a programmatic account of the view that Carnap attributes to him in his posthumously published piece ‘Theories’ (Ramsey, [1929]1978). Ramsey noted: ‘The best way to write our theory seems to be this $(\exists \alpha, \beta, \gamma)$: dictionary · axioms’ (see p. 120).

is, the sentences in L_O' which follow from S, follow also from S′. So, as far as observational consequences are concerned, S and S′ are equivalent. That is {what is} meant by O-equivalent. Now Ramsey showed that this existential sentence—which we call now the Ramsey-sentence—is O-equivalent to the theory TC. And he made the following practical proposal. He said: The theoretical terms are rather bothersome, because we cannot specify explicitly and completely what we mean by them. If we could find a way of getting rid of them and still doing everything that we want to do in physics with the original theory, which contains these terms, that would be fine. And he proposes this existential sentence. You see, in the existential sentence the T-terms no longer occur. They are replaced by variables, and the variables are bound by existential quantifiers, therefore that sentence is in the language L_O', {i.e.,} in the extended observation language. And he said: let's just forget about the old formulation TC about the T-terms; let's just take this existential sentence, and from it we get all the observational consequences which we want to have, namely, all those which we can derive from the original theory.

The form of the system which I propose makes essential use of the Ramsey-sentence, but I do not want to take this radical step, at least not here in this form which I shall describe now. I rather say: let's keep the theoretical terms, let's keep the old form. But then I make certain distinctions, in order to make the desired distinction between A-truth and factual truth. I make that in the following way: I wish to split up the theory TC, the theory in its ordinary form with the T-postulates and the corresponding postulates C, in another way, not into T and C, {but} into two sentences. { . . . } One represents the factual content—and I call that P; that is then the physical postulate, and that is synthetic. { . . . } The other {is} only an A-postulate, which gives only meaning specifications—partial meaning specifications, no more is possible here. {This sentence} does not convey factual information—and I call that A_T. By A_O I mean the conjunction of all the A-postulates, or meaning postulates, which we had in the observation language—I will not give them, something like 'no bachelor is married' might occur there, or 'warmer is a transitive relation' or something of that kind. And then for the T-terms I will give a postulate, which I will call A_T. After having specified these two, then I will define A-truth in the following way: A sentence is A-true, if it is a logical consequence of A_O and A_T together. I use there the assertion symbol of the *Principia* for L-true. So that is expressed by what is written there with this symbol. If A_O and A_T together L-imply S, in that and only in that case will I say that S is an A-true sentence.

	TC	
Problem:	P	A_T
	(physical postulates)	(A-postulates)
	(synthetic)	(analytic)
Answer:	^{R}TC	$^{R}TC \supset TC$

Form I. Postulates: A_O, T, C.
Form II. Postulates: A_O, A_T, P.

S is A-true in L $=_{Df} \vdash A_O \,\&\, A_T \supset S$.
S is P-true in L $=_{Df} \vdash A_O \,\&\, A_T \,\&\, P \supset S$.

Now I must specify what I mean by P and by A_T and then show, first, that the two together really are no more and no less than the old theory, so that I can regard them as a splitting of the old theory into parts; and second, that the one really is merely factual and the other merely {a} specification of meaning. I take as the P-postulate the Ramsey-sentence of the theory, which I write as TC with an upper left superscript R. So ^{R}TC is the Ramsey-sentence {which} I take as my P-postulate, as the physical postulate, as the factual content.

And then as the A-postulate A_T I take a conditional sentence, namely, if ^{R}TC then TC, where the if–then is material implication. So these are the two postulates which I propose—but this is merely a reformulation of the old theory, it is logically equivalent to the old theory. I want to stress especially this point: I do not propose a new theory. If a theory is given, I merely split it another way into three kinds of postulates, {which} I call {the} corresponding forms. I call the first Form I, that is the customary form, namely, A_O, the meaning postulates of the observation language, T, and C. Form II {is}: A_O unchanged, A_T and P. A_T takes something from C and something from T. {S}o A_T and P are not simply parts of T and C as they occur in the ordinary formulation, but are entirely reformulated, as you see for instance that P is the Ramsey-sentence. So the whole of T and C occur there, but in a changed form, with the theoretical terms having been eliminated and replaced by variables.

Does this system of A_T and P really fulfil the purpose, which I said it should fulfil? I will not show it, but it can be shown in a very simple way that the following three results hold:

(a) $\vdash (TC \equiv P \,\&\, A_T)$
(b) P is O-equivalent to TC
(c) $\vdash {}^{R}A_T$

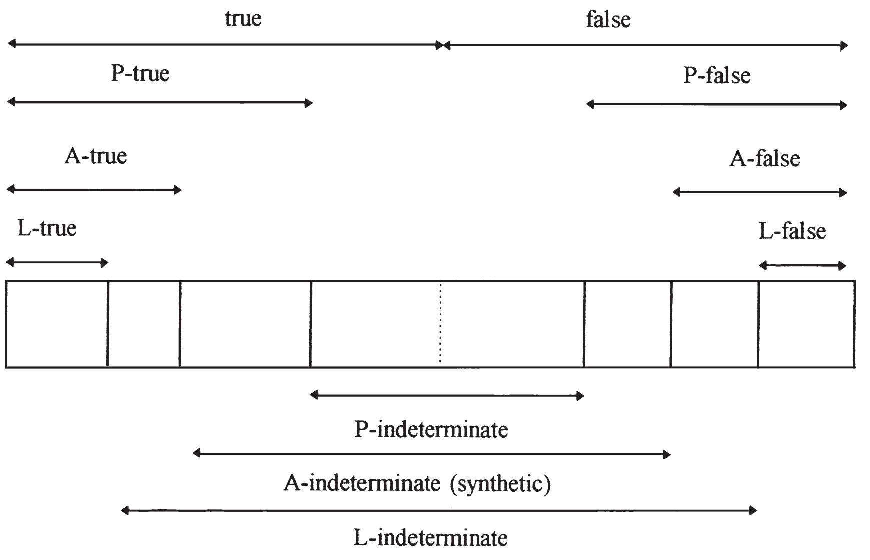
The first is that TC is logically equivalent to P and A_T. In the one direction it follows directly from modus ponens, and in the other direction, obviously from TC we can derive any conditional that has TC as a consequence. So this is quite clear. In other words, P and A_T together is just a reformulation in another form of the old theory.[26] Second, P is O-equivalent to TC, and P is—as I said before—in the extended observation language. It does not contain the theoretical terms, but it contains all what we might call the observational content. Therefore it seems to me that P really fulfils its role. It gives us the factual information as far as obser-

[26]Note that $A_T \equiv {}^{R}TC \supset TC$ and $P \equiv {}^{R}TC$. Given that the theory TC implies its Ramsey-sentence ^{R}TC (i.e., given that $TC \supset {}^{R}TC$), it can be easily shown that $TC \equiv {}^{R}TC \,\&\, ({}^{R}TC \supset TC)$.

vations are concerned and it certainly does not give us any specification of the meaning of the T-terms because they do not occur at all in P.

On the other hand, A_T is of such a kind that as result (c) says: the Ramsey-sentence of A_T—so imagine A_T: $^R TC \supset TC$ {is} now Ramsey-ized {as a whole}. In the second part we have some T-terms occurring, replace them by variables v_1, v_2, and so on, and then put existential quantifiers not before that part but before the whole sentence. That is then the Ramsey-sentence; not of TC, but the Ramsey-sentence of A_T. And it turns out that that sentence is logically true. Since it is a Ramsey-sentence, it is in the observation language. And it is logically true in the observation language, so it does not have any factual content. In other words: A_T does not say anything about the world of facts. All that it does is: it gives us some connection between the terms, namely the T-terms among themselves and the T-terms with the C-terms, of such a kind that it helps to give a partial interpretation for it. That is the purpose of an A-postulate. So in this way I propose to write the theory in the second form {i.e., Form II: $\langle A_O, A_T, P\rangle$, where $A_T \equiv {}^R TC \supset TC$ and $P \equiv {}^R TC$}. I do not say that this form is essentially superior. I do not say: let's forget the old form {i.e., Form I: $\langle A_O, T, C\rangle$} and only use this one. The old form is very convenient and for many purposes perhaps more convenient, because it is the customary form. We find there the Maxwell laws and the law of gravitation and such and such physical laws in their customary form, and then we have the C-postulates in their customary form—so that is certainly a very convenient form. The second form has only this purpose: if we want to make the distinction between A-truth and factual truth, then this form shows this interpretation in a clearer way. Once this interpretation has been made, then we might also introduce the term of P-truth: all those sentences of the total language L, which are logical consequences of A_O and A_T, (that is, all the A-postulates of all the parts of the language), and P, (the physical postulates), are called P-true, physically true, or factually true.

Here we have then the diagram of the classification of the sentences of the total language. They fall, of course, into true and false sentences. Among the true sentences we have a small sub-class, the L-true; a somewhat larger—including the L-true—the A-true (this is the analytic class); {and} then the P-true {sentences}. And on the other hand, correspondingly, L-false, A-false, and P-false; and intermediate{ly} then the indeterminate sentences, L-indeterminate, A-indeterminate, which means then synthetic, and P-indeterminate, which means so-to-speak contingent, not determined either negatively or positively, by the basic physical laws. So, this is the classification of the sentences which we have.



I will make a reference to publications. The Ramsey-sentence has been published posthumously in the book *Foundations of Mathematics* which was published in 1928—it was written some years before that. It has found very little attention until the very last years. Braithwaite refers to it and discusses Ramsey's method in his book,[27] but otherwise very little is to be found in the literature. But then Hempel, in a paper 'The Theoretician's Dilemma' which was published in the second volume of the Minnesota Studies for the Philosophy of Science, emphasised the great importance of the Ramsey-sentence, and discussed a number of methodological questions and logical questions of the language of science on the basis of the Ramsey-sentence.[28] On the other hand, he used it also in order to raise unsolved problems, and among them the problem which I mentioned, which he thought might perhaps not be solvable, namely, making a distinction between analytic and synthetic. Now, in addition, he has a more detailed discussion of the whole in an unpublished paper, which will appear in the Schilpp volume on my philosophy, which we hope will appear toward the end of 1960—that is not yet quite determined.[29] What I just explained, my explication of A-true also including the theoretical language, is contained in my unpublished reply to this paper by Hempel, which will also appear in the Schilpp volume.[30] A much briefer discussion of it has been published in a paper in German 'Beobachtungssprache und theoretische sprache', which has been published first in the periodical *Dialectica* (published in Zurich

[27]Braithwaite (1953), pp. 80–81.
[28]Hempel (1958), sections 9 and 10.
[29]Hempel (1963).
[30]Hempel's piece that Carnap refers to appears in the Schilpp volume (Hempel, 1963). Carnap's own reply 'Carl C. Hempel on Scientific Theories' appears as section 24 in his 'Replies and Systematic Expositions' (Carnap, 1963).

one year ago), and then the whole double-issue of *Dialectica* is separately published as a Festschrift for Paul Bernays under the title *Logica*.[31]

Now I want to raise some questions which I think are interesting from a philosophical point of view. It is the question that was already raised by Ramsey, namely, could we perhaps in some way get rid of the bothersome theoretical terms and restrict ourselves to the observation language? He did not yet make the distinction between L_O' and L_O, but, I think he might have envisaged when he said 'observation language' the whole L_O'. From now on, when I say 'observation language' I mean it in the wider sense, including as descriptive terms only the simple observation terms, but a very rich logic. I shall use a special logical symbol which I shall explain soon, which comes from Hilbert. Can we restrict ourselves to the observation language and still do everything that physicists want to do? This is the practical question. Ramsey's proposal is, of course, one way of doing it, namely using the Ramsey-sentence instead of TC. But that has certain—not essential—objections, only strong inconveniences. Think of the following fact: If somebody asks a physicist: Give me the whole of your theory, then really Ramsey is right: it does not make for much greater inconvenience whether he gives it in the old form—{a} long series of sentences, {of which} he says: These are my theoretical postulates, and a still longer series of correspondence postulates—or whether he makes it a little bit longer by prefixing some let's say 20 existential quantifiers and replacing some constants by variables in it. But if we now think of those sentences, which are much more frequent, when you read in the paper that the temperature yesterday in Santa Barbara was so many degrees and then a prognosis for tomorrow, {e.g. that}, tomorrow the temperature probably would be so many degrees, how would we express that in L_O'? We have no symbol there for 'temperature'. 'Temperature' in the old language was perhaps t_8, let's say, just the 8-th theoretical term. It has disappeared now: we are in L_O'. But there we have a variable u_8, which takes its place. But, in order to use it and say: u_8 for such and such geographical coordinates at such a time point is l00 degrees, we have to write all the n, let's say 20, existential quantifiers, all the theoretical postulates, all the correspondence postulates, and then in the same operand, which is the common operand for all the 20 existential quantifiers, {we can say that} u_8 for such and such coordinates is 100 degrees. Because if you were merely to write u_8 of such and such coordinates, that would not even be a sentence, because there is a free variable in it; it would not mean anything. And it would not help to just add the one quantifier, because that does not tell you that it is temperature. {The} essential characteristics {of 'temperature'} come from the combinations and connections with other theoretical terms {which} are expressed in the T-postulates, and the combinations and connections with the observation terms, {which are} expressed

[31]Carnap (1958). This piece has been translated into English as 'Observation Language and Theoretical Language' (Carnap, 1975).

in the C-postulates. So you must give *them* in order to give the full sentence. Now that is, of course, rather cumbersome and we would like very much to get rid of the T-terms and still have simple sentences for those simple sentences, which the physicist uses every day. Can we do that?

Well, we could do it if we found a way of giving explicit definitions for all the theoretical terms in the observation language. And this is the question which I want to raise now: is that possible? I thought very briefly about that question years ago and I just dismissed it from my mind, because it seems so obvious that it is impossible. Everybody knows that the theoretical terms are introduced by postulates just because we cannot give explicit definitions of them on the basis of the observational terms alone, even if we add a strong logic. At least, that seemed to be the case and therefore I did not think more about it, although, if we could do it, that would be a great advantage.

Now, it is possible. I found that only a few weeks ago and I hope I have not made a mistake—I have not discussed it yet with friends, except for telling David Kaplan about it, but only briefly—so I will present it here and if somebody can show that I am mistaken I shall be very glad to learn it, before I take all the trouble of writing it in a paper—or the trouble for my wife of transcribing all that is here now on the tape. So, in the hope that there is something in it, I will now present the way of doing this by explicit definitions, which is really so surprising that I still can hardly believe it myself.[32]

Before I do it, I will introduce a simplifying notation in the old language for TC. I write t for the n-tuple of the T-terms, the ordered n-tuple $\langle t_1, \ldots, t_n\rangle$. I write o for the m-tuple of the O-terms o_1 down to o_m. The Ramsey-sentence has then the simple form: there is a u such that TC(u,o). My old A_T postulate, in the theoretical language, has the form: if there is a u such that TC(u,o), then TC(t,o).

$$t =_{Df} \langle t_1, \ldots, t_n\rangle.$$

$$o =_{Df} \langle o_1, \ldots, o_m\rangle.$$

$$A_T\text{: } (\exists u)TC(u,o) \supset TC(t,o).$$

Let's make it clear to ourselves what really in effect is said by this A-postulate in the Form II. It says: if there is at least one theoretical entity u, if it exists at all, such that it has the relation TC to o—of course, you remember, u is really now an abbreviation for all the theoretical terms, so {the A-postulate in Form II} means: if these 20 entities exist, which have such and such relations among themselves, and such and such other relations to the observational entities, then let the terms t_1, t_2, and so on, down to t_n, be understood in such a way that this n-tuple is one of those which are in that not-empty class. If the class of those n-tuples is

[32]Carnap eventually published a paper on the ε-operator: Carnap (1961).

not empty, then let t be one of them. This is the meaning of A_T. We will come back to that in a moment. {This} intuitive meaning of A_T suggested to me the way in which to give explicit definitions.

In order to give those explicit definitions, I make use of a logical constant, that has been introduced by Hilbert, and extensive use of it has been made in the work by Hilbert and Bernays, *Die Grundlangen der Mathematik*. It is discussed in great detail at the beginning of the second volume. It is the so-called ε-operator, as Hilbert calls it. He writes an expression of it with ε_x, followed by a sentential formula, containing x as a free variable, let's say, Fx. $\varepsilon_x F(x)$, roughly speaking, means this: if Fx is not empty, if there is something that is F, {if that is} the class F is not empty, then $\varepsilon_x Fx$ stands for any element of that class; {but} it is not specified which one. You see, this is useful in mathematics, because according to mathematical reasoning, we often do the following. As we have no example of an instance for a certain class, for a certain property of natural numbers or property of real numbers, but we have proof that that class is not empty—by showing that by assuming that it were empty, it would lead to a contradiction, or in some other way—or even if we have instances, but we do not bother to specify which one we mean, then we say: let A be any one element of that class, {by which I mean that} I will now go on under the assumption that A is an element of the class F and {that} I will not presuppose anything else about A. All I will presuppose about A is that it is an element of the class, and then I go on to draw my conclusions from it. And indeed this ε-operator has been found extremely useful, especially in metamathematical considerations. Hilbert and Bernays give a detailed discussion of its value and its use. They {first} make use of it, and then later {they show} its eliminability, in order to show that it is not essential, that it is introduced for convenience, but we can dismiss it and still prove the same theorems for another mathematical system, which does not contain it.[33] If we use this symbol, we don't need to {admit} the ordinary quantifiers, existential and universal, as primitive. Hilbert {introduced an} axiom which says {that} if x is an F, in other words, if we know at all that there is something which is F, then $\varepsilon_x Fx$ is an F.

$$\text{Hilbert}\begin{cases}\text{axiom:} & Fx \supset F(\varepsilon_x Fx) \quad (H_1)\\ \text{definitions:} & (\exists x)Fx \equiv F(\varepsilon_x Fx) \quad (H_2)\end{cases}$$
$$(x)Fx \equiv F(\varepsilon_x \neg Fx) \quad (H_3)$$

So, what I said is the intuitive meaning of the ε-expression is expressed by this axiom {H_1}. And it is the only axiom for the ε-operator. Now, he defines explicitly the existential quantifier and the universal quantifier {definitions H_2 and H_3}, and

[33] As Fraenkel and Bar-Hillel note, if in a theory which is based on an ε-calculus, whose specific axioms do not contain ε-terms, a formula is derivable which does not contain ε-terms either, then there exists a proof of this formula in which no ε-terms occur. This eliminability of ε-terms assures the consistency of the ε-calculus relative to the consistency of the same theory based on first-order logic (Fraenkel and Bar-Hillel, 1958, pp. 184–185).

then on the basis of his one axiom $\{H_1\}$, he can show the theorems of all the first-order logic for the quantifiers. So it is actually a very elegant and effective basis of it. It is also very strong and that may give rise to some doubts about the legitimacy of its use—I will come to that later.

Now, I will characterise the Form III of the system.

Form III, in L_O': A_O, A_T', P
A_T': A_T^0 & A_T^1 & . . . & A_T^n.
A_T^0: $t = \varepsilon_u TC(u,o)$.
$(i = 1, \ldots, n)$ A_T^i: $t_1 = m_i(t)$.

It contains A_O and P in the old form, but replaces A_T by A_T', {where} A_T' consists just of explicit definitions, first of t, then of t_1, and so on. The explicit definition of t is very simple. I call it A_T^0 and it is the following: t equals $\varepsilon_u TC(u,o)$. In other words, t is the ε-object of TC. The ε-operator was sometimes also called by Hilbert a selection operator, because it selects an arbitrary element of the class. This definition was suggested to me by the meaning of the A_T, which told us: if there is anything at all that stands to o in the relation TC, then t should have this relation. Therefore, I said: then t is just the selection object, that is, the object which we can name by the ε-operator applied to TC. And so it suggested {to me} this theorem. Of course, it suggested it only; it did not prove that the theorem comes to the intended result. But that can easily be shown—I will not go into detail, but merely mention what can be shown on its basis.

First, having the t, we can easily define any t_i (i runs from 1 to n) as m_i of t, where m_i is a functor, meaning the i-th member in the n-tuple, which can very simply be expressed by the customary ι-operator and existential quantifiers—this I will not show here.[34] I will rather show the following.

From (H_2): $(\exists u)TC(u,o) \equiv TC((\varepsilon_u TC(u,o)),o)$.
with A_T^0: $\supset TC(t,o)$.
This is A_T.

Let's start with (H_2), which is Hilbert's definition of the existential quantifiers. From that we see that on the right-hand side the ε-expression $\varepsilon_u TC(u,o)$ occurs, which is just the definiens in the definition I {have} just proposed. So according to that definition, we can now replace it {the ε-expression} by t. If we do so, we have there a bi-conditional, and if we change that to a simple conditional, we have: if there is a u TC(u,o), then TC(t,o), which is our old postulate A_T in the Form II. So you see here that from our new postulate A_T', which contains these m + 1 explicit definitions,

[34]Given the definition of A_T^0: $t = \varepsilon_u TC(u,o)$, we can then define each theoretical term t_i $(i = 1, \ldots, n)$ as the i-th member of the n-tuple, using the schema A_T^i: $t_i = \varepsilon_x[\exists u_1 \exists u_2 \ldots \exists u_n (t = \langle u_1, \ldots, u_n\rangle \,\&\, x = u_i)]$. As Carnap notes, $t_i = \varepsilon_x[\exists u_1 \exists u_2 \ldots \exists u_n (t = \langle u_1, \ldots, u_n\rangle \,\&\, x = u_i)]$ admits of the logically equivalent form $t_i = (\iota x)[\exists u_1 \exists u_2 \ldots \exists u_n (t = \langle u_1, \ldots, u_n\rangle \,\&\, x = u_i)]$, since the formula $[\exists u_1 \exists u_2 \ldots \exists u_n (t = \langle u_1, \ldots, u_n\rangle \,\&\, x = u_i)]$ fulfils the uniqueness condition with respect to x (Carnap, 1961, p. 161).

we can logically derive the old A_T.[35] {We} knew already that from P and A_T we can derive TC. Therefore from the new A_T' together with P, which remains unchanged, we can derive TC. That is, if we keep the whole language, then in Form III we can derive TC from this system here. So this system really fulfils everything that we have: it is logically equivalent to the old theory—not only O-equivalent as Ramsey's form was, but logically equivalent to the whole theory. So it fulfils its purpose, it seems to me, and still we are in the observation language.

There is only one question left. You might say: what I said does not quite fit to each other. First I said, there is only an incomplete interpretation for the theoretical terms: nobody can specify their interpretation by merely referring to the observational terms. And then later, you may say, I claim to give explicit definitions of them. Now, obviously, that is incompatible. If we have no complete interpretation, then we have no possibility of an explicit definition. Well, you are right, or you would be right, if we were allowed only the observational terms and customary logical constants. But the Hilbert ε-operator belongs to a small class—there are a few other examples—of logical constants of a very peculiar kind. I will call them indeterminate. They are such that their meaning is not completely specified. {Here is an example.} {Let's} write down the ε-operator ε_n, {where} n {is} a variable for natural numbers, for the expression: n = 1 or n = 2 or n = 3. Then this total ε-expression is an element of the class consisting of the elements 1, 2, or 3. Let's abbreviate it by *a* {where} either *a* is 1, or *a* is 2, or *a* is 3. But if I write *a* is 1, or, it is not the case that *a* is identical to 1, there is no way of finding out the truth of this. Not because of lack of factual knowledge—it is not a factual sentence: there are only logical constants in it. {Nor} because of lack of logical information in the sense that I do not see quickly enough the logical consequences. Its meaning, the meaning of *a* has been specified by the ε-operator only up to a certain point: it is not any of those numbers which are outside of the class consisting of l, 2, 3. More is not said, just that is said. So we cannot determine whether *a* is 1 or is not. Now you see, this indeterminacy is just the one which we need for the theoretical concepts, if we use this explicit definition, which I used in my definition of t, because I defined t as that selection object which has the relation TC to o. Not meaning 'that', really any one, if there are such objects: { . . . } any one of them may be taken as denoted by this ε-expression. And this is exactly what we want to say by the meaning postulate. So this definition {gives} just so much specification as we can give, and not more. We do not want to give more, because the meaning should be left unspecified in some respect, because otherwise the physicist could not—as he wants to—add tomorrow more and more postulates, and even more and more correspondence postulates, and thereby make the meaning of the same term more specific than {it is} today. So, it seems to me that the ε-operator

[35]From Hilbert's definition H_2, it follows that $\exists u TC(u,o) \supset TC(\varepsilon_u TC(u,o),o)$. But $t = \varepsilon_u TC(u,o)$. Hence, $\exists u TC(u,o) \supset TC(t,o)$, which is A_T. So, A_T' entails the A_T-postulate of Form II, i.e., $^R TC \supset TC$.

is just exactly the tailor-made tool that we needed, in order to give an explicit definition, that, in spite of being explicit, does not determine the meaning completely, but just to that extent that it is needed. So that will conclude my talk.

Acknowledgements—The publication of the Santa Barbara paper was made possible thanks to the permission of the Carnap Archive in the University of Pittsburgh. All archival material from the Carnap and Feigl Archives is quoted with the permission of the University of Pittsburgh. All rights reserved. I should also thank Ms Brigitte Ulhemann, Curator of the Carnap Archive in the University of Konstanz, for her generous help, and the British Academy for financial support. Carnap's paper was discussed in the Graduate Seminar on the 'Philosophy of Rudolf Carnap' which I conducted with my colleague Thomas Uebel at the London School of Economics during the summer term of 1997–1998. Many thanks to Thomas and the graduate students who participated in the seminar for their perceptive comments. Many thanks are also due to Bill Demopoulos for important stylistic suggestions.

References

Braithwaite, R. B. (1953) *Scientific Explanation* (Cambridge: Cambridge University Press).

Carnap, R. (1937) *The Logical Syntax of Language* (London: RKP).

Carnap, R. (1952) 'Meaning Postulates', *Philosophical Studies* **3**, 65–73; reprinted in R. Carnap, *Meaning and Necessity: A Study in Semantic and Modal Logic* (Chicago: University of Chicago Press, 1956).

Carnap, R. (1956) 'The Methodological Character of Theoretical Concepts', in H. Feigl and M. Scriven (eds), *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, Minnesota Studies in the Philosophy of Science 1 (Minneapolis: University of Minnesota Press), pp. 38–76.

Carnap, R. (1958) 'Beobachtungssprache und Theoretische Sprache', *Dialectica* **12,** 236–248.

Carnap, R. (1961) 'On the Use of Hilbert's ε-operator in Scientific Theories', in Y. Bar-Hillel *et al*. (eds), *Essays on the Foundations of Mathematics* (Jerusalem: The Magnes Press), pp. 156–164.

Carnap, R. (1963) 'Replies and Systematic Expositions', in P. Schilpp (ed.), *The Philosophy of Rudolf Carnap* (La Salle, IL: Open Court), pp. 859–1013.

Carnap, R. (1974) *An Introduction to the Philosophy of Science* (New York: Basic Books).

Carnap, R. (1975) 'Observation Language and Theoretical Language', in J. Hintikka (ed.), *Rudolf Carnap, Logical Empiricist* (Dordrecht: Reidel).

Carnap, R. ([1950] 1990) 'Quine on Analyticity', in R. Creath (ed.), *Dear Carnap, Dear Van: The Quine–Carnap Correspondence and Related Work* (Berkeley and Los Angeles: University of California Press).

Fraenkel, A. A. and Bar-Hillel, Y. (1958) *Foundations of Set Theory* (Amsterdam: North-Holland P.C.).

Hempel, C. (1954) 'A Logical Appraisal of Operationism', *Scientific Monthly* **79**, 215–220; reprinted in C. Hempel, *Aspects of Scientific Explanation* (New York: Free Press, 1965).

Hempel, C. (1958) 'The Theoretician's Dilemma: A Study in the Logic of Theory Construction', Minnesota Studies in the Philosophy of Science 2 (Minneapolis: University of Minnesota Press).

Hempel, C. (1963) 'Implications of Carnap's Work for the Philosophy of Science', in P. Schilpp (ed.), *The Philosophy of Rudolf Carnap* (La Salle, IL: Open Court), pp. 685–709.

Hilbert, D. and Bernays, P. ([1939]1970) *Grundlagen der Mathematik II*, 2nd ed. (Berlin: Springer-Verlag).

Lewis, D. (1970) 'How to Define Theoretical Terms', *Journal of Philosophy* **67,** 427–446.

Papineau, D. (1996) 'Theory-dependent concepts', *Philosophy of Science* **63,** 1–20.

Quine, W. V. (1985) 'Carnap's Positivistic Travail', *Fundamenta Scientiae* **5,** 325–333.

Ramsey, F. ([1929]1978) 'Theories', in D. H. Mellor (ed.), *Foundations: Essays in Philosophy, Logic, Mathematics and Economics* (London: RKP), pp. 101–125.