only if $S$ is also deducible from $M$ when $M$ replaces $T$ as the accepted (normal syntactic) theory.

Suppose, now, that $T$ does, in fact, have a meaning abstract $M$. In what way, if any, is the meaning given to $M$ through acceptance of $T$ different from the meaning that $M$ would receive if it were to replace $T$ as the accepted theory? It is easy to see that the only difference between the syntactical accomplishments of $T$ and $M$ is the omission, when $M$ is the accepted theory, of whatever cognitively meaningless expressions are introduced by acceptance of $T$. All theoretical terms, consequences, and transformation pairs which are given meaning by acceptance of $T$ are preserved when $T$ is replaced by $M$. To our admittedly primitive understanding of these matters, it seems strange to suspect that the meaning of a portion of an accepted theory would be altered by deletion of another portion which is neither cognitively meaningful nor necessary for the remainder to maintain the same syntactical interconnectedness as before. In fact, the Thesis of Semantic Empiricism can be construed to rule out this possibility in that if deleting part of an accepted theory does not change the relation between the remainder of the theory and the observation language, it should not change the meaning of the remainder, either. Hence there is ample reason to assume that

Postulate 3d. If $M$ is a meaning abstract of normal syntactic theory $T$, the cognitive meaning, and hence factual content, of $M$ when $T$ is accepted are the same as the cognitive meaning and factual content, respectively, of $M$ when $M$ replaces $T$ as the accepted (normal syntactic) theory.

Lemma 2. Every normal syntactic theory has a meaning abstract.

Proof: Let ‘$T(\tau_1, \ldots, \tau_m)$’ be a normal syntactic theory in which $m$ ($0 \leq m \leq n$) of the theoretical terms are cognitively meaningful when $T$ is accepted, while the remainder are not. Then by P 3c, the sentencelike formulas deducible from $T$ which are cognitively meaningful when $T$ is accepted are exactly those consequences of $T$ which contain no theoretical terms other than the $m$ meaningful ones. Call these formulas the $M$-class of $T$. Now let $R^m_T$ be the sentencelike formula deduced from $T$ by existential quantification over the theoretical terms which are left meaningless when $T$ is accepted—i.e., when the meaningful terms are ‘$\tau_1, \ldots, \tau_m$’,

$R^m_T = \text{det} \left( \exists \phi_{m+1} \ldots \phi_n \right) T(\tau_1, \ldots, \tau_m, \phi_{m+1}, \ldots, \phi_n)$. Then $R^m_T$ belongs to $T$’s $M$-class, and it is simple to show, by Lemma 1, that any formula which belongs to the $M$-class of $T$ is also deducible from $R^m_T$

THE FACTUAL CONTENT OF THEORETICAL CONCEPTS

when $T$ is replaced by $R^m_T$ as the accepted theory. Hence $R^m_T$ is a meaning abstract of $T$. Q.E.D.

In particular, if all of the theoretical terms in $T$ are cognitively meaningful when $T$ is accepted, $T$ is its own meaning abstract; while if none are, the meaning abstract of $T$ is its prime consequence, $R_T$.

While the concept of "meaning abstract" bears sufficient technical interest to warrant discussion for its own sake, its present use is solely to provide a proof for Lemma 3. Consequently, the reader who feels uncomfortable with P 3c or P 3d may instead treat Lemma 3 as a postulated generalization of P 3b to replace P 3a–d.

Lemma 3. If an accepted normal syntactic theory $T$ does not itself signify a fact, then $T$ has a refuter only if its prime consequence has a refuter.

Proof: If $T$ is itself cognitively meaningful, the lemma follows immediately from P 3b. Conversely, suppose that $T$ is not cognitively meaningful. Then by P 3a, $T$ has a refuter only if it transforms a true cognitively meaningful sentence $S_1$ into another, $S_2$. Since $T$ transforms $S_1$ into $S_2$ if and only if $S_1 \supset S_2$ is a consequence of $T$, and $S_1$ and $S_2$ are true and false, respectively, if and only if $S_1 \supset S_2$ is false, $T$ then has a refuter only if it has a false cognitively meaningful consequence. Now, every cognitively meaningful consequence of accepted theory $T$ is also deducible from its meaning abstract, $R^m_T$; so if $T$ has a false consequence, $R^m_T$ must also be false when $T$ is accepted. Hence $T$, when accepted, has a refuter only if its meaning abstract also has a refuter. But by P 3d, $R^m_T$ has a refuter when $T$ is accepted only if it has a refuter when $R^m_T$ replaces $T$ as the accepted theory. But then, by P 3b, the prime consequence of $R^m_T$ also has a refuter; and since the prime consequence of $R^m_T$ is identical with the Ramsey sentence, $R_T$, of $T$, the prime consequence of $T$ likewise has a refuter. Q.E.D.

Postulate 4. Any refuter of an O-consequence of an accepted theory $T$ is also a refuter of $T$.

This simply brings forward in official form the obvious fact about theory usage noted earlier. By constraining an observation sentence to be a theory with no theoretical terms, P 4 also subsumes the point noted earlier (p. 298) that an observation sentence is likewise refuted by any refuter of its O-consequences.

This step is not quite automatic if the possibility raised in fn. 14 is taken seriously. However, the fact that $R^m_T$ contains every theoretical term in any meaningful consequence of $T$ obviates the difficulty.
William W. Rozeboom

Postulate 5. An accepted theory has a verifier if and only if it has no refuter.

This postulate merely formalizes for theories what was observed earlier, namely, that the terms ‘correct’ and ‘incorrect,’ unless generically inapplicable, are mutually exclusive and exhaustive. P 5 asserts that the behavioral role of an accepted theory is such that it has factual content. To say that an accepted theory would be incorrect if such-and-such were the case makes sense only if it is also the case that the theory would be correct if no condition were to obtain under which it is incorrect. Hence any fact which, together with any needed information such as that contained in SP I–IV and P 1–5 about the behavioral role of T, authorizes the conclusion that T has no refuter, then also authorizes, by P 5, the conclusion that T has a verifier and is hence itself a verifier of T.21

We have now extracted the factual content of an accepted normal syntactic theory, and it only remains to put the results in polished form.

Theorem 3. An accepted normal syntactic theory has the same factual content as its prime consequence.

Proof: Let \( R_T \) be a prime consequence of accepted normal syntactic theory \( T \). Then we have to show that the verifiers and refuters of \( T \) are identical with the verifiers and refuters, respectively, of \( R_T \). Since \( R_T \) is an O-consequence of \( T \), any refuter of \( R_T \) is also, by P 4, a refuter of \( T \), and hence by P 5, any verifier of \( T \) must also be a verifier of \( R_T \). Therefore, to complete the proof, it suffices to show that under SP I–IV and P 1–5, \( T \) can have no refuter when \( R_T \) is true. For then, given a fact \( f \) that verifies \( R_T \), it follows by P 5 that \( T \) has a verifier. Since this reveals that \( f \) and the facts about the behavioral role of \( T \) entail that \( T \) is correct, a verifier of \( R_T \) is also a verifier of \( T \)—which also shows, by P 5, that a refuter of \( T \) must also be a refuter of \( R_T \).

Suppose, now, that \( T \) signifies a fact. Then by P 1, \( T \) has a verifier and hence, by P 5, no refuter. On the other hand, suppose that \( T \) does not signify a fact. Then if \( T \) has a refuter, it follows by Lemma 3 that \( R_T \) must also have a refuter, a situation incompatible with \( R_T \)'s being true. Consequently, whether an accepted normal syntactic theory \( T \) itself signifies a fact or not, our premises (SP I–IV and P 1–5) about the behavioral role of theories show that \( T \) can have no refuter when its prime consequence is true. But as already pointed out, this suffices, by P 5, for a theory and its prime consequence to have the same factual content. Q.E.D.

One of the particularly important issues which demands attention by any serious methodological analysis of scientific theories is that of rivalry, conflict, or opposition among variously proposed theories. That is, what are the circumstances under which two alternative theories are incompatible? For we know from extended historical experience that many of the quarrels which arise in science and philosophy spring more from verbal misunderstandings and discordant interests than from genuine cognitive disagreement. When scientist A insists that his theory challenges the theory of scientist B, it would be highly useful to have means of determining in precisely what way, if at all, this is an actual clash of factual commitments and not just of personalities, especially if it is difficult to discern any testable differences between these theories. A major virtue of Theorem 3 is the illumination it brings to this question.

According to any intuitive understanding of the notion of “incompatibility,” Theorem 3 implies that two theories are incompatible if and only if they have incompatible observational consequences. For if a theory is factually equivalent to its prime consequence, then two theories are incompatible if and only if their prime consequences are incompatible. To demonstrate this clearly, however, calls for a definition of ‘incompatibility’ as applied to theories without necessarily assuming that theories are themselves cognitively meaningful, and this is not quite so simple as it might at first appear.

What do we mean by saying that two theories are incompatible? It will be helpful first to examine the concept as it applies to observation sentences, and then seek a suitable extension to theories. A condition of “incompatibility” which immediately comes to mind is that \( S_1 \) and \( S_2 \) are incompatible when a contradiction can be deduced from them jointly. This is not a necessary condition, however, for “incompatibility” is more than just a syntactical relation, and \( S_1 \) and \( S_2 \) may be incompatible even though they formally entail no contradiction. For example, \( F(a) \) and \( \sim F(b) \) are formally consistent, but are incompatible if ‘a’ and ‘b’ are synonymous. Thus a more satisfactory explication might be that two sentences are in-
both have verifiers in the roles allocated to them. That is, if the (potential) factual commitments of $E_1$ are incompatible with those of $E_2$, then any fact which verifies $E_1$ when the latter is used in its prescribed manner also refutes $E_2$ when the latter is used in its prescribed manner. This is not a sufficient condition for the incompatibility of $E_1$ and $E_2$, however, for if $E_1$ and $E_2$ are (or would be) both incorrect, it is vacuously true that all (potential) verifiers of $E_1$ (potentially) refute $E_2$ and all (potential) verifiers of $E_2$ (potentially) refute $E_1$. For a definition of "incompatibility" in terms of factual content, then, we need to identify something about the refuters of two incorrect expressions which reveals that they could not both be correct. The following is offered as one possibility which may or may not require modification as the ontology of "facts" becomes more clearly understood. For purposes of Theorem 4, however, any alternative definions would do, so long as it is wholly a condition on the verifiers and refuters of the expressions involved.

Definition 8. Expressions $E_1$ and $E_2$ are incompatible (relative to usage $U$) $\iff E_1$ (under $U$) has a verifier and every verifier of $E_1$ (under $U$) refutes $E_2$ (under $U$); or $E_2$ (under $U$) has a verifier and every verifier of $E_2$ (under $U$) refutes $E_1$ (under $U$); or there exists a tautological fact which is the disjunction of a refuter of $E_1$ (under $U$) and a refuter of $E_2$ (under $U$).

By the verifiers or refuters of $E_i$ "under $U$" is meant, of course, the verifiers or refuters that $E_i$ would have if used in accordance with procedure $U$. The definition does not presuppose that $U$ necessarily allows $E_1$ and $E_2$ to be used jointly.

To appreciate that the final clause in Definition 8 preserves the notion that if $E_1$ and $E_2$ are incompatible in their stipulated roles, they must not be, rather than merely are not, both incorrect, suppose that $f$ is a refuter of $E_1$ (under $U$) and $g$ is a refuter of $E_2$ (under $U$). Now, a tautological fact is such that we would say that it must be the case. Hence, if $f \lor g$ is tautologous, it must be the case that either $f$ or $g$; thus either $E_1$ (under $U$) or $E_2$ (under $U$) must have a refuter.

Theorem 4. (a) Theories $T_1$ and $T_2$ are incompatible if and only if their prime consequences are incompatible. (b) Theory $T$ and observation sentence $S$ are incompatible if and only if $S$ is incompatible with the prime consequence of $T$.

Proof: Let $S_1$, $S_2$, and $S_3$ be observation sentences or theories. Since application of Definition 8 depends only upon the verifiers and refuters concerned, if $S_1$ has the same verifiers and refuters, respectively, as $S_2$, then
S1 and S3 are incompatible if and only if S2 and S3 are also incompatible. But by Theorem 3, the verifiers and refuters of a theory (when accepted for normal syntactic usage) are identical with those of its prime consequence. Hence two theories are incompatible if and only if their prime consequences are incompatible, and similarly for a theory and an observation sentence. Q.E.D.

A theorem of great philosophical importance follows immediately from Theorem 3. If two theories have identical O-consequences, their prime consequences must be formally equivalent. Hence from Theorem 3,

**Theorem 5. If two theories have identical observational consequences, they have the same factual content.**

Theorems 4 and 5 dispel a number of perplexities that have traditionally been associated with the epistemic status of scientific theories. Much philosophical Angst and operationistic impatience have been vented over the prima-facie possibility that two conflicting theories might have no observational disagreement, or that nonequivalent or even incompatible theories might have identical observational consequences and be equally supported (or disconfirmed) by any empirical evidence. The present analysis suggests that the famous pragmatic dictum “A difference which makes no difference is no difference,” should be put even more strongly as a logical contention: “There is no difference which makes no difference.” Theorems 4 and 5 show that given the Thesis of Semantic Empiricism, it is not possible for two theories to be incompatible without being observationally incompatible as well, or to be nonequivalent without differing in their observational consequences. (Note that I am not saying that two theories with the same O-consequences are necessarily synonymous, but only that any verifier or refuter of one is also a verifier or refuter of the other.) The fallacy has been to presume that the same theoretical symbol necessarily has the same meaning in one theory as it has in another, overlooking that it is the particular theoretical usage which gives the symbol its meaning.

**Technical Note:** It will be recalled that the theorems which have been developed in this section rest upon certain assumptions in addition to those made explicit in the theorems proper. All but one of these, which were set forth in Section I, concern the character of the observation language. However, we also made one stipulation that might seem to be a gratuitous restriction on the form of a normal syntactic theory; namely, that theoretical terms enter the theoretical postulates syntactically only as constants. It would appear that at least some of the variables of a language may be descriptive (i.e., nonlogical) terms in the sense that the classes over which they range are nonlogical categories. One might then wonder whether a theoretical concept might not find expression as the range of a variable. Syntactically, this would be accomplished by introducing variables and perhaps constants of a new, uninterpreted formal type. If it is possible to introduce theoretical terms in this way—and indeed, so long as variables need not be purely logical, it seems unreasonable to deny that they can—we might question whether the present theorems apply to such theories.

A satisfactory discussion of formal types and the ranges of variables is too lengthy to be undertaken here. There are, however, compelling reasons for believing that a language which contains nonlogical variables and is also adequate to formulate and cope with the various problems which arise from the use of these variables must be such that for any nonlogical variable ‘x’ in the language, it must contain or permit introduction of (a) a descriptive constant which refers to the class (or the defining property thereof) ranged by ‘x’, and (b) a purely logical variable, ‘x’, whose range includes that of ‘x’.

**Thus, if formal type i represents a nonlogical category, an empirical problem immediately arises as to whether or not a given entity t can be designated by a constant of type i. But if the language contains the predicate ‘x’ can be designated by a symbol of type i,’ this is equivalent in force to ‘x’ is a member of C,’ where ‘C’ designates the class corresponding to formal type i. Moreover, if ‘x’ is also a nonlogical variable, we would have to determine that t belongs to the range of ‘x’ before using this predicate to inquire whether or not t belongs to the range of ‘x’; so if the language is to be able to formulate the question about an entity’s membership in the range of a nonlogical variable, it must contain a logical variable to terminate the regress.**
In this section, we shall investigate the semantical status of theoretical terms and postulates. Before proceeding further, however, let us introduce a simplification in notation. Unless the analysis turns on the number of theoretical terms, there is no need to maintain explicit reference to \( n \) theoretical constants. Hence, with the exception of a few places needing careful formulation, we may restrict discussion to theories with only one theoretical term.

In preceding sections, the question repeatedly arose whether a (normal syntactic) theory is merely an instrument for generating its \( O \)-consequences, or whether \( T \) is itself cognitively meaningful. My first contention in this section is that the latter is indeed the case. For if we deny that a normal syntactic theory is, in some fashion, itself an assertion, we find ourselves committed to the view that no expression containing a descriptive term which does not designate a sense datum can be an assertion. If we take ‘observed’ not in the phenomenalistic sense of ‘directly experienced,’ but in the broader usage of science and everyday life, there is no hard and fast distinction between observational (i.e., “empirical”) and theoretical concepts. The cytologist, for example, considers cells and their grosser properties as “observable,” even though the observation depends upon an intervening distortion of light rays by the lens of a microscope. More generally, it has long been accepted that our access to the events in the objective world to which our observational terms are commonsensically presumed to refer is only through the medium of causal chains which bridge between these events and our nervous systems. Most of our “observational” concepts, upon philosophical scrutiny, may be seen to lose their halo of immediacy and to stand in very much the same relation to more immediately given events as theoretical concepts stand to events in the commonplace world. Now, I am by no means convinced that the phenomenally “given” is as mythological as some would have it, nor do I deny that some terms of ordinary language appear to designate phenomenal entities. However, I think it would be absurd to maintain that sentences in the everyday observation language, except for a proper phenomenal subset, are nothing but instruments for committing one to a set of statements wholly in a phenomenal language. This would be plausible only if we did, in fact, habitually use ordinary language for this purpose. The fact is, of course, that purely phenomenal statements, at least those recognized as

such, play a minor if not virtually nonexistent role in linguistic practice. If only sentences known to be wholly phenomenal were able to make assertions, virtually the whole of our linguistic machinery would lack cognitive meaning. I would thus maintain that ordinary observation sentences, whether recognizably phenomenal or not, do in general have semantic properties, and that by the same token, such properties are also possessed by theoretical postulates.

However, in arguing from the negligible incidence of recognizably phenomenal statements in ordinary discourse to the conclusion that theoretical postulates and observation sentences in the broad sense must, in general, themselves be assertions, a possibility which must not be overlooked is that while a symbol can designate only a previously experienced sense datum, we may construct, by intricate and presumably to a large extent unwitting definitional processes, sentences which are wholly about phenomenal facts but which are not readily identified to be so. This, of course, is the standard phenomenalistic move—not to deny that everyday sentences are assertions, but to claim that upon analysis, they can be discovered to be wholly phenomenal. In like manner, the positivist, however he construes the “observed,” can argue that while a theory is indeed an assertion, it is analyzable into a sentence of the observation language. Thus granting that theories may be assertions, we must consider whether there might not be some sentence ‘\( S \)’ constructable in \( L_o \) for which it can be argued that \( T(\tau) \equiv_{\text{def}} S \)—i.e., that ‘\( T(\tau) \),’ in virtue of its definition, has the same meaning as ‘\( S \).’

Let us first dispose of the possibility that the way in which ‘\( T(\tau) \)’ signifies a fact is by being a peculiar, syntactically improper, notational form for some more orthodox sentence in the observation language, similar to the way, for example, that a code signal may be said to signify a fact because, while the code signal is not syntactically a sentence, it has been stipulated to abbreviate an ordinary sentence. What we are now asking is, if ‘\( T(\tau) \)’ not merely carries factual content but also signifies a fact, whether it could do this in any way other than by ‘\( \tau \)’ standing in a relation of reference to some entity of appropriate type. It would be possible, for example, to stipulate that ‘\( T(\tau) \)’ is to mean that \( p \), where ‘\( p \)’ is some observation sentence of syntactical form different from ‘\( T(\tau) \).’ Actually, the Thesis of Semantic Empiricism and SP I confute this; however, since we now wish to consider possible alternatives to the Thesis, we must look further into the possibility that the meaning of ‘\( T(\tau) \)’ is not governed by the syntax of
William W. Rozeboom

In particular, since natural theory usage seems to show that we normally regard a theory to have the same observational force as its prime consequence, we are especially interested in the possibility that $T(\tau) = \text{det} (\exists \phi) T(\phi)$.

That this is untenable, however, may be shown in at least two ways:

1. If ‘$T(\tau)$’ is an assertion, then presumably the reason that acceptance of $T$ commits one to the $O$-consequences of $T$ is because the latter are logically entailed by ‘$T(\tau)$.’ But if ‘$T(\tau)$’ asserts $(\exists \phi)T(\phi)$, then applying formal inference rules to the sentence form ‘$T(\tau)$’ is not logical deduction. That logical inference is more than merely a set of operations within an arbitrary formal calculus is due to certain relations which obtain among the cognitive meanings of statements as a result of their logical forms, as normally mirrored by their syntactic forms. Hence it is incorrect to stipulate that a syntactically sentencelike formula $S$ is to assert the same fact as another, syntactically different, statement, and then to claim that the syntactic consequences of $S$ must be its logical consequences. To be sure, if it were true that $T(\tau) = \text{det} (\exists \phi) T(\phi)$, we could vindicate taking a sentence in $L_0$, syntactically deductible from ‘$T(\tau)$,’ to be a logical consequence of ‘$(\exists \phi)T(\phi)$’ and hence of ‘$T(\tau)$,’ because as it happens, the two formulas have the same $O$-consequences. But the existence of Ramsey sentences has been known only since 1929, and even since then has been virtually ignored. Hence, if it were the case that $T(\tau) = \text{det} (\exists \phi) T(\phi)$, users of theories would heretofore have been unjustified in taking acceptance of $T$ as necessary commitment to the $O$-consequences of $T$.

2. It seems unacceptable to grant semantic status to ‘$T(\tau)$’ and withhold it from its theoretical consequences, especially since in the de facto use of theories, the full conjunction of theoretical postulates, ‘$T(\tau)$,’ is seldom, if ever, actually constructed. For example, suppose that $T$ is the conjunction of two theoretical postulates, ‘$S_1(\tau)$’ and ‘$S_2(\tau)$.’ If ‘$S_1(\tau) \cdot S_2(\tau)$’ is a statement, we should certainly wish also to say that ‘$S_1(\tau)$’ and ‘$S_2(\tau)$’ are statements. But if $S_1(\tau) \cdot S_2(\tau) = \text{det} (\exists \phi)[S_1(\phi) \cdot S_2(\phi)]$, it is not the case that $S_1(\tau) = (\exists \phi)S_1(\phi)$ and $S_2(\tau) = (\exists \phi)S_2(\phi)$ except in degenerate instances. Without entering into formal details, let me simply assert that if $T(\tau) = \text{det} (\exists \phi) T(\phi)$, there appears to be no satisfactory translation for the constituent postulates in $T$ or their theoretical consequences. We may raise analogous objections against any other sentence $S$ which might be proposed as a definiens for ‘$T(\tau)$’ unless $S$ has the same syntactical structure as ‘$T(\tau)$.’ Since in this case, $S$ must be of the form ‘$T(d)$,’ where ‘$d$’ is a referring expression in $L_0$, setting $T(\tau) = \text{det} T(d)$ is the same as setting $\tau = \text{det} d$. We may therefore conclude that if ‘$T(\tau)$’ signifies a fact, it does so because ‘$\tau$’ functions referentially in a syntactically proper context.

Our next problem is to determine the extent, if any, to which theoretical terms extend the referential capabilities of language. In particular, we must now consider whether ‘$\tau$’ might not be equivalent to some descriptive expression lying wholly within $L_0$. (By saying that one descriptive expression is “equivalent” to another, I mean that the one may be replaced in any sentence by the other without change in the factual content of the sentence.) For a positivist would in no way be dismayed by the fact that ‘$T(\tau)$’ is a cognitively meaningful statement in which ‘$\tau$’ functions referentially so long as he were allowed to argue that there existed an expression ‘$d$’ in $L_0$ with the same designative force as ‘$\tau$’. He would then contend that ‘$T(\tau)$’ is equivalent to the observation sentence ‘$T(d)$,’ and similarly, that any other expression ‘$E(\tau)$’ is equivalent to ‘$E(d)$’.

What conditions must obtain if ‘$\tau$’ may be regarded as equivalent to some observational expression ‘$d$’? (Note that we cannot, for the moment, draw upon the Thesis of Semantic Empiricism for assistance, for the positivistic contention is an alternative to the Thesis.) An obviously necessary condition is that ‘$T(\tau)$’ and ‘$T(d)$’ have the same factual content. But what, according to the positivist, is the factual content of a theory? If we are to proceed further, he must tell us his criterion for an observational fact not to be a refuter of ‘$T(\tau)$’. Only two alternatives seem available to him:

1. He might assert that the refuters of ‘$T(\tau)$’ are exactly the refuters of ‘$T(d)$,’ where ‘$d$’ is that observational expression to which he claims ‘$\tau$’ is equivalent. But to select ‘$d$’ for this purpose, rather than some other of the many expressions of the same formal type as ‘$\tau$’ available in $L_0$, is simply to take the alleged equivalence of ‘$\tau$’ and ‘$d$’ as a premise from which to analyze the force of ‘$T(\tau)$,’ which would be justified only if ‘$\tau$’ had been explicitly introduced to have the same meaning as ‘$d$’.

2. He may hold that ‘$T(\tau)$’ is refuted by an observational fact $f$ only if $f$ refutes an $O$-consequence of $T$, and hence by the corollary to Theorem 2, only if $f$ refutes ‘$(\exists \phi)T(\phi)$’. But then, if ‘$\tau$’ is equivalent to ‘$d$’, ‘$(\exists \phi)T(\phi)$’ must analytically entail ‘$T(d)$’ (and also, of course, conversely), since any
refuter of "T(d)"—which thus also refutes "T(τ)"—must then also refute "(∃ϕ)T(ϕ)." And indeed, whether the positivist wishes to adopt this position or not, it would seem that in the natural usage of theories, the fact that T(d) is not the case would not be taken as disproof of "T(τ)" unless the theory has an O-consequence which is refuted by \( \sim T(d) \); otherwise, by parity of reasoning, any observational fact \( \sim T(d) \) should disprove the theory. Hence, unless \( (∃ϕ)T(ϕ) \) entails that T(d), the force of T(d) is stronger than that of T(τ), and τ and d' cannot be equivalent. Consequently, a prerequisite of the positivist's thesis is that whenever a theory "T(τ)" is itself a cognitively meaningful sentence, there exists a descriptive expression d' in \( L_0 \) such that "(∃ϕ)T(ϕ)" and T(d') are analytically equivalent. Let us refer to the italicized condition as the Positivistic Criterion.

Now, there can be no denial that for some theories, the Positivistic Criterion is satisfied. For example, if "(ϕ)[r(ϕ) \supset P(ϕ)] \cdot [∃a(ϕ)]" is a theory in which 'P' and 'a' are observational constants, then "(∃ϕ)[(ϕ \supset P(ϕ)] \cdot [ϕ(a)]" is formally equivalent to "(ϕ)[P(ϕ) \supset P(a)] \cdot [P(a)]", and a positivist could contend that 'P' is equivalent to 'P'. However, it is obviously not the case that for any predicate 'F(ϕ)', a descriptive expression d' can be found such that T(d') has the same force as "(∃ϕ)F(ϕ)"—otherwise, logical quantifiers could be entirely eliminated from the language without attenuating its strength. Hence, if the positivistic thesis is to provide a general account of the meaning of theories, it must be the case either that (a) the only expressions, T(τ), which are ever legitimately regarded as theories are those which satisfy the Positivistic Criterion, or that (b) only when the Positivistic Criterion is satisfied can T(τ) itself be an assertion. But (a) is obviously false—not only do we fail to invoke the Positivistic Criterion when passing judgment upon theories, it is highly unlikely that it is met by any theory of current scientific importance. As for (b), this contention would also apply to the reduction of common-sense "observational" terms to purely phenomenal phrases, and would imply that ordinary-language observation sentences are dichotomized into those which are genuine—i.e., phenomenal—statements, and those which are merely mechanisms for passing from one phenomenal statement to another. But this is wholly implausible. Sentences which are recognized as being purely phenomenal play at best a minor role in actual linguistic usage, while it is just not true that observation sentences which are not recognizably phenomenal, we differentiate in use between "genuine" statements which we think could be given a phenomenal reduction, and "formal devices" which have a second-class linguistic status. It seems to me highly gratuitous to postulate a semantic distinction which corresponds neither to a difference in use nor to any feature in our normal conceptualization of language. I can only conclude that the Positivistic Criterion for the cognitive meaningfulness of theoretical expressions is untenable, and that in general we must be prepared to find that a theoretical term, though meaningful, need not be equivalent to any phrase in the observation language.

But how, then, are we to analyze the meanings of theoretical terms? The solution lies, I believe, in regarding such meanings not as something brought to the theory by the theoretical constants, but as something acquired by the theoretical terms in virtue of their participation in the theory. This, of course, is simply the Thesis of Semantic Empiricism which was invoked in proof of Theorem 3, except that we are now adding that theoretical terms do have cognitive meaning acquired in this way. I do not see how we could possibly hold otherwise if we wish both to maintain the empiricist tradition and yet to grant extra-observational reference to theoretical terms. Hence,

Postulate 6. The semantic properties imparted to a normal syntactic theory T by its acceptance are such that T is itself able to signify a fact.

From P 1, P 2 and P 6, it follows immediately that

Theorem 6. If "T(τ_1, ..., τ_n)" is an accepted normal syntactic theory in which "T(ϕ_1, ..., ϕ_n)" is an observational predicate and "τ_i, ..., τ_n" are theoretical terms, then: (a) "T(τ_1, ..., τ_n)" signifies a fact if and only if there exist entities t_1, ..., t_n such that t is the fact that T(t_1, ..., t_n).
(b) If t is an entity such that "(∃ϕ_1, ..., ϕ_1 = t_1, ϕ_i + 1, ..., ϕ_n)T(ϕ_1, ..., ϕ_i = t, ϕ_i + 1, ..., ϕ_n)", then "τ_i" designates t.

That is, reverting to the simplest case, accepted theory "T(τ)" signifies every fact of form T(ϕ), and 'τ' designates every entity which satisfies T(ϕ). It should be noticed, however, that Theorem 6 supplies sufficient but not necessary conditions for 'τ' to designate t. Neither do Ps 1–6 suffice to determine the factual significata of all theoretical consequences of T. We shall attempt to do something about this deficiency shortly.

The plausibility of the present interpretation of theoretical reference will be strengthened, perhaps, by the behavioral theory of designation to be sketched at the end of this section. For the present, let us consider the obvious objection which arises. I say 'the objection' because it seems to me that basically there is only one—the fact that according to the present
formulation, a theoretical expression may have more than one referent. For if there were at most one entity which satisfies ‘T(ϕ),’ we would regard ‘T(τ)’ as assigning a referent to ‘τ’ by means of description—i.e., we could assume that τ = def (ϕ)T(ϕ). And while the analysis of descriptions is far from agreed upon, it is not implausible that under suitable conditions, descriptions and expressions which contain them do, in some sense, designate. Hence the present view should appear at least reasonable, so long as it is possible to develop a workable semantical theory of multiple designation.

Since the notion that a theoretical term may have more than one referent is the key idea to emerge from the present analysis, it is very important to have a clear understanding of what is being contended. The relation that obtains between a predicate and an entity which satisfies the predicate is occasionally known as ‘denotation.’ Thus we might say that ‘human’ denotes Tom, James, Elmer, etc. Under this usage, a predicate may be said to have “multiple denotata,” and it is crucial that this be sharply distinguished from multiple designation. A primitive predicate designates, or refers to, an abstract entity which is exemplified (or belonged to, if the referent of a predicate is a class) by its denotata, and hence will in general have many denotata even though it has but one designatum. Since theoretical terms are syntactically primitive, they may be said to name the entities to which they refer. Then to say that a theoretical term may have multiple designata is to imply that a term may simultaneously name more than one entity, thus departing radically from classical semantics.

It is, moreover, most important to appreciate that this unorthodox suggestion which has emerged, namely, that theoretical expressions may designate without designating uniquely, is due neither to a personal perversity nor to some special, restrictive, arbitrary assumption during the earlier stages of the argument. Quite the contrary, it is an apparently inescapable joint consequence of two popular and highly plausible epistemological beliefs, namely, (a) that a theoretical sentence may genuinely signify a fact which cannot be signified by a sentence in the observation language; and (b) that the semantic properties of theoretical expressions are given to them by their use with the observation language. For even if we grant that observation terms have unique referents, there just does not seem to be any way for the observation language to provide a criterion which may admit an unobserved entity as a referent of a theoretical term and yet also guarantee uniqueness. It would seem, therefore, that a theory of multiple designation is an inescapable correlate to any coherent form of empirical realism, where by the latter we mean epistemological theories which affirm both that knowledge about unobserved entities is possible and that this knowledge is given only through what is observed. If this be so, however, it becomes binding upon philosophically responsible empirical realists to carry through a comprehensive analysis and at least partial reformulation of basic semantical principles, for it must frankly be admitted that a theory of multiple designation is not, prima facie, wholly compatible with classical semantics.

It is a cornerstone tenet of semantics that a statement has at most one truth value—i.e., that it is not both true and false. It is further customary to hold that if ‘S’ designates the property P, and ‘s’ designates entity t, then the sentence ‘S(s)’ is true if it is the case that P(t), and false if it is the case that ~P(t). But this would constitute a fatal objection to any semantical theory which allows a term to have more than one designatum. For suppose that ‘s’ designates both t₁ and t₂. If t₁ and t₂ are different entities, there must be some property P such that P(t₁) and ~P(t₂). But if ‘S’ is a predicate which designates P, it would then appear that ‘S(s)’ must be both true and false. Applied to the present contention that theoretical terms designate all entities which satisfy the observation predicate characterizing the theory, this objection charges that it admits truth-inconsistent statements in violation of the Principle of Contradiction. To be sure, so far as theory ‘T(τ)’ itself and any theoretical sentences derivable from it are concerned, no ambiguities in truth value arise, for it is a condition on the designata of ‘τ’ that they satisfy ‘T(ϕ)’; hence if ‘T(τ)’ formally entails ‘E(τ),’ the case cannot arise where ‘τ’ designates t and it is not the case that E(t) (since by an easily proved corollary to Lemma 1, ‘E(τ)’ is deducible from ‘T(τ)’ only if every satisfier of ‘T(ϕ)’ also satisfies ‘E(ϕ)’). However, if ‘F(τ)’ is a theoretical sentence not entailed by ‘T(τ),’ then if ‘τ’ has more than one designatum, say t₁ and t₂, it is entirely possible that F(t₁) while not F(t₂), which would seem to imply that ‘F(τ)’ may be both true and false.

Now, it should first of all be noted that the semantical assumption just

footnote 24 For an informal discussion of this point through common-sense examples, see [14].
William W. Rozeboom

employed, namely, that if ‘S’ designates P and ‘s’ designates t, then ‘S(s)’ is true if it is the case that P(t) and false if it is the case that ~P(t), is significantly stronger than SP I–IV, in which were formalized the semantical principles on which the present analysis is based. For SP I–IV leave open the possibility that even though ‘S’ designates P and ‘s’ designates t, ‘S(s)’ may not assert that P(t)—i.e., signify P(t) truly or ~P(t) falsely, depending on which is the case—and hence that a sentence may fail to signify a fact even though all its descriptive terms have designata. We now see that if both multiple designation and the Principle of Contradiction (SP IV) are to be maintained, this possibility must remain. However, this does not leave matters in a very satisfactory state, for if ‘S(s)’ does not necessarily truly signify the fact that P(t), or falsely signify the fact that ~P(t), even though ‘S’ designates P and ‘s’ designates t, what then are the conditions sufficient for a fact to be a verifier or refuter of a sentence in virtue of the semantic properties of the latter?

To understand the origins of the predicament in which our analysis now finds itself, and to sympathize with its departure from classical semantics, it is necessary to remain sensitive to a truistic but not always properly appreciated prerequisite for semantical relations to obtain. This is, simply, that it is not words and sentences qua sign-designs which stand in semantical relations to entities, but words and sentences in use—i.e., symbols which have come to play a suitable role in language behavior. It is customary and quite proper for “pure” semantics to axiomatize certain properties of semantical relations abstracted from the total linguistic situation, but it must not be forgotten that when semantical relations obtain between symbols and extralinguistic entities, it is because these symbols are being used in a certain way. While it is perfectly acceptable for a semanticist to lay down sentences of the form ‘s designates x’ as postulates for analysis without committing himself to the nature of this relation, the results of his analysis are not applicable to either de facto or idealized language practices unless the bare signs of the language are embedded in a pragmatic context by virtue of which a coordination is established between signs and their designata.

Now this point may seem trivial at first, but it ceases to do so when one reflects that the “use,” or “linguistic role,” of a sign-design is more clearly described as some aspect of the psychological state of a language user o at time t with respect to that sign-design, and that it is by no means necessarily the case that the psychological state of person o at time t with respect to a sentencelike sign-design S is such as to endow S (in its linguistic role for o at t) with all the semantic properties presupposed by classical semantics, even though S has an appreciable incidence in o’s linguistic behavior. On the other hand, merely because the psychological state of o at t with respect to certain expressions in his language does not fully qualify these expressions for analysis by classical semantics, it would be most rash to conclude that these expressions are not in any way cognitively meaningful or do not function referentially for o at t. For example, classical semantics has no place for vague concepts; yet it would be absurd to argue that because the “borderline fuzziness” of most if not all terms in actual use reveals them to be more or less vague, ordinary language is cognitively meaningless. Moreover, it would be hazardous to construe the discrepancies between actual languages and classical semantics as due wholly to noncognitive contaminations of a theoretically pure semantical state described by the classical postulates. It seems much more reasonable to suspect, or at the very least to entertain as a possibility, that the classical account is a limiting form of what is generally a more complex pattern of semantical relations, while the latter is just as much a pure cognitive system as its classical limit and may likewise (though more comprehensively) represent a theoretical ideal to which actual languages are but an approximation.

If one does admit the possibility that classical semantics may be only a special instance of more general semantical principles, however, then clearly we should expect that in order to analyze the cognitive function of theoretical expressions it will be necessary to develop a semantical theory adequate to the broader case. For as will shortly be examined in greater detail, concepts which qualify as “theoretical” are transient stages of a linguistic growth process and are hence incomplete in a way that the concepts presupposed by classical semantics are not. Consequently, if the present analysis of theoretical expressions is basically sound, P 1–6 provide a framework within which we may begin to explore the nonclassical dimensions of cognitive processes.

It has already been pointed out that P 1–6 do not fully delimit the designative properties of theoretical expressions. While any reasonably adequate development of a generalized theory of semantics, and discussion of its relation to the classical limit, is far beyond the present scope, let me at least offer a provisional set of hypotheses which seem to make a
certain amount of intuitive sense, and which, moreover, reconcile the possibility of multiple designation with the Principle of Contradiction.25

Definition 9. \( E(\tau_1, \ldots, \tau_m) \) is an autonomous subtheory of theory \( T \triangleq \text{det} E(\tau_1, \ldots, \tau_m) \) is a sentencelike formula deducible from \( T \) in which \( E(\phi_1, \ldots, \phi_m) \) is an observational predicate and \( \tau_1, \ldots, \tau_m \) are \( m \) of the \( n \) (\( 0 < m < n \)) theoretical terms contained in \( T \); and the cognitive meaning imparted to \( E(\tau_1, \ldots, \tau_m) \) by (normal syntactic) acceptance of \( T \) is the same as the meaning acquired by \( E(\tau_1, \ldots, \tau_m) \) when accepted as a (normal syntactic) theory by itself.

The purpose of this definition is to facilitate handling of the possibility suggested earlier that a totality, \( T \), of accepted theoretical postulates may contain subsets which function independently of the remainder. Whether there are, in fact, autonomous subtheories of \( T \) which are not formally equivalent to \( T \), and what the conditions must be for a subtheory to be autonomous, we shall not here attempt to explore. There is reason to believe that a theoretical consequence \( E \) of \( T \) is an autonomous subtheory of \( T \) if every consequence of \( T \) containing one or more of the theoretical terms in \( E \) is equivalent to a sentence of \( C \cdot S_E \), in which \( C \) contains no theoretical terms in \( E \) and \( S_E \) is deducible from \( E \) alone. However, this may not be a necessary condition for autonomy.

Definition 10. \( E \) is a unified subtheory of theory \( T \triangleq \text{det} E \) is a theoretical consequence of \( T \) and there is no autonomous subtheory, \( E_a \), of \( T \) such that \( E_a \) is deducible from \( E \) and \( E \) is not deducible from \( E_a \).

That is, a unified subtheory of \( T \) cannot be resolved into components whose meanings are acquired independently of the remainder. The units of meaning acquisition when \( T \) is accepted are then those theoretical consequences of \( T \) which are autonomous and unified.

Hypothesis A. If \( E(\tau_1, \ldots, \tau_m) \) is an autonomous and unified subtheory of an accepted normal syntactic theory \( T \) in which \( \tau_1, \ldots, \tau_m \) are theoretical terms and \( E(\phi_1, \ldots, \phi_m) \) is an observational predicate, then \( \tau_1 \) designates an entity \( t \) if and only if it is the case that \( \exists \phi_1, \ldots, \phi_1 - 1, \phi_1 + 1, \ldots, \phi_m \) \( E(\phi_1, \ldots, \phi_1 - 1, t, \phi_1 + 1, \ldots, \phi_m) \).

Since a set of entities \( t_1, \ldots, t_n \) will satisfy \( T(\phi_1, \ldots, \phi_m) \) only if its subset \( t_1, \ldots, t_m \) satisfies \( E(\phi_1, \ldots, \phi_m) \), an entity \( t \) will qualify as a designation of \( \tau_1 \) under Theorem 6 only if it also qualifies under

While these hypotheses concern only the acquisition of designa by theoretical terms through their use with observation language \( L_o \), a similar set of principles would be expected to govern the endowment of expressions in \( L_o \) with meanings derived from immediate experience in the event that \( L_o \) is not a phenomenal language (cf. fn. 32).

THE FACTUAL CONTENT OF THEORETICAL CONCEPTS

Hypothesis A. Hence this hypothesis extends Theorem 6 in such a way as to provide necessary as well as sufficient conditions for the designata of theoretical terms.

Hypothesis B. If \( E(\tau_1, \ldots, \tau_m) \) is a theoretical sentence in which \( E(\phi_1, \ldots, \phi_m) \) is an observational predicate and \( \tau_1, \ldots, \tau_m \) are theoretical terms introduced by an accepted normal syntactic theory \( T \), then \( E(\tau_1, \ldots, \tau_m) \) signifies a fact \( f \) if and only if there exist entities \( t_1, \ldots, t_m \) such that \( \tau_1, \ldots, \tau_m \) designate \( t_1, \ldots, t_m \), respectively, and \( f \) is the fact that \( E(t_1, \ldots, t_m) \).

From this and SP I it follows that a theoretical sentence can signify a fact only truly. Hypothesis B is an extension of P 1 to all theoretical sentences, deducible from \( T \) or not. Actually, it needs to be generalized to describe what a set of theoretical sentences simultaneously signify (see footnote 27), but this is a further development which may be forgone here.

Hypothesis C. If \( E \) is a theoretical sentence whose theoretical terms have been introduced by an accepted normal syntactic theory \( T \), \( E \) is true or false according to whether or not there exists a fact signified by \( E \).

That is, cognitive meaningfulness does not presuppose factual reference, and a sentence may be false precisely because there is nothing in external reality which conforms to the criteria built into the sentence's meaning.

It will be observed that Hypotheses A–C agree with classical semantics in the limiting case where (unified) theory \( T(\tau) \) is adequate to confer exactly one designatum, say \( t \), upon \( \tau \) (i.e., when the situation \( (\phi) [T(\phi) \equiv \phi = t] \) obtains). For then a sentence \( E(\tau) \) is true if it is the case that \( E(t) \) and false if it is the case that \( \sim E(t) \). Where they differ from classical semantics is that it is not universally the case that if \( \tau \) designates an entity \( t \) and \( \sim E(t) \) obtains, then \( E(\tau) \) is false. Rather, for \( E(\tau) \) to be false, every designatum of \( \tau \) must fail to satisfy \( E(\tau) \). That is, the factual content of \( E(\tau) \) according to Hypotheses A–C is the same as that of \( (\exists \phi) [T(\phi) \cdot E(\phi)] \), although the facts, if any, which are signified by these two sentences are by no means the same. Another prima-facie difference between classical semantics and the present generalization is rejection by the latter of the relation described earlier as “false signification.” In order to deal with the semantical status of false observation sentences, we have so far implicitly assumed that when a sentence \( S(s) \) of \( L_o \) is false, it is so because \( S(s) \) falsely signifies a fact \( \sim P(t) \) in virtue of \( S \) designating \( P \) and \( s \) designating \( t \). In such an interpretation, false observation sentences and
true observation sentences are on a par with respect to designating—both are conceived to be about some state of extralinguistic reality. But it may well be questioned whether such a concept of "false signification," in the sense defined by SP I–III, really can be extracted from classical semantics, which has always tended to confound the meanings of sentences with their factual significance. Even without drawing upon the theoretical dimension of language, it may be argued that a sentence can be false even though—or rather, because—it has no designatum (see [13]). Wholly aside from the problem of theories, it may be that "falsehood" is best characterized as a derivative semantical condition wherein a sentence is false if and only if it is cognitively meaningful but fails to signify a fact. If so, then classical semantics and the present hypotheses also agree—completely, not merely in the limit—with respect to the concept of falsehood.

Because it brings out an important property of theoretical concepts, I would now like briefly to present an informal argument in favor of the contention—i.e., Hypothesis B—that it would never be correct to say that a theoretical sentence 'E(r)' is not deducible from accepted theory T, falsely signifies a fact 'E(t)', even though 'r' as introduced by theory 'T(r)', designates t. Suppose that there exists a t such that T(t1) and ~E(t1), and also a 2 such that T(t2) and E(t2). Then by the present interpretation, 'r' is introduced by 'T(r)', designates both t1 and t2 (cf. Theorem 6), and one might argue on classical grounds that if this is granted, then we should have to say that 'E(r)' falsely signifies the fact that ~E(t1) as well as truly signifying the fact that E(t2). Now, the concept of "incorrectness," of which "falsehood" is a special case, is pragmatic—i.e., an entity is "incorrect" in a certain behavioral role only if it leads, actually or potentially, to error. But a sentence can lead one into error only when it is believed or accepted, for only then does one act upon the behavioral prescriptions of the sentence. Moreover, to believe or accept 'E(r)' in addition to accepting the theory 'T(r)' is to accept the enriched theory 'T(r) · E(r)'. Since by hypothesis it is the case that T(t2) · E(t2), it follows that 'E(r)' is then a consequence of the unambiguously true theory 'T(r) · E(r)', and so by Hypothesis A and SP I does not falsely signify ~E(t1); hence there are then no grounds on which to argue that 'E(r)' is incorrect. That is, when 'r' is introduced by theory 'T(r)', so long as it is the case that (∃φ)[T(φ) · E(φ)], the correctness of 'E(r)' is uncontaminated by the existence of a t such that 'r' designates t and it is the case that ~E(t), for 'E(r)' can lead one into error only in the course of adopting a new, improved theory 'T(r) · E(r)', and with respect to the latter theory under the conditions stipulated, 'E(r)' is in no way incorrect. But by SP II, if 'E(r)' is not incorrect, it does not falsely signify a fact. Hence 'E(r)' cannot signify a fact falsely so long as 'T(φ)' and 'E(φ)' are jointly satisfied. To drop the latter condition, we need but reflect that as brought out in the discussion of P 1 (see footnote 17), whether or not 'E(r)' falsely signifies ~E(t1) depends only on the relation of ~E(t1) to the use of 'E(r)', and not, in addition, on whether or not some other entity t2 satisfies 'E(φ)'. To be sure, 'E(r)' is false when there is no joint satisfier of 'T(φ)' and 'E(φ)', but only because 'E(r)' then fails to signify a fact truly, not because it signifies some fact falsely.

What the preceding argument reveals—and this is its real importance—is that when a theoretical sentence 'E(r)' is not entailed by the theory which gives 'r' its meaning, then the pragmatic effectiveness and hence the truth or falsity of E is essentially given by whether or not addition of E to the postulates of the theory would yield a correctly enriched theory. Hence a theoretical sentence E not entailed by an accepted theory cannot be said to have meaning in quite the same way that the theory and its consequences have meaning; rather, the meaning of such an E is best characterized as a disposition to have the meaning it would have were the theory enriched in a certain way. This makes more palatable the rather unpleasant consequence of Hypotheses A–C that although two theoretical sentences E1 and E2 may each be true separately, their conjunction may be false. For example, if two distinct entities t1 and t2 each satisfy the observational predicate 'T(φ)' of accepted theory 'T(r)', both 'r = t1' and 'r = t2' are separately true under Hypotheses A–C, yet 'r = t1' ·

27 When classical semantics analyzes the linguistic properties of the observation language through sentences of the form "'S(s)' in L0, asserts that P(t)," what is primarily being indicated is a relationship among the meanings of sentences in L0 and L; and it is necessary to be very careful in moving from this kind of an account to one analyzing the relations of expressions in L; to their designata, since not all meaningful expressions have designata, even when their syntactic role is that of a descriptive term.
(τ = t₂),’ which entails that t₁ = t₂, is false. But while this violates classical semantics, it makes a certain amount of intuitive sense upon reflection that while either ‘T(τ) • (τ = t₁)’ or ‘T(τ) • (τ = t₂)’ is a perfectly good (i.e., correct) enrichment of ‘T(τ)’ under the conditions stipulated, the stronger enrichment ‘T(τ) • (τ = t₁) • (τ = t₂)’ is false. However, a more penetrating analysis of this situation must await another occasion.

While the preceding considerations are rather fragmentary, they nonetheless expose a particularly vital aspect of the meanings of theoretical terms. It was commented earlier that there is an important sense in which such meanings are incomplete. We now see that in order to give pragmatic effectiveness to a theoretical sentence not entailed by the theory then in force, it is necessary to augment the theory until it does entail that sentence. Given any enrichable accepted theory T, there will be sentences, containing terms whose meanings are acquired through their participation in T, whose truth or falsity cannot be judged without thereby enriching the meanings of these terms. Any enrichable theory has an inherently concomitant envelope of unresolved theoretical questions which demand that the theory be supplanted by a better, more complete theory.

In regard to such enrichments, however, there is an apparent paradox which needs resolution. Suppose that there are entities t₁ and t₂ such that T(t₁) and T(t₂), but that E(t₁) and not E(t₂). Then theory ‘T(τ),’ if accepted, is true, and ‘τ’ designates both t₁ and t₂. When it comes to enriching the theory in regard to whether or not τ satisfies ‘E(ϕ),’ however, it would appear that we can have it both ways; the theory ‘T(τ) • E(τ)’ and the theory ‘T(τ) • ¬E(τ)’ are both true when accepted. But this might seem paradoxical; for if we can enrich the theory in two directions, why can’t we enrich it in both at once, giving us the logical inconsistency ‘T(τ) • E(τ) • ¬E(τ)?’ Or even if we do not try both directions at once, is not ‘T(τ) • E(τ)’ incompatible with ‘T(τ) • ¬E(τ)?’ The answer, of course, is that ‘τ’ has a different meaning in ‘T(τ) • E(τ)’ than it has in ‘T(τ) • ¬E(τ).’ By the Thesis of Semantic Empiricism, the meaning of a theory is unchanged by substitution of new theoretical terms for old; and in the present example, the enrichment ‘T(τ) • E(τ)’ is quite compatible with the enrichment ‘T(μ) • ¬E(μ).’ Similarly, it is possible to move in both semantical side, the significata of a given theoretical sentence will still depend, in part, on the nature of the other assumption formulas, but this is as it should be, for whether or not a given theoretical sentence is an acceptable enrichment of a theory depends in part on what other sentences are also to be added.

Suggestions for a behavioral theory of semantics. Let us conclude this section with a few words in outline of a behavioral theory of semantical relations. We saw earlier—and indeed, I do not see how it would be possible to dispute this truism—that whatever semantic properties are possessed by a set of sign-designs for a person o at time t are due to the way in which these signs are used by o at t. Since ‘use’ is an ambiguous and rather misleading term carrying teleological connotations of “purpose” or “intended causal effect,” this is better put by saying that the semantic properties of a sign-design s for o at t depend upon the kind of behavioral effect that s has on o at t, or, more generally, the kind of effect that presentation of s would have on o at t. From this it is but a short step to propose that (1) the cognitive meaning of a sign-design s for a person o at time t is some

---

I am indebted to W. Sellars for this adroit phrase.
aspect of the behavioral effect that s has, or would have, on o at t; and that (2) the designata of s (for o at t), if any, are determined by the cognitive meaning of s (for o at t). The adjective ‘behavioral’ has minimal restrictive force here—in particular, it is not meant necessarily to rule out “mentalistic” interpretations of meanings, for there is no reason why behavioral and mentalistic accounts of linguistic processes may not be describing the same events (cf. [6]). Rather, it is to emphasize that meanings are to be found in the dynamics of person-symbol interactions. When s is a sentence, its “cognitive meaning” may be described as a “prescribed behavioral adjustment,” since when a statement is under declarative consideration, certain behavioral tendencies or “sets” controlled by this sentence, perhaps highly removed from gross motor activities, are brought into play under a provisional status, where the degree of the latter (i.e., the degree of belief) is dependent upon factors additional to the cognitive meaning of the sentence. It should be noted that (1) does not suggest that a term must have a referent or that a sentence must signify a fact in order for the term or sentence to be meaningful.

We may elaborate this theory by two further hypotheses: (3) The cognitive meaning of a statement is compounded out of the cognitive meanings of the constituent terms in a definite way determined by the syntactical structure of the statement. This is not to imply that the meanings of constituent terms are always in some sense causally prior to the meanings of the sentences in which they occur, for we wish to interpret the meanings of theoretical terms as derived from the theoretical postulates. Yet, if the sense of a sentence is determined by its component terms and formal structure, as any acceptable theory of semantics must recognize,

[It is not true that the total “use,” or behavioral force, of an expression on the occasion of a particular occurrence is relevant to its semantic properties. Thus a given sentence may be employed as a simple declaration of what is believed to be the case (“The barracks will be cleaned tonight.”), as an interrogation (“The barracks will be cleaned tonight?”), or as a command (“The barracks will be cleaned tonight!”), to list but three broad categories of a vast number of possible uses. Yet, in each case the cognitive meaning of the sentence—i.e., the proposition in virtue of which it is able to be about the external world—is the same. While we do not normally think of questions and commands as designating anything, if their total function did not preserve a cognitive component, they could not serve their purpose—for example, an effective command must describe the desired state of affairs which comes to exist when the command is properly executed. The total linguistic status of the occurrence of an expression would appear to require description on two dimensions: (1) the cognitive meaning, or designative potential, of the expression; and (2) the function—i.e., assertion, query, command, etc. in the case of sentences, reference (and other uses?) in the case of descriptive terms—for which that cognitive meaning is being employed.

THE FACTUAL CONTENT OF THEORETICAL CONCEPTS

meanings must be compoundable according to certain definite principles. (4) If a statement S signifies a fact f, it does so because the cognitive meaning of S is “appropriate” to f. Just what ‘appropriate’ means in this context is difficult to pin down. By saying that the meaning of S is appropriate to f, we wish to indicate that the behavioral adjustment prescribed by S somehow prepares for, or adapts to, the fact f. In those cases where it would make sense to say that a person is aware that f is the case, we might say that S signifies f (for person o at time t) when the behavioral adjustment prescribed by S is suitably similar to the behavioral adjustment that would be set off by awareness that f. More generally, if we presume there is some behavioral adjustment, which might be called the “behavioral significance” of a fact f (for o at t), which is maximally and specifically appropriate (for o at t) to f, we may then propose that a statement S (truly) signifies fact f (for o at t) when the cognitive meaning of S (for o at t) is sufficiently similar to the behavioral significance of f (for o at t).

Before the substantive details of this relation can adequately be filled in, we shall need a much more developed science of behavior than is now available. For this reason, the present theory can be no more than a crude outline. However, this is enough to make plausible the possibility that a statement may signify more than one fact. For suppose that S is a statement suitably rich in meaning that it signifies exactly one fact. Then is it not possible that by “weakening” the meaning of S—i.e., by withdrawal of a certain portion of its prescribed behavioral adjustment—S would now be “appropriate,” in that way which characterizes designation, to a set of facts differing only in respect to a feature to which the weakened meaning of S no longer prescribes a differential adaptation? What I am suggesting, in other words, is that if semantic relations are grounded upon a similarity (or perhaps a more complex relation) between the behavioral prescriptions of symbols and behavioral significances of things symbolized, then designation may be a matter of degree, rather than an all-or-none affair. The more “weakly” S signifies f, the more it is possible for S also to signify other facts which are similar to f in suitable respects.

These suggestions may be sharpened by proposal of a similar analysis for the way in which a descriptive constant ‘s’ designates an entity t. There would appear to be some sense in which an entity may be said to have “behavioral import” for an organism. This cannot be analyzed simply
as the reaction produced in the organism by the entity acting as a stimulus, for organisms respond to facts, not stimuli as such (although the organism may respond to the fact that the stimulus is present—see [12]). Nonetheless, since the behavioral significance of a fact is determined by the entities it comprises—for if it were not so determined, the behavioral significances of facts sharing one or more constituents would not need to be related in the manner they in actuality are—we may regard the behavioral imports of entities as behavior elements out of which the behavioral significances of facts are compounded according to the way (i.e., logical structure) in which the entities constitute the fact. Now, we have already hypothesized that the (cognitive) meanings of statements are compounded out of the (cognitive) meanings of their constituent terms. Hence, if it is the case that when a constant ‘s’ designates an entity t, the meaning of ‘s’ is sufficiently similar to the behavioral import of t, it becomes clear (at least in overview) how a statement might signify a certain fact in virtue of the statement’s formal structure and the meanings of its constituent terms. It is now especially easy to suggest the conditions under which multiple designata for descriptive terms, and derivatively for statements, might come about. For if the relation between the meaning of a symbol and the behavioral import of its designatum is that the former is, or closely resembles, a part of the latter, then the meaning of a sufficiently weak symbol might be a behavioral effect common to the behavioral imports of several entities. The process of concept formation would then consist of endowing a symbol with behavioral force—i.e., cognitive meaning—which, in turn, determines the entities, if any, to which this symbol refers. The stronger, or richer, the meaning of the symbol, the fewer the entities designated by it, while if it is possible to make the symbol sufficiently strong in meaning, it will have a unique designatum.

There are obviously many serious problems and ramifications to this theory which the present outline has not begun to explore. However, if the theory appears to have any merit at all, the purpose for which it has been suggested here has been accomplished, namely, to show it to be not unreasonable that a term might simultaneously designate more than one entity. Moreover, this sketch is further helpful in clarifying the manner in which a theory, though equivalent in force to an observation sentence, nonetheless manages to enrich the language. We have argued that although the factual content of a theory is identical to that of its Ramsey sentence, ‘T(τ)’ and ‘(∃φ)T(φ)’ do not signify the same fact; ‘T(τ)’ signifies each member of a (possibly empty) set of (in the main) non-observational facts, each of which entails the observational fact, if any, signified by ‘(∃φ)T(φ).’

Now, I see no reason why, if a certain complex behavioral effect can be compounded out of other effects, it might not also be compoundable in more ways than one. It is then not implausible that an organism of sufficient behavioral intricacy could take a complex effect E, compounded from behavior components acquired previously, and restructure it in such a way that some of the constituents of the restructured E are behavior elements which were not previously available. Thus it seems quite conceivable, under the present semantical theory, that a theoretical term ‘r’ could be infused with just that degree of meaning which would make the behavioral force of ‘T(τ)’ essentially the same as that of ‘(∃φ)T(φ)’ when the latter already exists in the organism’s behavioral repertoire. Presumably, this could be accomplished simply by using the same symbol ‘r’ in the various theoretical postulates of T—it should not be necessary actually to construct the full conjunction, ‘T(τ),’ of theoretical postulates. For if each theoretical postulate contributes a meaning component to ‘r,’ the combined effect should be the same as if ‘r’ acquired its meaning directly from use in the conjunction, ‘T(τ).’ Further, it is important to note that while the force of ‘T(τ)’ is the same as that of ‘(∃φ)T(φ),’ if ‘E(τ)’ is entailed by ‘T(τ)’ but not conversely, the meaning of ‘E(τ)’ is richer than that of ‘(∃φ)E(φ).’ It will be recalled that one difficulty in regarding ‘T(τ)’ as a peculiar way of asserting that (∃φ)T(φ) was that no comparable translation exists for ‘E(τ).’ But theoretical statements derived from T pose no interpretative difficulties once we realize that ‘r’ functions as a name in that it has a fixed meaning (so long as the theory is not enriched or otherwise altered) which may be carried from one statement to another. The only difference between theoretical terms and observational terms, under this interpretation, is that the meanings of the former are weaker, and their referents thus possibly more numerous, than those of the latter. Hence, theoretical terms constitute a genuine enrichment of language, rather than peculiar formal devices for deriving observational predictions, and may themselves become “observational” if their meanings are given sufficient strength through an accepted, true and sufficiently forceful theory.
THE FACTUAL CONTENT OF THEORETICAL CONCEPTS

It would be highly surprising if any explication of a problem so philosophically basic as the meanings of theoretical terms did not have important implications for many other related problems as well. In closing, I would like to consider, very briefly, the import of the present analysis for certain unresolved problems of current interest.

Identification and reduction. For those who prefer a realistic interpretation of theoretical terms, it is unnecessary to conceive of theoretical entities as partaking, somehow, of a different kind of “reality” from observational entities of the same type. There is no reason why, in principle, a theoretical entity cannot become “known” in the same way that observational entities are known. In fact, the referent (or a referent) of a theoretical term may be an entity already independently accessible to the observation language. (The “phantom burglar” postulated by the police to account for a sudden upsurge in larceny may turn out to be the police chief himself.) Again, we may seek to “reduce” the theoretical terms of one theory to those of another, success in which is sometimes regarded as confirmation of the “reality” of the reduced entities. (Thus in genetics, the gene has appeared increasingly real as cytological theory has proliferated.) In either instance, we speak of finding the “identity” of the entity for which the theoretical term was at first only a “promissory note.” How is such an identification to be analyzed?

In all cases where a theoretical entity is “identified,” the crucial step consists in an assertion ‘\( \tau = d \),’ where ‘\( \tau \)’ is the theoretical term whose identity is being proposed and ‘\( d \)’ is a designative expression which is either (1) wholly in the observation language; (2) a demonstrative (e.g., “So that’s what \( \tau \) is!”), the analysis of which case is essentially that of (1); or (3) contains other theoretical terms, ‘\( \mu_1 \), . . . , ‘\( \mu_n \),’ in which case ‘\( \tau \)’ has been “reduced” to ‘\( \mu_1 \), . . . , ‘\( \mu_n \).’ We need not be concerned here with the precise analysis of the identity relation (except, preferably, to assume that ‘a = b’ does, in fact, make an assertion about a and b, rather than being an ellipsis for a semantical statement such as ‘(x) (‘a’ designates x if and only if ‘b’ designates x)). What we are now investigating is the meaning of an assertion of identity when one of the expressions involved is a theoretical term.

Suppose that ‘\( \tau \)’ is a theoretical term whose meaning is defined by the theory ‘\( T(\tau) \),’ and that ‘\( d \)’ is a descriptive term in the observation lan-


guage. Under what circumstances would we be willing to say that ‘\( d \)’ is the identity of ‘\( \tau \)—i.e., to claim that ‘\( \tau = d \)? It is, of course, obvious that if it is not the case that ‘\( T(d) \),’ it would be most incorrect to assert ‘‘\( \tau = d \),’’ for as was shown earlier, whatever ‘\( \tau \)’ designates, it must be something that satisfies ‘\( T(\phi) \).’ But suppose that it is the case that ‘\( T(d) \).’ Would we not then be justified in claiming that ‘\( d \)’ is the identity of ‘\( \tau \)? It is hard to deny this claim, for what other criterion could we invoke in deciding whether or not ‘\( \tau = d \);’ yet the matter is not so simple as all that. First of all, we must recognize, presuming our earlier analysis to be correct, that if ‘\( T(d) \)’ is the case, then ‘\( \tau \)’ designates ‘\( d \).’ But this is in itself insufficient to conclude that ‘\( \tau = d \).’ For suppose the fact that ‘\( T(d) \)’ necessitated the conclusion that ‘\( \tau = d \).’ Then if some other observational entity ‘\( d^* \),’ different from ‘\( d \),’ also satisfies ‘\( T(\phi) \),’ we would have to conclude also that ‘\( \tau = d^* \),’ which by the transitivity of identity would entail, contrary to hypothesis, that ‘\( d = d^* \).’ The reason the fact that ‘\( \tau \)’ designates ‘\( d \)’ does not necessitate the conclusion ‘\( \tau = d \)’ is that the latter adds a further restriction on the designata of ‘\( \tau \)’ beyond that imposed by the theory ‘\( T(\tau) \).’ Assertion that ‘\( d \)’ is the identity of ‘\( \tau \)’ involves not only the judgment that ‘\( T(d) \),’ but also the decision to enrich the theory in this way.

To say that asserting ‘\( \tau = d \)’ involves a decision is not to imply that the decision is a difficult one to make. For if it can be determined with high certainty that an observational entity ‘\( d \)’ satisfies ‘\( T(\phi) \),’ then to accept the enrichment ‘\( \tau = d \)’ is to accept the theory ‘\( T(\tau) \cdot (\tau = d) \)—i.e., ‘\( T(d) \)— which not only is verified by the fact that ‘\( T(d) \),’ but also becomes supplemented by further facts known about ‘\( d \).’ That is, acceptance of the enrichment ‘\( \tau = d \)’ not only changes the status of the theory from hypothesis to known fact in this case, it also increases its usefulness. Conversely, to deny the identity of ‘\( \tau \)’ with ‘\( d \)’ is to adopt the counterenrichment ‘\( T(\tau) \cdot (\tau \neq d) \),’ while the latter not only is unlikely to have any worthwhile consequences beyond those of ‘\( T(d) \),’ but also stands a reasonably good chance of being false. Hence it is an almost automatic process, and rightly so, to identify a theoretical entity with the first entity observed to satisfy the theory.

The situation is somewhat more complicated in the case of reduction, and my remarks here can be no more than fragmentary. Suppose that an accepted theory ‘\( T \)’ can be written as the conjunction of two autonomous subtheories, ‘\( T_1 \)’ and ‘\( T_2 \),’ which contain no theoretical terms in common—i.e., that ‘\( T \)’ is of the form ‘\( T_1(\tau) \cdot T_2(\mu) \).’ We may then describe ‘\( T_1 \) and
William W. Rozeboom

$T_2$ as separate theories, say a macrotheory and a microtheory, adopted simultaneously. Suppose further that there is an expression ‘$d_\mu$’ containing theoretical terms of $T_2$, such that ‘$T_1(d_\mu)$’ is entailed by ‘$T_2(\mu)$’. That is, suppose the microtheory $T_2$ implies the existence of, and supplies a descriptive expression for, an entity which exemplifies the macrotheory $T_1$. It follows that (a) the truth of $T_1$ is entailed by the truth of $T_2$, and (b) the entities designated by ‘$d_\mu$’ are a subset of the entities designated by ‘$\tau$’. Under such circumstances, we should be tempted to identity $\tau$ with $d_\mu$—i.e., to assert ‘$\tau = d_\mu$’—an enrichment which is equivalent simply to dropping ‘$T_1(\tau)$’ as a separate hypothesis. (Thus as is also true in the case of observational identification, the enrichment sustained by a theory through reduction of its theoretical elements to constructs in another theory consists in assimilating the theory to a set of beliefs external to the theory, and abandoning the theory as a separate hypothesis.) And to be sure, if we are certain that ‘$T_2(\mu)$’ is true, the reasons for identifying $\tau$ with $d_\mu$ are as strong and as legitimate as identifying $\tau$ with some observational entity, $d_0$ when it is known that $T_1(d_0)$.

But there is an important difference between observational identification and theoretical reduction. In the former case, we considered the legitimacy of asserting ‘$\tau = d_\mu$’, given knowledge that $T_1(d_0)$. In the latter case, on the other hand, we are judging the assertion of ‘$\tau = d_\mu$’, given knowledge that ‘$T_2(\mu)$’ entails ‘$T_1(d_\mu)$’. The difference is that while the theory ‘$T_1(\tau)$’ is confirmed by the fact that $T_1(d_0)$, the fact that ‘$T_2(\mu)$’ entails ‘$T_1(d_\mu)$’ does not confirm ‘$T_1(\tau)$’—it only shows that ‘$T_1(\tau)$’ must be true if ‘$T_2(\mu)$’ is true. There is thus the danger, if we accept ‘$\tau = d_\mu$’, that ‘$T_2(\mu)$’ is false, a contingency which, if realized, in general leaves ‘$d_\mu$’ without a designatum and hence falsifies ‘$T_1(d_\mu)$’ even though ‘$T_1(\tau)$’ by itself may remain quite true (since ‘$\equiv (\exists \phi)T_1(\phi)$’ does not entail that ‘$\equiv (\exists \phi)T_1(\phi)$’ unless $T_1$ and $T_2$ are analytically equivalent). That is, to identify $\tau$ with $d_\mu$ is to risk replacement of a theoretical expression which has a referent by another which does not, and thus to gamble the success of one theory upon that of another. To draw out the implications of this for the practical aspects of theory building, and to buttress the argument by citing specific examples, would require a more extensive discussion than is practical here. The conclusion which would ultimately be drawn is that although the relation between a macrotheory and a microtheory (or, for that matter, between two sets of theoretical terms on the same level) may be such as to suggest strongly that certain microstructures are the “identities” of the theoretical macroentities, it is best to remain noncommittal about the identities as long as the microtheory sustains a reasonable doubt, or, at most, to carry the identity assertion as a kind of auxiliary hypothesis which may be discarded, if necessary, without otherwise necessitating any change in the macrotheory.

To summarize: To “identify” a theoretical entity is to make both a factual judgment and a decision about the subsequent use of the theoretical term. To enrich the theory ‘$T(\tau)$’ by adoption of the identity assertion ‘$\tau = d$’ is a legitimate and desirable move when (a) there is an entity which is designated by ‘$d$’, and (b) an entity designated by ‘$d$’ satisfies ‘$T(\phi)$’. To the extent there exists doubt that ‘$d$’ meets either condition, assumption that ‘$\tau = d$’ is a dubious maneuver which should never be made unless the line of retreat remains clearly visible.

Implicit definition. One of the stickier problems of analytic philosophy has been what to say about (stipulative) “implicit” definitions. Since theoretical postulates have traditionally been taken as the paradigm case of implicit definition, the present account of theoretical concepts, if tenable, should substantially clarify this issue.

A stipulative definition is a sentence ‘$D(a)$’ (or set of sentences, the conjunction of which reduces to the first case) through which meaning is assigned to one (or more) of its constituent terms ‘$a$’. When a stipulative definition is of the form ‘$a \equiv d$’ (or a conjunction of sentences of the form ‘$a_1 \equiv d_1$’), it is known as an “explicit” definition. When ‘$D(a)$’ is of a form other than ‘$a \equiv d$’, it is known as an “implicit” definition.

Since any analysis of implicit definition sufficiently broad to cover all forms of ‘$D(a)$’ other than ‘$a \equiv d$’ will undoubtedly be applicable to the latter as well, it would seem more logical to regard (stipulative) explicit definition as a special case of (stipulative) implicit definition.

According to the views developed earlier, if the meaning of a term ‘$a$’ is determined (solely) by its usage in the sentence ‘$D(a)$’, then ‘$a$’ designates every entity $t$ such that $D(t)$. This account does not explain how

---

Footnote 30: Explicit definitions are frequently written ‘$a \equiv_{def} d$’. How ‘$\equiv_{def}$’ should be analyzed is not easy to decide. Since the force of ‘$\equiv_{def}$’ does not appear to be the same as ‘$=$’, the subscript does not occur vacuously, yet the Identity in ‘$a \equiv_{def} d$’ does not seem to differ from the Identity in ‘$a = d$’. One interpretation of ‘$\equiv_{def}$’ is in terms of linguistic norms, where the sentence ‘$a \equiv_{def} d$’ is regarded not as an assertion, but a rule, conformity to which necessitates the truth of ‘$a = d$’. Still another interpretation is to regard ‘$a \equiv_{def} d$’ as an ellipsis for a more complex descriptive statement relating why it is the case that ‘$a = d$’. We shall here treat ‘$a = d$’ (and more generally, ‘$D(a)$’) as the “definition” itself (cf. [17], pp. 139, 149f). Whether this is correct, or whether
William W. Rozeboom

‘a’ acquires this meaning—such an explanation lies within the province of the psychology of language. What is of philosophical relevance is that ‘a’ designates in this way. Thus whatever role the symbol complex ‘D(a)’ may play in the acquisition by ‘a’ of meaning, the semantical status of ‘D(a)’ is simply that of a descriptive sentence. What needs to be spelled out in greater detail, however, are the truth conditions of ‘D(a)’: (1) An implicit definition is not logically true. While it is difficult to find a wholly satisfactory explication of the classical concept of “logical truth,” the underlying notion is that a statement to which this term applies is true or false by virtue of its logical form. But an implicit definition ‘D(a)’ is true by virtue of its logical form only if the expression formed by replacing ‘a’ with any descriptive constant of the same formal type is necessarily true. This obtains only when the property D(φ) is necessarily possessed by every entity of the appropriate type, in which case (as will be elaborated below) ‘D(a)’ is empty of definitional force. Hence an implicit definition cannot be tautological. However, (2) an implicit definition, if true, is true ex vi terminorum. Given that ‘a’ designates an entity t, it is unnecessary to inquire further as to whether or not D(t) is the case in order to pass judgment on the truth of ‘D(a).’ It is in the meaning of ‘a’ that any entity designated by ‘a’ satisfies ‘D(a).’ 81 On the other hand, it is not the case that ‘D(a)’ has no factual content, or that ‘D(a)’ is not empirically falsifiable, for (3) the empirical force of an implicit definition is contained in the defined term’s success or failure at designating. While ‘a’ designates any entity which satisfies ‘D(φ),’ it by no means follows that there is any such entity. Hence, ‘D(a)’ is empirically true or false according to whether or not there exists an entity designated by ‘a’—i.e., according to whether or not (∃φ)D(φ).

To summarize: A statement, ‘D(a),’ which implicitly (or, as a special case, explicitly) defines a term, ‘a,’ does not fit conveniently into the traditionally analytic-empirical dichotomy of the truth grounds of statements. Since it is inconceivable that a (so defined) should not exemplify D, one might think that ‘D(a)’ should be analytically true. On the other hand, the most important cases of implicit definition, scientific theories,

81 Some complications may arise here if ‘D(a)’ is not unified.

reveal clearly that implicit definitions are not compatible with all possible facts, and hence must embody a factual commitment. Traditional semantical analysis presupposes that all primitive extralogical constants of cognitively meaningful statements do, in fact, designate. The present analysis suggests, on the other hand, that it is not true that all sentences which violate this presupposition are meaningless, and that there is an important class of empirically significant statements whose truth values depend wholly upon whether or not all their primitive extralogical constants have designata. (Actually, pursuit of this line of thought in light of Hypotheses A–C, above, leads to a radical reinterpretation of the traditional concepts of “analytic” and “synthetic,” but this is far beyond the scope of the present discussion.)

Definite descriptions. This problem has been a philosophical headache for many years. The difficulty is not so much a lack of interpretations as it is a surfeit of them. While Russell’s [15] famous analysis is perhaps the most widely accepted, and seems to reproduce most satisfactorily the intuitive truth conditions of statements using definite descriptions, it has its own drawbacks, while alternative interpretations find themselves parting with common sense or the Law of Excluded Middle in the case of unsatisfied descriptions.

Part of the difficulty in finding an intuitively convincing explication of descriptions probably lies in an ambiguity in common usage. It seems to me that in de facto language practices, descriptions are frequently used as demonstratives. After all, one can call attention to an object by naming some of its distinguishing features as well as by pointing at it, and when used in this way, asserting that the A is a B would have essentially the same force as saying that this is a B—it need not even be the case, in this instance, that there is only one A (cf. [18], p. 186), or even that the entity referred to is an A, so long as the context of usage is such that the sign sequence, ‘the A,’ momentarily designates the appropriate entity.

However, while descriptions may in fact occasionally be used as demonstratives, this is certainly not the case which has stimulated philosophical concern. What needs to be determined is what is meant by saying, “The A is a B,” when the A is not necessarily accessible to a demonstrative. The Russellian analysis, which takes such an assertion to be equivalent to “There is an x such that B(x), and for any y, A(y) if and only if y = x,” has one fatal drawback: Under this analysis, descriptions do not designate. Russell himself was quite explicit on this point. Although the assertion
William W. Rozeboom

“The A is a B’ appears to be of the logical form ‘ψ(x),’ in which ‘x’ designates the (only) entity which possesses a property ϕ, Russell contends that the genuine logical form is ‘(∃x)(ψ[x] · (y)(ϕ(y) ≡ y = x)],’ and in the latter expression, there is no term, or complex of terms, which designates any entity which exemplifies ϕ. Hence under the Russellian analysis, definite descriptions are not designators, but syntactical condensations. But surely this seriously undermines the Russellian account as an acceptable analysis of statements involving definite descriptions, for it seems to me inescapable that a description is, in actual language practice, used syntactically in essentially the same manner that we would use a descriptive constant of the same type level, and that when we say, ‘the A,’ we intend to refer to the A. On the other hand, use of the definite article to assert that the A is a B when one could otherwise say that an A is a B, would seem to indicate that ‘The A is a B’ entails that there is one and only one entity which is an A; while conversely, existence of exactly one entity which is an A and which, moreover, is also a B, is certainly a sufficient condition for the truth of ‘The A is a B.’ Hence it would appear that ‘The A is a B’ and ‘(∃x)(B[x] · (y)[A(y) ≡ y = x])’ have exactly the same truth conditions, and a thoroughly satisfactory explication of the former would seem to require the force of the latter, but the logical form ‘ϕ(x).’

While we need not here make any definite commitments as to what this explication might be, it is instructive to observe that in important respects, definite descriptions appear to be very similar to theoretical terms. According to the position developed earlier, if ‘T(τ)’ is a theory, then while ‘T(τ)’ has the same truth conditions as ‘(∃ϕ)T(ϕ),’ ‘τ’ actually designates that entity (or entities), if any, in virtue of which the latter is true. Implied by this analysis is the idea that a language user does not necessarily require having had direct awareness of an entity in order to refer to it—by appropriate synthesis of meaning components available through other sources, he is able to construct an expression which designates the entity. If this conclusion is correct, then it is conceivable that the phrase ‘the A’ may also be a designative expression of this kind. In

I strongly suspect that the ultimate components from which all cognitive meanings are synthesized are those aroused by direct experience. This possibility must not be confused, however, with the question of whether all meaningful linguistic expressions are constructed from a phenomenal language. Contrary to frequent philosophic misconception, meanings are to be found among psychological processes even when there is no corresponding language framework to govern them.

THE FACTUAL CONTENT OF THEORETICAL CONCEPTS

particular, if ‘the A’ is regarded as a theoretical term introduced by the unified theory ‘(x)[A(x) ≡ x = the A],’ then by Hypotheses A–C, ‘the A’ has a referent if and only if there is exactly one entity which satisfies ‘A(x),’ while ‘The A is a B’ is true or false according to whether or not there is exactly one A which, moreover, is also a B. Thus the present analysis of theoretical concepts makes it plausible that a definite description could carry the force of an existential operator in the Russellian fashion and yet serve as a genuine designator. In fact, this line of reasoning also suggests an explication for that neglected waif of linguistic analysis, the indefinite description. Suppose we regard the phrase ‘an A’ as a theoretical term introduced by the unified theory ‘A(an A).’ Then by Hypotheses A–C, ‘an A’ designates every satisifier of ‘A(x);’ while the sentence ‘An A is a B’ is true if and only if (∃x)[A(x) · B(x)], yet is of the logical form ‘ϕ(x).’

The meaning criterion. One of the dominant themes of modern analytic philosophy—certainly a guiding motive of the logical empiricist movement—has been the search for the “meaning criterion,” a principle by which can be determined whether or not a given expression is cognitively meaningful. For difficulties, if any, which reside in the meaningfulness of expressions constructed wholly in the observation language, the present views have little relevance. On the other hand, to the extent that the meaning problem is concerned with the meanings of nonobservational terms, the present analysis of theoretical concepts provides a simple and plausible solution. It has been here contended that the meaning of a theoretical term is not something brought with it to the context of usage, but is given to it by the (accepted) postulates which contain it. If this is correct, then it is misleading to construe the desired meaning criterion as a yes-or-no test to be applied to terms whose meaningfulness is in doubt. Rather, cognitive meaningfulness is better seen as a matter of degree, and we should ask what meanings the usage of such terms has conferred upon them.

Suppose that the (cognitive) meaning, if any, of a term ‘a’ is imparted to it by its use in a set of (perhaps provisionally) accepted sentences, the conjunction of which is ‘D(a).’ That is, ‘D(a)’ is the implicit definition of ‘a.’ Under what circumstances would we say that ‘a’ is meaningless? One intuitively plausible criterion, that a term is cognitively meaningless when it has no designatum, does not seem to be acceptable. For, if the present interpretation of theoretical concepts is correct, the assertion ‘D(a)’ is
false if and only if ‘a’ has no designate—i.e., if and only if it is the case that \(~ (\exists \phi) D(\phi)\). Then if ‘a’ were meaningless when it has no designate, (a) it would be possible for a sentence containing meaningless terms to be false, and (b) decisions about meaningfulness would necessitate appeal to extralinguistic facts, so that meaning judgments would be logically subsequent, rather than prior, to truth judgments. Moreover, not all apparently meaningful expressions in ordinary use have designate. For example, most philosophers would hold that definite descriptions are meaningful, even when there exists no entity which uniquely satisfies the description. Or again: we should surely not wish to say that ‘square-circleness’ is meaningless, even though it would seem most peculiar to say that there exists a property, Square-circleness.

It would thus appear, since a syntactically well-formed implicit definition is true when it succeeds in assigning designate to its definienda and false otherwise, that there must be some rudimentary sense in which all terms actually in use have meaning—which is really not so surprising, since assignment of a syntactic role to a sign-designta must surely confer some behavioral effect upon it. On the other hand, it is by no means the case that all terms must have a pragmatically significant meaning. Suppose that ‘a’ is defined by ‘a = a.’ We should scarcely feel that ‘a’ has thereby acquired any useful meaning, for the property of self-identity is possessed by any entity whatsoever. Similarly, in the other extreme, we should be reluctant to grant that a term defined by a logically inconsistent definition had been given any useful force. More generally, if ‘D(\phi)’ is a predicate whose applicability can be determined on logical grounds alone, the assertion ‘D(a)’ contributes no useful meaning to ‘a.’ Conversely, if the implicit definition of a term ascribes to it a logically contingent predicate, then any sentence containing this term has an empirically falsifiable existential commitment, and the term must have useful meaning. I propose, therefore, that a term is “effectively meaningful” (i.e., having pragmatic force) when and only when its usage is logically consistent and imposes extralogical limits on its possible referents. Thus, when ‘D(\phi)’ is a logically consistent monadic predicate in which ‘\phi’ is a purely logical variable, ‘D(a)’ gives ‘a’ effective meaning if and only if, when ‘D(a)’ is adopted, it is not the case that ‘a’ necessarily designates every entity in the range of ‘\phi’—i.e., if and only if ‘(\phi) D(\phi)’ is not necessarily true. The reason for stipulating that ‘\phi’ must be a logical variable is that if its range were a nonlogical category, it would not be logically decidable whether

The factual content of theoretical concepts

‘D(\phi)’ is applicable to a given entity t even though the empirical fact that t is in the range of ‘\phi’ logically entails that D(t). Thus if ‘x\phi’ is a variable ranging over humans, ‘(x\phi)(x\phi = x\phi)’ is logically true, but the implicit definition ‘a\phi = a\phi’ gives ‘a\phi’ effective meaning, since its designate are restricted to humans. (Actually, in a language which contains nonlogical variables, a term is endowed with empirical commitments simply by choosing it to be of a formal type which represents a nonlogical category. It is by tracing the implications of this and similar considerations that we can appreciate the necessity for terms whose formal types represent purely logical categories.)

For the general case of an n-adic implicit definition ‘D(a_1, \ldots, a_n),’ the formalized meaning criterion is more complex than in the monadic case, since some but not all of the defined terms may be given effective meaning. For example, suppose that ‘D_1(a_1)’ effective-meaningfully defines ‘a_1,’ that ‘D_2(a_2)’ is tautologous, and that ‘D(a_1, a_2)’ is equivalent to ‘D_1(a_1) \cdot D_2(a_2).’ Then ‘D(a_1, a_2)’ gives effective meaning to ‘a_1’ but not to ‘a_2.’ Here, as in the monadic case, ‘a_2’ is effectively meaningless since its designate remains unrestricted; yet ‘D(a_1, a_2)’ is not necessarily true, and, if unified (cf. Definition 10), would give a designate to ‘a_2’ only if it also provides one for ‘a_1.’ To be sure, ‘D_1(a_1)’ and ‘D_2(a_2)’ are undoubtedly autonomous subdefinitions (cf. Definition 9) of ‘D(a_1, a_2),’ in this instance, and if so, permit ‘D(a_1, a_2)’ to assign designate to ‘a_2’ whether ‘D(\phi)’ is satisfied or not. However, we have not so far attempted to specify the conditions of autonomy, and if a plausible criterion of effective meaningfulness can be found without prior assumptions about autonomy, then this may also help to clarify the latter. It will be noticed in the present example that whether ‘D(a_1, a_2)’ is unified or not, if there exists one pair of entities t_1, t_2 such that D(t_1, t_2), then ‘a_2’ designates all entities in the range of its formal type, since for any entity t_1 in that range, it is the case that D(t_1, t_1). But this would seem more generally to be an adequate formalization of the notion that if a definition gives effective meaning to some but not all of its definienda, failure of an effectively meaningless definiendum to designate all entities of its type should result only from lack of designate for the effectively meaningful definienda. I suggest, therefore, that the following formal criterion (of which the monadic instance already treated may readily be seen to be a special case) is characteristic of the conditions under which a new descriptive term is given effective meaning through its usage with other terms.
THE FACTUAL CONTENT OF THEORETICAL CONCEPTS

Postulate 7. The term '$a_i$' is given effective meaning by a logically consistent implicit definition $'D(\alpha_1, \ldots, a_n)'$, where none of the definienda, $'a_1', \ldots, 'a_n'$, occur in the predicate $'D(\phi_1, \ldots, \phi_n)'$, and $'\phi'_i$ is a purely logical variable, if and only if it is not the case that $(\exists \phi_1, \ldots, \phi_n) D(\phi_1, \ldots, \phi_n)$ entails $(\phi_i) (\exists \phi_1, \ldots, \phi_i - 1, \phi_i + 1, \ldots, \phi_n) D(\phi_i, \ldots, \phi_n)$.

The postulate may be applied to the case where $'\phi'_i$ has extralogical restrictions on its range by first putting $'D(\alpha_1, \ldots, a_n)'$ into L-normal form (see p. 323, above). The restriction of P 7 to logically consistent definitions is to allow for the possibility that if $D$ can be decomposed into several autonomous subdefinitions, the logical inconsistency of one of these should not deny the remainder an opportunity to confer effective meaning on their definienda.

What can be said in justification of P 7 from the standpoint of intuitive conditions of meaningfulness? On the whole, these are so vague as to be of little assistance in this respect. However, I shall conclude with three observations which, I think, lend weight to the adequacy of the proposed criterion.

1. Postulate 7 is not merely a syntactical criterion. Those philosophers who have taken the search for a meaning criterion most seriously have usually attempted to characterize the conditions of meaningfulness in terms of syntactical relations among sentences containing the term in question and other sentences whose meaningfulness is not in doubt. But the meanings of expressions are by no means fully determined by their syntactical properties, and in particular, the implications of a sentence are not necessarily exhausted by its formal consequences. Hence a meaning criterion which draws upon only the syntactical features of language is bound to prove inadequate. In contrast, by making entailment—i.e., a relation between the factual contents of sentences (see p. 297, above) due to their meanings as well as to their syntax—a critical ingredient of the criterion, P 7 addresses itself directly to the full linguistic force of the term whose meaning is under consideration.

2. While P 7 does not draw specifically upon syntactical properties, it nonetheless satisfies a certain syntactical condition which has been thought to pose difficulties for a proposed meaning criterion, namely, the requirement that meaningfulness be invariant under syntactical equivalence transformations (see [4], p. 55). If $'a'$ is meaningless when defined by $'D(a)'$, and $'D(a)'$ is formally equivalent to $'D^*(a)'$, then $'a'$ must also be meaningless when defined by $'D^*(a)'$; hence it is a condition on the adequacy of a meaning criterion that it yield this result. To prove this follows from P 7, we have to show that if $[a]'D(a) \equiv D^*(a)'$ is formally true and if $[b]'(\exists \phi) D(\phi) \supset (\phi) D(\phi)'$ is analytically true, then $'(\exists \phi) D^*(\phi)'$ also entails $'(\phi) D^*(\phi)'$. We observe first of all that since $D(a)$ formally entails $'(\exists \phi) D(\phi)'$, it follows from $[a]$ that $D^*(a) \equiv (\exists \phi) D(\phi)$ is formally true. Hence by Lemma 1, $[c]'(\exists \phi) D^*(\phi) \supset (\exists \phi) D(\phi)'$ is also formally true. Now as easily proved from Lemma 1, a sentence of the form $'F(a)'$, in which the matrix $'F(\cdot)'$ does not contain $'a'$, is formally true if and only if $'(\phi) F(\phi)'$ is formally true. Hence from $[a]$, $'(\phi)[D(\phi) \equiv D^*(\phi)]'$ and thus also $[d]'(\phi) D(\phi) \equiv (\phi) D^*(\phi)'$ is formally true. Then from $[b]$ and $[d]$, $'(\exists \phi) D(\phi)'$ entails $'(\phi) D^*(\phi)'$; and hence from $[c]$, $'(\exists \phi) D^*(\phi)'$ entails $'(\phi) D^*(\phi)'$. Q.E.D. Thus under P 7, the effective meaningfulness and, conversely, meaningfulness of an implicitly defined term is invariant over formally equivalent forms of the definition.

3. My final observation is less concerned with P 7 as such than with the inconsistency of certain widespread beliefs about the conditions of meaningfulness for terms not introduced into the language ostensively. It is widely held that in order for a theoretical term to have meaning, its defining postulates must lead to some empirical conclusion—i.e., that if a theory $'T(\tau)'$ confers meaning upon $'\tau' , 'T(\tau)'$ must have an $O$-consequence which is not analytically true. This stipulation can be made precise, of course, only by defining 'analytic,' but we may presume that persons who subscribe to this belief would also include statements of form $'d = d' among the analytic truths. It is universally agreed, moreover, that an explicit definition, $'a =_d d' , is a perfectly legitimate way to confer meaning upon 'a.' Now, it has already been argued that there is no difference in kind between a (stipulative) explicit definition and a set of theoretical postulates--both give meaning to previously neutral symbols by using them in a context with other symbols which have already attained meaning. If this view is accepted, then any test of the meaningfulness of theoretical terms must also apply to explicitly defined terms. But the prime consequence of $'a = d' is $'(\exists \phi) (\phi = d)' , which is analytic if $'d = d' is analytic. Hence, if 'analytic truth' applies to formally valid sentences containing nonlogical terms, the belief that a meaningful theory must have nonanalytic consequences is incompatible with the belief that explicitly defined terms are meaningful. The relevance of these remarks to P 7 is that the latter does not imply that an implicit definition must have
nonanalytic consequences in order for its definienda to receive effective meaning. That this is how matters should be may be appreciated in greater generality by realizing that so long as the predicate ‘D(ϕ)’ contains meaningful descriptive constants, the sentence ‘(∃ϕ)D(ϕ),’ even when analytic, contains effective meaning components carried by its descriptive terms which may then be mobilized to give effective meaning to ‘a’ when defined by ‘D(a).’

Thus not only does P 7 meet certain general and rather difficult conditions of adequacy, it also avoids the inconsistency in what is probably the most widely held intuitive condition on a meaning criterion. Moreover, the line of reasoning of which it is a culmination makes clear just what sort of semantic desiderata are lacking in a term which fails to meet the criterion. I submit, therefore, that even if P 7 fails to capture all the nuances that we might wish of a meaning criterion, it will serve at the very least to define a certain interesting kind of meaningfulness.

REFERENCES