Minnesota Studies in the Philosophy of Science, Vol. III, 1962, pp. 273–357, (eds) Herbert Feigl and Grover Maxwell. University of Minnesota Press.

The Factual Content of Theoretical Concepts

Philosophers of science have more and more united in rejecting the older positivistic judgment that all descriptive words of an ideal language are, or are explicitly definable on the basis of, terms whose referents are phenomenally "given." It is far from surprising that tendencies toward a phenomenal reductionism should be a serious philosophical pressure within many critical thinkers, for it is indeed difficult to see how the actual "content" of thought (whatever such an expression might mean) could transcend the limits of direct experience. Yet repeated failure to realize such a program increasingly dims the likelihood that scientific or everyday language can be reduced to phenomenal terms alone. To be sure, this might be interpreted as revealing merely the semantic imperfections of existent linguistic practices, but such a gambit is tantamount to abandoning the analytical scalpel for a dogmatic bludgeon, especially since a number of highly competent philosophers have seriously questioned the very possibility of a phenomenal language.

The problems of "meaning" and reductionism come into especially sharp focus in the analysis of scientific theories, for here they find expression in that conceptual framework which we use with maximal clarity. For it is in common-sense object talk, its usage refined and molded by years of pragmatic repercussions, that philosopher and layman alike carry on the business of living. And given this everyday "observation language,"

NOTE: This essay owes its existence to the vantage point erected by the philosophical tradition currently known as "logical empiricism." This movement has with increasing penetration and acuity spotlighted the epistemic and ontological problems that underlie the use of theoretical concepts, and with the assistance of the modern renaissance in formal logic, has been developing an ever more powerful conceptual frame with which to attack these problems. The basic issues involved have been set forth with particular clarity by Feigl [5] and Hempel [7], while the reader will also profit from the articles by Carnap [4] and Hempel [8] in the earlier volumes of this series. I also wish to acknowledge my indebtedness to the National Science Foundation for the postdoctoral fellowship during whose tenure this essay was written.

in terms of which most practical (if not philosophical) problems appear to be resolvable, we may ask in a matter-of-fact, commonsensical way: What, if anything, can be said with the theoretical terms of a science which cannot alternately be said in the observation language? Or more or less alternately: Are or are not theoretical statements "about" the same things that observation statements are "about"?

The various answers which have been proposed to questions such as these fall into two main categories. On the one hand we find *positivistic* positions, which hold theoretical terms to be either meaningless computational devices or explicitly definable by observation terms, so that statements using theoretical terms can assert nothing inexpressible in the observation language. In contrast, there are the *realistic* interpretations, which regard the designata of theoretical terms to be (in general) beyond the scope of observational reference, a view which might seem to imply that the factual commitments of a theoretical statement are incapable of expression in the observation language. Each view has its difficulties, the former in that its application to specific cases has met with repeated failure, while the latter flirts with transcendentalism. It is my opinion that, as is so frequently true of philosophical disputes, the insights of both positions are substantially sound. I shall argue that the factual commitments of a scientific theory can be expressed—in existential hypotheses, to be sure—solely in the observation language, but that theoretical terms function in a true theory as *names* of the hypothetical entities and cannot be explicitly defined in the observation language.

The remainder of this introductory section will exhibit in greater detail the problem with which we are here concerned, the presuppositions upon which it rests, and the steps to be taken in search of a solution.

The controversy over the meanings of theoretical terms would seem to be founded on the following presuppositions:

1. There exist in the world certain "particulars" (or "objects," if one wishes to accept the additional commitments of ordinary language). These are differentiated from one another by the "properties" they possess or "classes" to which they belong, while the latter, in turn, variously exemplify or belong to still higher level properties or classes, etc. Particulars, properties, properties of properties (where relations may be regarded as properties of ordered n-tuples) and any other components of reality may collectively be referred to as "entities."

2. Entities in combination constitute "facts."[1] Thus if a is an entity and P is a property of a, a and P are constituents of the fact that $P(a)$—i.e., the fact that a exemplifies P.

3. It is possible for one entity to "designate," "represent," "stand for," "refer to," or be "about" another entity. Designators belong to a wider class of entities known as "symbols," certain compounds of which are also able to designate. In particular, if (a) the symbols $a_1, \ldots, a_n$ (e.g., words) are constituents of a larger symbol S (e.g., a sentence), (b) entities $\alpha_1, \ldots, \alpha_n$ are the constituents of a certain fact f, (c) $a_1, \ldots, a_n$ designate $\alpha_1, \ldots, \alpha_n$, respectively, and (d) S conforms to certain other conditions (such as exemplifying an appropriate formal structure), then S represents the fact f in a way that we shall describe by saying that S *signifies* the fact f.[2]

Presuppositions 1–3 assert merely that there *is* some sort of reality about which we can talk, speculate, and perhaps have knowledge, and that these cognitive events are possible because certain elements of our symbolic processes stand in some sort of referential relation to components of that reality. Since these beliefs manage to subsume virtually all the problems of ontology and epistemology, they can hardly be said to call for no further explication. Nonetheless, there is an important sense in which they are philosophically neutral—*some* such beliefs are presupposed by any

[1] The ontological status of facts has been questioned by some philosophers, especially those of an "ordinary language" turn, who, for reasons which seem to me to be either confused or obscure, are unwilling to countenance "facts" as being among what there is, and indeed, even appear unwilling to grant the term any cognitive significance whatsoever. Since the developments in Sec. II, as they now stand, depend essentially on quantification over fact variables, it should be pointed out that it is formally possible to dispense with facts by replacing them with certain uniquely correlated sets. For example, we may replace the class of facts of form $x \epsilon y$ with the class of ordered pairs of sets such that the first member of the pair is an element of the second. Still another alternative would be to replace "facts" with true statements in a hypothetical omniexpressive metalanguage. In some such fashion, the present analysis could be reworked to arrive at the same conclusions but without depending upon any assumptions about the ontological status of facts. However, the present willingness to quantify over facts is due not merely to the additional difficulties such modification would add to an already complicated story, but even more to the observation that in natural language discussion of such topics as "events," "causes," "phenomena" (in the scientific sense), etc., quantification over fact variables is spontaneous and indispensable. In other words, there are facts, and no theory of semantics can be adequate which does not examine the relationship of sentence to (extralinguistic) fact.

[2] It is tempting to indicate the semantical relation between a sentence S and a fact f signified by S by saying that S "refers" to f. However, this would be misleading, for the relationship that we wish to indicate is a cognitive one, whereas strictly speaking, "reference" is but one of the many uses to which an expression with given cognitive

serious intellectual undertaking. The aim of epistemology and ontology is not to establish them but to clarify and elaborate upon them. Hence we need not feel uneasy, for present purposes, in taking "entity," "fact," "designate," etc., as primitive concepts. In particular, the reader should not try to read more into the present use of 'designate' than is necessary. We here presuppose no particular *analysis* of this concept (though the outline of a behavioral theory of reference will be suggested later), but simply recognize that if it is possible for a statement to represent a fact, there must be some sort of relation between the constituents of the statement and the constituents of the fact. In particular, we need not assume that there is only one kind of "aboutness"—the analysis of '*x* designates *y*' may conceivably differ in important respects according to whether *x* is a primitive term, a compound phrase, a sentence, or some other component of language which may in some sense be said to point out an aspect of reality extrinsic to itself.

One further background assumption will set the stage for the problem at hand. By a "language," let us mean a stock of symbols together with certain principles of usage such that when properly used, some of the symbols ("descriptive" terms) designate various entities, and certain complexes ("sentences") of symbols can be formed which then signify facts whose constituents are designated by the descriptive terms of the sentences. Then we presuppose that languages do exist and that

4. If a person has "observed" an entity e, then he can add to his language a symbol which, when used by that person, designates e.

Just what is meant by saying a person "observes" an entity is difficult to decide. Fortunately, effective use of 'observed' as a primitive concept does not depend upon clarity in its analysis, for so far as everyday language (upon which the philosopher is no less dependent for communication than anyone else) is concerned, this term is used with as much assurance and precision as any other. There is a very important intuitive sense in which we speak of certain facts, in contrast to others, as having been "observed." For science, in particular, the concept of that which is

properties may be put (see fn. 29). Actually, ordinary English usage (which does more to confuse than to clarify the nature of semantical relations) does not seem to yield a satisfactory term for the relation between a statement and that aspect of reality in virtue of which the statement is true or false. Even to say that a statement is *about* a fact, as will sometimes be done here in informal commentary, is to stretch ordinary usage a bit, for we usually (though not always) say that what a statement is "about" is the entities referred to by its constituent descriptive terms.

observed ("data") plays an especially basic role. If we take 'e is an observed entity' to mean that e is a constituent of an observed fact, then this expression should here involve us in no major difficulties.

By 'observation term,' let us mean a term which has been introduced into a language in accordance with presupposition 4. Then the controversy over the meanings of theoretical terms is basically the question of *whether or not a language can contain descriptive terms* (i.e., terms that designate) *which are not observation terms*. Or, phrased somewhat differently: *Is it possible for a symbol to designate an unobserved entity?* Let a fact, all of whose constituents are observed entities, be called an "observational" fact. Then a third formulation of the positivistic-realistic issue is this: *Can a sentence ever signify a nonobservational fact?* Or yet again, if we call a sentence all of whose descriptive terms are observation terms an "observation sentence": *Can sentences be constructed which signify facts, yet which are not observation sentences?* To these questions, the positivist returns an emphatic "No." He by no means necessarily holds that nothing exists which has not been observed—such a view is absurd no matter how one restricts one's ontology. He does insist, however, that only those entities which have been observed can be talked about.[8] The realist, on the other hand, just as emphatically denies that only observational facts can be signified in our language. He not merely admits the existence of unobserved entities but insists that we can and do talk about them. It is important to note that this issue cuts across the question of what *can* be observed. One need not hold, for example, that only sense data are observable, to be a stanch positivist—witness the operationistic movement in contemporary science.)

The difference between these contrasting views emerges with especial clarity when we try to analyze the factual content of scientific theories. It is a well-known and disquieting fact that the most powerful theories invariably contain symbols which are not logical terms, yet apparently refer to no entity which has ever been "observed" in any intuitive sense of this notion. The positivist is forced to hold either that (a) appearances are deceptive and such "theoretical" symbols do, in fact, represent concepts definable wholly in logical and observational terms, or that (b) expressions containing theoretical terms are merely computational devices which are

[8] This principle permeates the writings of Bertrand Russell (e.g. [16], p. 91), although reference to his theory of descriptions [15] is usually necessary to make the thesis explicit. For a more modern statement of the positivistic position, see [2].

no more semantically *about* facts than are calculating machines. The realist, on the other hand, is able to maintain that theoretical terms may designate existent but hitherto unobserved entities, and that scientific hypotheses containing these terms may simply signify certain facts which happen to be as yet nonobservational. The realistic position is a seductive one, but is incompatible with an empirical epistemology unless it is possible to show how, given an observation-based language, a person might acquire additional terms which designate hereto unobserved entities even though his scope of observation remains unchanged. One result of the present analysis is to suggest how this might come about and, correlatively, the limits of such a language enrichment.

Section I will attempt to formalize the concept of (scientific) "theory." The analysis will be idealized in that we presuppose the theory user to have at his command a fully formalizable observation language, all descriptive terms of which designate observed entities (where 'observed' is to be understood in any appropriately broad or narrow sense). It will be heuristically helpful to regard this observation language as an idealized version of the observation language we use in everyday life. In Section II, we turn to the problem of the "factual content" of a theory. We shall be able to determine this without first prejudging whether or not the theory is itself an assertion, though not without making certain general assumptions about the nature of semantical relations and the way in which theories are actually used. In Section III we shall explore the *semantical* status of theories, and conclude that under suitable circumstances, theoretical terms do, in fact, designate unobserved entities. Finally, Section IV considers briefly the implications of this analysis for several long-standing philosophical problems.

Let us conclude this introduction with some needed semantical preliminaries. While it is all very well to undertake analysis of the possible semantical properties of theories, such an effort is especially handicapped by lack of a well-understood and generally accepted theory of semantical relations upon which the analysis can draw. In particular, classical semantics, as formulated most explicitly in the work of Tarski and Carnap, does not adequately deal with the relations between cognitively meaningful sentences and extralinguistic reality (see the next paragraph). Hence the present essay labors under the double burden of developing a semantical theory even as it argues for the meaningfulness of theoretical expressions. While suggested postulates for a generalized theory of semantics will be found in Sections II and III, it is highly advisable to advance a few preliminary considerations at this point.

Any comprehensive theory of semantics must come to grips with (a) the manner in which the semantical properties of a sentence derive from those of its constituent terms, (b) the extralinguistic designata of sentences—i.e., what aspects of reality sentences themselves are about, over and above the referents of their constituent terms, and (c) the truth conditions of sentences. Of these, Tarski's [19] famous schema "*S* is true if and only if it is the case that p," where *S* is a meaningful sentence whose metalinguistic translation is 'p,' concerns only the third. In Carnap's [3] more complete theory, a sentence *S* designates a "proposition" or "state-of-affairs" *p* (which might be put more idiomatically by saying "*S* asserts that p") when the descriptive terms in *S* designate the constituents of *p*, and *S* and *p* show similar composition. Sentence *S* is then said to be true if and only if there is a proposition *p* such that *S* designates *p* and *p* is the case.

Unfortunately, this formulation is still not satisfactory for present purposes. To begin with, there is the problem of the ontological status of propositions. These cannot be identified with facts, for propositions are true or false—i.e., are or are not the case—whereas facts are what *determine* the truth values of propositions. Neither can we identify a *true* proposition as a fact, for then we have no way to cope with *false* propositions—to say that a false proposition is a possible but not actual fact is to propose a strange ontology in which nonexistence is a category of Being. By far the most satisfactory interpretation of propositions is to regard them as the *meanings*, or *senses*, of sentences—i.e., those aspects of the linguistic process through which sentences are able to make contact with an external reality (see Section III; also [13]). But if so, it is then incorrect to say that a sentence *designates* a proposition; for "designation" is the relation of aboutness between linguistic and extralinguistic entities, whereas a word or sentence *expresses* (i.e., produces, evokes, has) a meaning *in virtue of which* it may designate something else.[4] Hence to analyze the semantical properties of a sentence merely in terms of "propositions" is to leave unexamined the manner in which sentences communicate with the facts that determine their truth.

Now it might be thought that statements of form "*S* expresses proposi-

[4] Ordinary language is in agreement (for what this is worth) that a sentence expresses, not designates, a proposition.

tion p," "*S* is true if and only if it is the case that p," or "*S* asserts that p," though unable to deal with the designative properties of sentences, might still suffice to determine the truth conditions and hence the factual content of meaningful expressions. Within limits which need not detain us, this is true; however, the difficulty for present purposes is how such statements are to be obtained. For expressions translatable into our metalanguage, the matter is fairly straightforward: If S is a sentence under analysis whose translation is 'p,' there is surely nothing amiss about accepting the metalinguistic statement "*S* is true if and only if it is the case that p," or even the stronger claim "*S* asserts that p." However, our major concern here is with the *problematic* semantical status of theoretical sentences, and to presuppose their translatability into the metalanguage would simply be to beg the whole issue at the outset. We shall indeed attempt to arrive eventually at a sentence schema of the form "Theory *T* is true if and only if it is the case that p," but for reasons which need not be explored here, it seems possible to reach such a conclusion only through analysis of the semantical relations which may obtain between sentences and those aspects of reality which determine their truth values, namely, *facts*, not merely between sentences and their *meanings*, i.e., propositions.

When we abandon propositions in favor of facts as the designata of sentences, however, a complication arises. If it is correct to say that sentence *S* asserts that *p*, and it is a fact that *p*, then it seems unobjectionable to conclude that what S designates, or signifies, is the fact that *p*. But what shall we say when *S* asserts that *p*, but it is the case that $\sim p$? What we *cannot* say is that *S* signifies the fact that *p*, for there is no such fact. In this instance, however, *S* stands in an especially intimate relation to the fact that $\sim p$, for just as *S* is true in virtue of p when it is the case that *p*, it would seem that S is false in virtue of the fact that $\sim p$ when it is not the case that *p*. Apparently we need to admit two kinds of semantical relations between sentences and facts; one for sentences which make true assertions and another for sentences which make false assertions. More generally, it follows from the assumptions of classical semantics that for each cognitively meaningful sentence *S*, there is a fact *f* whose constituents are designated by the descriptive terms of *S* and which determines the truth value of *S*. If *S* is true in virtue of such an *f*, we shall say that *S* *signifies f truly*, whereas if S is false in virtue of *f*, then *S signifies f falsely*. For example, under the classical assumption that a sentence 'P(a)' in which 'P' designates the property *P* and 'a' designates the individual a, is true if it is the case that $P(a)$ and false if it is the case that $\sim P(a)$, it would follow in the first case that 'P(a)' truly signifies the fact that $P(a)$, and in the second case that 'P(a)' falsely signifies the fact that $\sim P(a)$. Under classical theory, then, each cognitively meaningful sentence signifies, truly or falsely, exactly one fact, namely, the fact that p when the sentence asserts that p and it is the case that p, or the fact that $\sim p$ when the sentence asserts that p and it is not the case that p. There will later be occasion to question portions of the classical view. Nonetheless, the concepts of true and false signification, as roughly sketched here and defined more precisely in Section II, should allow the reader to pass without undue intuitive strain from the more familiar notion of what a sentence expresses or asserts to the needed appreciation of semantical relations between sentences and facts.

I

If we are to determine the factual content of scientific theories, it is first necessary to decide what we mean by a "theory." If we restrict "theory" to "hypothesis formulated in the observation language," we have, of course, cut ourselves off from our problem. On the other hand, "hypothesis (or statement) in a theoretical language" poses difficulties. For in what sense is a "theoretical language" entitled to be called a *language*? It is not sufficient for a string of signs to conform to certain topographical characteristics in order for it to be a "statement," for in its normal usage, this term implies that the sign complex has *meaning*. So long as the meanings of theoretical terms are in question, we are not entitled to call the expressions in which they occur "statements," "hypotheses," or other similar concepts which presuppose a certain semantical status for their subjects. Thus we must find an identifying feature of theories which does not prejudge their meaning content.

It seems quite plain, in the final analysis, that the ultimate purpose of a theory is cash-value prediction—i.e., to assist anticipation of the truth values of observation sentences,[5] given the truth values of other observation sentences. Thus whatever else a theory may be, it is *at least* a tech-

[5] For simplicity, we shall use "sentence" in the sense of "cognitively meaningful declarative sentence," or "statement." It is important to note that an "observation sentence," as defined above (p. 277), may include logical terms, hence permitting molecular and generalized sentences. Therefore, a fact signified by an observation sentence is not necessarily "observable" in the sense in which this expression is frequently used. For example, if 'R(x,y)' is a dyadic observational predicate, the observation sentence '(x)(∃y)R(x,y)' cannot, except in special cases, be either verified or refuted by any

nique by which some observation sentences are transformed into others. This confronts us with an interesting problem area which does not seem to have been previously explored: Given certain requirements such as consistency, is any *kind* of transformation which (partially) maps the domain of observation sentences into itself acceptable, prior to empirical evaluation, as a legitimate scientific theory? For example, suppose we had (a) a set of rules for generating geometric figures from one another (e.g., "x derives from y and z if x results from the superimposition of y upon z") and (b) a set of "coordinating definitions" setting up a (not necessarily exhaustive) pairing of sentences and geometric figures. Could a transformation technique based upon such a system *in principle* be construed as a theory? If not, on what grounds do we reject it? On the other hand, if a transformation such as this counts as a theory, in what sense can a theory be regarded as an *assertion?*

It would be too lengthy a digression from the main purpose of this paper to explore further the general concept of *theory as transformation* at this time. However, those theories which have actually seen application to human affairs appear to be of a special kind which I shall refer to as "normal syntactic" theories. Such a theory is identified by (1) a set of "inference rules" which are applicable to (but not only to) observation sentences, and when so applied, yield valid deductions; and (2) a set, *K*, of sentencelike sign complexes, or "theoretical postulates," such that application of the inference rules to the union of *K* and a set of observation sentences, O, yields another set of observation sentences. By "sentencelike" I mean sign complexes which are syntactically similar to observation sentences in such a way that if certain components (the "theoretical" terms) of the theoretical postulates were to possess designative meaning in the same way that descriptive terms of the observation language have designative meaning, the theoretical postulates would themselves be mean-

finite set of observations. On the other hand, the meaning of 'observable' is difficult to pin down. In what sense, for example, is the fact signified by a statement about the current number of coconuts on an uninhabited atoll *observable*? Presumably, because if I were there, I would be able to observe how many there are. But how does this differ in kind from saying that if I were acquainted with all pairs of objects, I would be able to observe whether or not $(x)(\exists y)R(x,y)$? To be sure, we believe it to be *physically* impossible for me to observe all pairs of objects, but it is likewise physically impossible for me to be at some spatial position other than where I am, for (presumably) I am where I am because of physical laws controlling the motions of material bodies. If the issue were germane to present purposes, I would argue that the only tenable analysis that can be given to '*f* is an observable fact' is something with roughly the force of 'There is an observation sentence which signifies *f*.'

ingful sentences. There is thus the possibility that a normal syntactic theory is not *merely* a transformation, but that perhaps its theoretical terms do, in fact, designate; in which case the theoretical postulates semantically *express* a hypothesis, and the observation sentences which are derived from the theory are the logical consequences of that hypothesis.

(Syntactic theories of a more general kind would comprise theoretical postulates not necessarily syntactically isomorphic to observation sentences, and, perhaps, inference rules not validly applicable to observation sentences. However, there is then no reason for the theory user to think that the postulates of a nonnormal syntactic theory might themselves signify facts. The present analysis will be restricted to those theories where there *is* good reason to suspect that the theory may be more than just a transformation technique—i.e., the normal syntactic case—although the more general case is certainly of philosophical interest and in fact opens some rather exciting epistemic possibilities.)

For a completely general account of normal syntactic theories, we would have to discuss a wide variety of observation languages. Fortunately, our main point of departure, Theorem 2, can be established with a minimum set of stipulations about the syntax of the observation language which, moreover, would presumably be satisfied by any satisfactory formalization of the language we in fact use in science and everyday life. As consequences of the formation rules we require:

1. Sentences are finite concatenations of certain syntactically primitive symbols, where the latter are of three kinds: (a) *logical constants*, including the truth-functional connectives and existential quantification; (b) *descriptive constants*, which designate specific entities; and (c) *variables*, each of which ranges over entities of a specific kind.

2. All primitive symbols (the logical constants may be excluded if desired, since their inclusion here is trivial) are effectively classifiable according to "formal type," so that each symbol is of exactly one formal type, and each formal type i specifies a class of entities C^i which is the range of every variable of type i and contains all entities designated by constants of type i. (We could also allow a given term to be of more than one type, but this can be reduced to the first case.)

3. Let L_o be the observation language under consideration and L_M the metalanguage in which the present analysis is being conducted—or better, let L_M be the language in which this discussion *would* be conducted were its syntax fully formalized. Then we stipulate that expressions in L_o are

translatable into L_M. That is, the descriptive constants, variables, formal types, logical constants, and concatenation arrangements of L_o correspond distinctly to descriptive constants, variables, formal types, logical constants, and concatenation arrangements of L_M in such a way that if S_o is a sentence in L_o and S_M is the sentence in L_M formed by placing those symbols of L_M which correspond to the primitive symbols in S_o in the concatenation arrangement of L_M which corresponds to that of S_o, S_M is true in L_M if and only if S_o is true in L_o. Although this stipulation concerns more than just the syntax of L_o, it economically characterizes the latter as being isomorphic to part of the syntax of a formalized English. Since we are not concerned with the physical topography of expressions in L_o as such, we may for convenience identify expressions in L_o with their translations in L_M, thus allowing us to write expressions in L_o in the standard logical notation of formalized English (e.g., taking '$\exists$' as the existential operator in L_o, etc.). This will also allow us, in discussion of the semantical properties of L_o, to use expressions of L_o as well as *mention* them (e.g., "The sentence '$P(a)$' of L_o is true only if it is the case that $P(a)$"). Granting that the observational portion of a formalized English could be made to satisfy stipulations 1, 2, and 4, stipulation 3 is then trivially satisfied by the observational basis of English as well as by any other sufficiently similar observation language.

4. If a symbol complex of the form '$P(a^i)$' is a sentence in L_o, where 'a^i' is a syntactically *primitive* descriptive constant[6] of formal type *i* and 'P' abbreviates a simple or complex predicate, then there is also a corresponding existentially quantified sentence in L_o of the form '$(\exists\phi^i)P(\phi^i)$,' where 'ϕ^i' is a variable of type *i*.[7] Further, if 'S_i' and 'S_j' are sentences in L_o, then '$\sim S_i$' and '$S_i \cdot S_j$' are also sentences in L_o. (Hence we may also assume that L_o contains '$(\phi^i)P(\phi^i)$,' '$S_i \supset S_j$,' etc.)

(Conditions 2 and 4 call for some further comment. First of all, while the languages with which we are concerned admit abstract entities, this is not a result of stipulation 4 but of presupposition 1 of the introductory

[6] No notion of a "descriptive constant" which is *not* syntactically primitive has been or will be explicitly invoked here. However, the term 'constant' is sometimes applied to certain syntactically complex expressions (e.g., compound predicates and definite descriptions) as well as to primitive terms, and the need hence arises (in view of my arguments in [13]) to make clear that stipulation 4 authorizes quantification only over primitive terms in L_o.

[7] This stipulation can be weakened without affecting the argument to follow, so long as theoretical terms are limited to formal types over which existential generalization is authorized.

section. I would agree with Bergmann [1], *contra* Quine, that the ontological commitments of a language are made by its primitive descriptive constants rather than by its variables, though to be sure these commitments are *revealed* by its variables.[8] In fact, I have never understood the nominalistic thesis that only particulars exist, unless "exist" here means something like "possess space-time positions," in which case the thesis is trivial. If, as other possibilities, the nominalist wishes primarily to challenge that properties can exist without being exemplified, or that every predicate has a referent, he need not commit himself to the view that no abstract entities exist. The use of predicate variables in no way necessitates that every predicate expression constructable in the language must be a substitution instance of a predicate variable, or must be assumed to designate an abstract entity (see [13]).

(With respect to stipulation 2, it should be observed that in letting the formal type of a variable specify its range, we have not stipulated that these types are necessarily the syntactical representations of purely "logical" categories. That is, we need not assume that variables are wholly a part of the logical framework of language. For example, the variable 'x^s' might range only over the class of swans, in which case '$(\exists x^s)$ (x^s is blue)' would be true if and only if there is a blue swan. In particular, predicate variables need not be construed to range over the totality of properties or classes of a given Russellian type, especially if it be maintained (as I do not) that to every cognitively meaningful predicate, no matter how complex logically, there corresponds a property or class. The possibility that some variables may be nonlogical terms allows for the possibility of "theoretical" variables. However, the explicit definition of 'theoretical postulate,' below, upon which Theorems 2–5 are based, assumes that theoretical terms function syntactically in the theoretical postulates only as constants. The extent to which this attenuates the present analysis will be discussed at the end of Section II.

(Finally, a word about the status of definite descriptions, if any, in L_0: While contemporary philosophers have yet to reach general agreement

[8] Surely Quine is correct when he proposes (e.g., [10] that whether or not a term 'A' in sentence 'F(A)' can be construed to designate anything is to be tested by judging whether or not the corresponding existential generalization, '$(\exists\phi)F(\phi)$,' makes sense. But if such an existential generalization *would* make sense if we made it, we do not nullify the possibility that 'A' carries ontological commitments merely by legislating that '$(\exists\phi)F(\phi)$' is not to be considered a sentence of the language. Otherwise, we could rid philosophy of ontological problems altogether simply by refusing to countenance the use of any quantifiers.

on the syntactic and semantic status of definite descriptions, two main alternatives would seem to be available. Either (a) descriptions are genuine *referring* expressions which play the same syntactical role as do proper names of similar formal type, or (b) sentences containing descriptions are abbreviations for more complex statements which do *not* contain descriptions (see [15]). In the latter case, descriptions are notational conveniences which are not strictly part of the formalized language—or, more precisely, can be ignored insomuch as anything which needs be said about the syntactic or semantic status of a sentence containing a description (its deductive consequences, truth conditions, etc.) is already given by the corresponding statement about the sentence for which it is an abbreviation. On the other hand, if a definite description functions referentially in a language for which stipulations 1 and 4 hold, the description must be classed by the rules of the language as a primitive term, for if a description *refers* to an entity which satisfies it, quantification over only a part of the description yields nonsense. Hence even if L_o is construed to contain definite descriptions, their syntactic and semantic properties are already covered by the rules for primitive descriptive constants in L_o In particular, to the extent that definite descriptions are considered part of the *observation* language, they do *not* provide a way to refer to unobserved entities, for by definition all primitive descriptive terms of L_o have observed referents. In Section IV it will be argued that definite descriptions are most properly regarded as a form of theoretical term.)

On the basis of stipulations 1–4, the theoretical postulates of a normal syntactic theory may now be characterized more precisely as certain sign complexes which could be generated under the formation rules of the observation language if its primitive descriptive terms were augmented by a set of formally typed but otherwise uninterpreted constants. Let a descriptive constant of the observation language, whatever its formal type, be represented by the notation 'c_i, where the subscript is indexical. Similarly, let an uninterpreted but typed sign be written 'τ_i.' (For notational completeness, a superscript indicating formal type should be added, but this turns out to be unnecessarily cumbersome. We will be able to omit explicit reference to type so long as it is remembered that the members of a given set of terms are not necessarily of the same formal type unless so stipulated.) Application of the formation rules of the observation language to the observational constants augmented by the uninterpreted constants, 'τ_1,' . . ., 'τ_n,' then generates syntactically well-formed, sentencelike formulas of form '$S(c_1, \ldots, c_m, \tau_1 \ldots, \tau_n)$' ($m \geqq 0$), where '$S(\ )$' is a sentential matrix containing only logical terms (and perhaps observation-language variables, if these be non-logical) such that if the 'τ_i' were replaced by a set of observational constants, 'c^*_1,' . . ., 'c^*_n,' of corresponding types, the resulting expression, '$S(c_1, \ldots, c_m, c^*_1, \ldots, c^*_n)$,' would be an observation sentence.

For simplicity, we make the further assumption that our observation language L_o is "complete" in that the theorems of the language include all sentences which are formally valid. By "theorem," we mean a sentence which can be effectively deduced by the formal (i.e., syntactical) inference rules of the language from any other sentence. 'Formally valid' may be defined as follows: By a "model" of a language *L*, we mean any arbitrary assignment of ranges to the variables and designata to the descriptive constants of *L*, subject only to the restrictions (a) that to each formal type *i* is correlated a nonempty class of entities, C^i, such that C^i is the range of each variable of type *i* and contains all designata of descriptive constants of type *i*; and (b) also, perhaps, additional restrictions on the ranges assigned to the various variables.[9] Then each model of *L* assigns a specific truth value to each sentence of *L*, which may be different under different models. For example, '$P^j(a^i)$' is true under a given model if and only if the designatum assigned to 'P^j' is a property of (or, if the entity assigned to 'P^j' is a class, contains) the designatum assigned to 'a^i'; and '$(\exists x^i)P^j(x^i)$' is true if and only if the designatum assigned to 'P^j' is a property of (or contains) at least one of the entities in the range assigned to 'x^i.' We then define the "formally valid" sentences of *L* to be those which are true in

[9] The various alternative restrictions that can be placed on the ranges assigned by a model to the variables of *L* generate a whole family of concepts of "formal validity," several of which appear in the technical literature of formal logic. (Just how such formal concepts relate to the philosophical notion of "logical truth" is not at all easy to decide.) Fortunately, it is here unnecessary to be explicit about these additional restrictions, for our present concern with "formal validity" is only to characterize deducibility in L_o, and Lemma 1 is unaffected by any restrictions that may be placed on the ranges assigned by a model to the variables of *L*. This indifference of Lemma 1 to the ranges of the variables is rather convenient, for it thus becomes unnecessary to make any commitments here as to possible *logical* restrictions on the designata of descriptive terms of various types. In particular, we need not take a stand on whether a primitive predicate necessarily designates a property, a class, or either.

Actually, the present "definition" of 'model of *L*' is more a heurism than literally correct. A model does not literally assign designata to the descriptive terms of a language, for these already have referents determined by their meanings. More accurately, "model" should be understood as a purely formal concept having to do with the mapping of one domain into another.

all models of L. As is customary—and also necessary if, as assumed later, formal deducibility is to reflect a necessity relation among the truth values of sentences in L—we restrict the formal inference rules of a language to some subset of those transformations which yield only valid deductions; that is, rules such that if deduction of S_j from S_i is authorized, then S_j must be true in all models of L in which S_i is true. This restriction, together with the completeness assumption, entails that the formal inference rules (some of which may be formalized as "axioms" or "axiom schemata") of observation language L_o yield as theorems exactly those sentences in L_o which are formally valid. Recent discoveries in formal logic show this to be a reasonable assumption [see 9]. The completeness assumption is merely for convenience, however, and could be arbitrarily weakened without affecting the validity of Theorems 3–5 (see footnote 11).

For comprehensiveness, let us take as the inference rules of a normal syntactic theory all formal inference rules of the observation language (however, see footnote 11). Then each normal syntactic theory is uniquely characterized by a specified *finite* set of sentencelike formulas—the "theoretical postulates"—of the form '$S(c_1, \ldots, c_m, \tau_1, \ldots, \tau_n)$' as defined above. (The set of theoretical postulates must be finite if the theory is literally to be *believable*—axiom schemata, which are occasionally interpreted as infinite sets of axioms, and which, so construed, would entail an infinite set of theoretical postulates, are most satisfactorily conceived as rules of inference.) More specifically, we shall understand a normal syntactic theory to include in its set of postulates all formally independent sentencelike formulas containing theoretical terms, including "correspondence rules," which are in force when the theory is under consideration.[10] (The set of theoretical postulates may also include sentences wholly in the observation language, indeed, must do so under the definition of 'theory' just offered if there would otherwise be sentences with theoretical terms deducible from the theory plus extraneous observation-language postulates but not deducible from the theory alone.) Obvious-

[10] Actually, this may be stronger than necessary. In order to see what is involved in accepting a given theoretical formula '$S(\tau_i)$' containing the theoretical term 'τ_i,' we must determine the pragmatic force of 'τ_i' under the particular circumstances involved, whether this force be characterizable as a genuine cognitive meaning or only as that of part of a transformation mechanism. But while the force of 'τ_i,' and hence that of '$S(\tau_i)$,' will be determined by the role of 'τ_i' not only in '$S(\tau_i)$' but also in other accepted formulas in which it occurs, it does not seem to follow that in order to evaluate the (cognitive or noncognitive) significance of 'τ_i' and '$S(\tau_i)$' it is necessary to con-

ly, a set of theoretical postulates has exactly the same implicative force as the single postulate formed by logical conjunction of the members of the set. Therefore, the postulates of a normal syntactic theory may be written as a single, finite, sentencelike formula, '$T(\tau_1, \ldots, \tau_n)$,' in which the sentential matrix, '$T(\quad)$,' contains only logical and observational terms, and 'τ_1,' . . ., 'τ_n' are theoretical constants. It should be noted that the predicate '$(\lambda\phi_1, \ldots, \phi_n)T(\phi_1, \ldots, \phi_n)$,' henceforth written simply '$T(\phi_1, \ldots, \phi_n)$,' which is ascribed by a normal syntactic theory to its theoretical terms, is an expression constructable wholly within the observation language prior to any consideration of the theory.

It is convenient at this point to introduce the convention that if 'S' abbreviates what is construed as the principal predicate in a sentence of L_o (or, as a degenerate case, if 'S' abbreviates the sentence itself), then '*S*' is the name of that sentence in the metalanguage. Thus 'S_i' is a name for '$S_i(c_1, \ldots, c_n)$.' Similarly, '*T*' refers to the theory '$T(\tau_1, \ldots, \tau_n)$.' Further, if S_i and S_j are sentences, '$S_i \cdot S_j$' is the name of the conjunction of S_i and S_j, and similarly for the other connectives. Thus we will have two alternative ways of referring to theoretical expressions and sentences in L_o; the customary procedure of putting the sentence itself in quotes, and also, when brevity is in order, italic notations. (The particular notational convention here described was adopted in response to a last-minute discovery that boldface type could not be used. Unfortunately, since metalanguage transcriptions of object-language expressions have also been italicized, there is now a certain ambiguity between the use of italics to name object-language expressions on the one hand, and to *translate* them on the other. In most instances, the context makes unhesitatingly clear what interpretation is intended; however, a mild but regrettable confusion does tend to arise in some passages discussing the relevance of certain facts $T(t_1, \ldots, t_n)$ or $\sim T(t_1, \ldots, t_n)$ to theory T.)

A normal syntactic theory is a transformation in that, in general, one or more observation sentences S_j may be deduced, through use of the inference rules of the theory, jointly from T and some observation sentence S_i

sider *all* theoretical postulates accepted at the time. In particular, if the total set of accepted postulates can be split up into subsets for which it can be argued that the use of each subset is totally independent of the others, it would then seem that each should be regarded as a separate theory. Just what this "total independence" of usage might consist of, however, is a question which is not easily resolved. Fortunately, the present definition of "theory" to include *all* the accepted theoretical assumptions does not preclude the possible autonomy of certain subsets of its postulates, and as will be seen, suitable allowance for this contingency is made in the ensuing development.

where S_j is not deducible from S_i alone. When S_j is deducible from T alone, we may speak of S_j as a "consequence" of T. More generally, to include the case where T is not normal syntactic, we may define 'O-consequence of T' (where the 'O' is to distinguish the observational consequences of T from expressions derivable from T containing theoretical terms) by

Definition 1. C is an O-consequence of theory T (in language L) $=_{\text{def}}$ *C is an observation sentence (of L), and for any observation sentence S (of L), T transforms S into C.*[11]

While the concept of "O-consequence," so defined, is relative to a particular language L, it will henceforth be presupposed that L is the language which results when theory T is accepted by a person whose observation language is L_o. The extent to which L differs from L_o depends, of course, upon whether or not adoption of theory T effects a genuine language enrichment.

Since we have taken the inference rules for a normal syntactic theory to be sufficiently complete, a normal syntactic theory, T, transforms S_i into S_j if and only if $S_i \supset S_j$ is an O-consequence of T. Therefore, the total syntactical force of a normal syntactic theory for observation sentences is represented by the set of its O-consequences.

II

What does it mean to "believe," "accept," or "entertain" a theory T? We feel tempted to answer that it is *at least* to believe, accept, or entertain all the O-consequences of T. This will not quite do as it stands, however, for in general, T will have an infinite number of O-consequences, and it is unlikely that an infinite set of propositions can be entertained by a human mind. Hence, it is safer to say that to accept a theory is to be (at least) *committed* to its O-consequences—i.e., to be in a state such that belief that S is an O-consequence of T is sufficient cause for belief that S. More generally, we should say that to accept a theory is to be in a state such that belief that S_i and that T transforms S_i into S_j necessitates

[11] If T were here limited to normal syntactic theories, we could instead adopt: *C is an O-consequence of T* $=_{\text{def}}$ *C is an observation sentence and T formally entails C (i.e., T ⊃ C is formally valid)*. With similar replacement of 'is deducible from' by 'is formally entailed by' in Definition 2, Theorems 1–5 then follow without any assumptions whatsoever about the inference rules of L. (However, it must then also be argued—as indeed it may—that to accept T is to be committed to all observation sentences formally entailed by T whether they are also deducible from T or not.)

belief that S_j, but this may be reduced to commitment to the O-consequences of *T*, since commitment to $S_i \supset S_j$ and belief that S_i necessitate commitment to belief that S_j.

These formulations, however, raise a further problem: In what sense can *commitment* to the O-consequences of *T* be assimilated to *cognitive* processes? For to say that if circumstances are such-and-such, then a person will believe so-and-so, is to describe a *disposition* to acquire certain beliefs, a state which, unless additional conditions are also satisfied, would not ordinarily be regarded as a form of knowledge. Two possibilities arise, beyond which interpreting an accepted theory as itself a form of (possible) knowledge seems highly tenuous. (a) The theory, in itself, may be cognitively meaningful. This is not implausible in the case of a normal syntactic theory, since '$T(\tau_1, \ldots, \tau_n)$' will be an assertion if the 'τ_i' have appropriate semantical properties, but is questionable for other forms of theories, if other transformation techniques may be so designated. (b) The theory may have among its O-consequences a finite subset which entail the remainder. For convenience, let us call the conjunction of the sentences in such a subset a "prime consequence" of the theory:

Definition 2. C is a prime consequence of theory T $=_{\text{def}}$ *C is an O-consequence of T, and for any sentence S, if S is an O-consequence of T, S is deducible from C.*

If a theory does not itself make an assertion, the best candidate for the cognitive content of the theory would seem to be what is asserted by a prime consequence of the theory. It should be noted that it does no harm, as a figure of speech, to speak of *the* prime consequence of a theory, if it has one, for if both C_1 and C_2 are prime consequences of the theory, C_1 formally entails C_2 and conversely. That is,

Theorem 1. All prime consequences of a theory are formally equivalent.

We now prove a lemma, of great importance in the formal theory of quantification, which holds for any language in which '$\supset$' and '$\exists$' have the customary interpretation. Note that the lemma is *not* restricted to languages which are translatable into our present metalanguage, nor do we need a definition of 'formally valid' more precise than that provided in Section I.

Lemma 1. If '$A(a^i) \supset B$' *and* '$(\exists\phi^i)A(\phi^i) \supset B$' *are sentences in a language L (where* 'a^i' *and* 'ϕ^i' *are a primitive constant and a variable, respectively, of formal type i, and* 'A' *and* 'B' *abbreviate, respectively, a predicate* *and a sentence neither of which contain '*a^1*'), and '*$A(a^1) \supset B$*' is formally valid in L, then '*$(\exists\phi^1)A(\phi^1) \supset B$*' is also formally valid in L.*

Proof: By definition, a sentence is formally valid in L if and only if it is true in all models of L. Now, the sentence '$(\exists\phi^1)A(\phi^1)$' is true in a given model if and only if there is an entity in the range assigned to 'ϕ^1' which possesses the property, or belongs to the class, assigned to 'A.' (If 'A' is a complex predicate, the property or class assigned to it is a function of the assignments made to its constituent terms.) Hence if there exists a model, M_i in which '$(\exists\phi^1)A(\phi^1)$' is true, then there exists a model M_j, differing from M_i in at most the assignment to 'a^1,' in which '$A(a^1)$' is true—we simply form M_j from M_i by assigning to 'a^1' one of the entities which possess the property, or belong to the class, assigned to 'A' by M_i. (Since 'a^1' is primitive, we are free to do this, and since 'A' does not contain 'a^1,' the latter may be reassigned without changing the assignment to the former.) Consider, now, the set Σ_k of all possible models of L given a fixed assignment of designata to the descriptive constants and ranges to the variables in 'B.' 'B' must be either true in all models in Σ_k or false in all. (1) If 'B' is true in the models in Σ_k, then '$(\exists\phi^1)A(\phi^1) \supset B$' is obviously true in all models in Σ_k. (2) If 'B' is false in the models in Σ_k, then '$A(a^1)$' must also be false in all models in Σ_k, since by hypothesis, '$A(a^1) \supset B$' is true in all models of L. But then '$(\exists\phi^1)A(\phi^1)$' must also be false in all models in Σ_k; for if '$(\exists\phi^1)A(\phi^1)$' were true in a model M_i in Σ_k, then as just shown there would also be a model M_j, differing from M_i in at most the assignment to 'a^1' and hence also in Σ_k (since reassignment of 'a^1' does not affect the assignment to 'B'), in which '$A(a^1)$' were true. Thus '$(\exists\phi^1)A(\phi^1) \supset B$' is true in all models in Σ_k. But Σ_k is the set of all models of L given any particular (permissible) fixed assignment of designata and ranges to the terms in 'B'; since these sets jointly exhaust the models of L, '$(\exists\phi^1)A(\phi^1) \supset B$' must be true in all models of L and is hence formally valid in L. Q.E.D.

Theorem 2. Every normal syntactic theory has a prime consequence.

Proof: Let the sentencelike formula '$T(\tau_1, \ldots, \tau_n)$' be the conjunction of the postulates of a normal syntactic theory. Then by existential generalization over the theoretical constants, we obtain the Ramsey sentence[12] '$(\exists\phi_1, \ldots, \phi_n)T(\phi_1, \ldots, \phi_n)$,' which may be designated by

[12] So named after F. P. Ramsey ([11], pp. 212–215, 231), who first called attention to this construction.

'R_T.' (It is to be understood, of course, that each variable 'ϕ_i' agrees in formal type with the corresponding 'τ_i.') Since R_T contains only terms in L_o, and our earlier stipulations about the formal properties of L_o ensure that R_T is deducible from T, R_T is an O-consequence of T. To see that R_T is also a *prime* consequence of T, we consider the formal properties of the calculus L'_o formed from the observation language L_o by adding the theoretical constants 'τ_1,' . . ., 'τ_n' to the descriptive constants in L_o. Let 'C' be a sentence in L_o which is deducible from T. Then by the definition of 'normal syntactic theory,' '$T(\tau_1, \ldots, \tau_n) \supset C$' must be formally valid in L'_o; and since 'C' and '$T(\phi_1, \ldots, \phi_n)$' contain no theoretical terms, n applications of Lemma 1 shows that '$(\exists\phi_1, \ldots, \phi_n)T(\phi_1, \ldots, \phi_n) \supset C$' must likewise be formally valid in L'_o. But the latter formula contains no theoretical terms and is thus also a formally valid sentence in L_o. Hence '$(\exists\phi_1, \ldots, \phi_n)T(\phi_1, \ldots, \phi_n) \supset C$' must be a theorem of L_o; and since 'C' is any O-consequence of T, it follows that any O-consequence of T may be deduced, in L_o, from the Ramsey sentence of T.

Corollary. The prime consequence of a normal syntactic theory is its Ramsey sentence.

It should be intuitively apparent that the prime consequence of a theory must stand in a special relation to its factual content. We saw above that while there may be problems in interpreting a theory as itself constituting a knowledge claim, the most obvious alternative, that the cognitive content of a theory is the (infinite) set of its O-consequences, also meets with difficulty. The horns of this burgeoning dilemma are blunted, however, by Theorem 2, which shows that all normal syntactic theories, in which class presumably fall all theories which have actually been entertained by scientists or have been of philosophical concern, possess a prime consequence. For not only is the prime consequence of a theory a straightforward (albeit existential) hypothesis in the observation language, thus posing no more philosophical difficulties than any other quantified statement in L_o, it has also exactly the same O-consequences as does the theory. Hence the conclusion that the factual content of a theory is the same as that of its prime consequence is a most seductive one. To show that this conclusion is correct as well as seductive will be the burden of the remainder of this section. However, it will also be brought out, especially in Section III, that the relation between a theory and its prime consequence is by no means a matter of simple synonymy.

Since the detailed discussion and proofs of the theorems which terminate this section are somewhat involved, it will be helpful to develop these theorems intuitively before turning to a more rigorous analysis. The fundamental assumption concerning the semantical status of theories, on which the remainder of this article rests, is that the theoretical terms 'τ_1,' . . ., 'τ_n' *derive* their meanings, if any, from their occurrence in the sentential function '$T(\phi_1, \ldots, \phi_n)$' which they complete to form the theory '$T(\tau_1, \ldots, \tau_n)$' (see the Thesis of Semantic Empiricism, below). Now if it is indeed true that theoretical terms are *given* their significance by the theory in which they are imbedded, it follows that so long as distinctiveness is preserved, the theoretical terms in theory T may be exchanged for any other set of theoretical terms without altering the meaning or factual commitments of the theory. Moreover, it is intuitively convincing, and indeed can be proved with a modicum of assumptions unrelated to the question of theoretical meaningfulness (see [14]), that two theories which have no theoretical terms in common are incompatible if and only if they have incompatible observational consequences. But if T_1 and T_2 have theoretical terms in common—that is, if T_1 and T_2 make use of common theoretical sign-designs (though the meanings given to the common terms may be different in the two cases)—it follows from the fundamental assumption already noted that we can replace the theoretical terms in T_2 with new theoretical terms in such a way that the resulting theory T^*_2 is equivalent to T_2 in meaning and has no theoretical terms in common with T_1. Therefore, T^*_2, and hence T_2, is incompatible with T_1 if and only if T^*_2, and hence T_2, has observational consequences which are incompatible with those of T_1. That is, two theories, having common theoretical terms (sign-designs) or not, are incompatible if and only if they have incompatible O-consequences (Theorem 4, below), and as easily shown, the same holds for a theory and an observation sentence. Now, to *deny* a theory T is to make an assertion, theoretical or observational, which is incompatible with T. But by the conclusion just reached, this denial must therefore be incompatible with the O-consequences, and hence with the Ramsey sentence, R_T, of T. But if assertion of a theory T commits one to accept R_T (since T entails R_T), while denial of T commits one to denial of R_T (since denial of T is incompatible with R_T), then T and R_T must be factually equivalent—i.e., a theory and its Ramsey sentence have the same factual content (Theorem 3, below).

If this argument appears convincing to the reader, he may turn immediately to Section III, pausing in transit only to read the Thesis of Semantic Empiricism, Semantical Principles I–IV, and Postulates 1 and 2. However, despite the conviction sustained by this intuitive demonstration (and given the Thesis of Semantic Empiricism, its force is considerable), it nonetheless glosses over several issues which, though seldom adequately explored, are basic to semantical theory. For example, it was implicitly assumed here that a denial—and no more than this—of theory T *could* be constructed, either as an alternative theory or as a sentence in L_o. But is it not also conceivable that although T is stronger in content than R_T, it is impossible to extract the difference between T and R_T for separate denial while retaining R_T? Again, questions need to be raised about the status of the assumption concerning the incompatibility of two theories having no theoretical terms in common, questions which penetrate to the heart of problems about extrasyntactical incompatibility (i.e., incompatibility in virtue of the *meanings* involved). Therefore, we shall now attempt to reach the conclusions derived so expeditiously in the preceding paragraph by a surer but more arduous route.

I have so far spoken of the "factual content" of a theory or statement in an offhand manner. It now becomes necessary to give this rather vague notion a more precise definition. It has already been noted that the *semantic* status of theories is problematic. However, we speak of "believing," "accepting," or "entertaining" theories in very much the same sense that we apply these terms to observation sentences. In either case, acceptance of a theory or observation sentence has pragmatic repercussions—it involves a behavioral adjustment which, broadly speaking, is pragmatically appropriate or inappropriate according to the facts of reality. In this sense, at least, both theories and sentences stand in the same kind of relations to facts. More specifically, a theory or a sentence is unqualifiedly either "correct" or "incorrect" in virtue of some fact (though perhaps not necessarily an observable fact) which determines the adaptive value of the theory or sentence. This is obviously true of a sentence, which is correct or incorrect (i.e., its acceptance is appropriate or inappropriate) in virtue of the state of reality according to which it is true or false. If it were not also true of a theory—if there were no fact, observational or otherwise, which is a sufficient condition for the correctness or incorrectness of the theory, but rather, the theory were describable only as being more or less useful—then the use of theories would be intrinsically different from that of sentences, and there would exist not even a possibility that a theory might *signify* a fact.

It is quite apparent that whatever we mean by the "factual content," "truth conditions," "factual commitments," etc., of a sentence or theory, these notions are intimately related to the conditions under which the theory or sentence is correct or incorrect. It is also apparent that in the case of a sentence, its factual content is in some sense given by what the sentence asserts. Hence if we can define 'factual content' in terms of the conditions under which a sentence or theory is correct or incorrect without presupposing that the sentence or theory has semantic properties, and can then show the definition to be in suitable agreement with what a sentence asserts, we shall also have, correspondingly, a satisfactory definition of the "factual content" of a theory.

Let us call any fact which is a sufficient condition for a sentence *S* (theory *T*) to be correct a "verifier" of *S* (*T*). More precisely, a fact *f* is a verifier of *S* (*T*) if and only if it may validly be reasoned, "It is a fact that *f*; *therefore*, in view of the behavior-inducing properties of *S* (*T*)—its meaning, its transformational force, etc.—acceptance of *S* (*T*) is pragmatically appropriate."[13] Similarly, any fact which is a sufficient condition for the incorrectness of *S* (*T*) may be called a "refuter" of *S* (*T*). It is important to be clear that the verifiers and refuters of *S* (*T*) do *not* include all facts which are *evidence* for or against *S* (*T*). By "evidence," we mean any fact which (correctly) influences our belief that the sentence (or theory) *has* a verifier or refuter but which is not necessarily *itself* a verifier or refuter of *S* (*T*). For example, consider the sentence *Q*: 'No crows are pink.' Then the fact, say, that over 100,000 crows have been observed under a

[13] Since it is not the bare sentence shape (i.e., sign-design) in language *L* which has truth value, but sentence-shape-cum-meaning (i.e., statement), it is not possible to tell merely from the physical topography of a sentence *S* in *L* whether or not a given fact *f* refutes *S*. It is also necessary to have information, either implicitly in the form of our own language habits if we are actually using *L*, or explicitly if we are evaluating *L* from without, about the linguistic functioning of *S* and perhaps other expressions in *L*. Consequently, any metalinguistic argument that a sentence *S* of language *L* is refuted or verified by a fact *f* must appeal to a set of premises about the character of *L* as a (meaningful) *language*, including some propositions about the conditions under which sentences in *L* are true or false (correct or incorrect, adaptive or maladaptive) in accordance with extralinguistic reality. Different ways of constructing these premises amount to different theories of semantics, and a theory of what such premises should consist of is a theory of *metasemantics*, a discipline in which the ground has scarcely been broken, but which will almost certainly need extensive development before the near universal confusion that seems to exist concerning the nature of semantical concepts is appreciably diminished.

wide variety of conditions and none were pink, is strong evidence that Q is correct, but does not itself verify Q, for the existence of a pink crow is still possible.

If, now, it were the case that a sentence or theory had at most one verifier or refuter, we could simply take this fact to be its factual content. Quite the contrary, however, if a sentence or theory has one verifier or refuter it also has an indefinite number of them. For example, if the fact *f* is a verifier of *S* (*T*), the molecular fact $f \cdot g$ (i.e., the conjunction of *f* and *g*), where *g* is any other fact, is also a verifier of *S* (*T*). In general, we believe that two facts *f* and *g* may in some instances be related in such a way that *f* is a sufficient condition for *g*—i.e., that because *f* is a fact, *g* is, of necessity, also a fact (more briefly: *f*, hence necessarily *g*). That is, in any speculations about what *might* be the case, we should feel it not only unnecessary but logically absurd to consider the possibility that *f*, but not *g*. When such a necessity relation holds between facts *f* and *g*, we say that *f entails g*.

Just what the nature is of this relation of entailment, we fortunately need not attempt to determine here, except, perhaps, to observe that it is reflexive, transitive, but not symmetric. It is sufficient to recognize that there seems to *be* some such relation. Thus we should all agree, surely (barring the natural perversity of philosophical doubt), that if S_1 truly signifies (only) fact *f*, S_2 truly signifies (only) fact *g*, and S_2 is deducible by valid formal inference rules from S_1, then *f* entails *g*. Under such circumstances, we say that *f logically* entails *g*. The extent, if any, to which "entailment" is a broader relation than "logical entailment" is still very much an open question. If S_1 and S_2 truly signify (only) facts *f* and *g*, respectively, then *f* entails *g* if $S_1 \supset S_2$ is "analytically" true; but how the concept of "analyticity" should be analyzed is an issue which currently is raging merrily. (Note that *f* may logically entail *g* even though $S_1 \supset S_2$ is not logically true; for example, S_1 and S_2 may be syntactically independent, but signify the same fact in virtue of containing synonymous terms. Under such circumstances, $S_1 \supset S_2$ is analytically true without being logically true.) Since the present analysis does not depend upon any specific interpretation of "entailment," we may, without prejudice, leave room for the possibility that "*f*, hence necessarily *g*" may hold for certain facts *f* and *g* even though *f* does not logically entail *g*.

The reason we must here recognize the relation of entailment is its intimate connection with conditions of verification and refutation. Whether or not a given fact verifies or refutes a given sentence or theory is something which can be decided only by examining the specific case. However, the notions of "verifying" or "refuting" an entity *E* (where *E* is anything —sentence, theory, or whatever—being used to effect a behavioral adjustment which is appropriate or inappropriate with respect to facts external to the behavior) are deeply embedded in the practical core of linguistic behavior, and we may easily recognize certain principles according to which these concepts are used. Thus,

1. *E cannot have both a verifier and a refuter*—i.e., the terms 'correct' and 'incorrect' are mutually exclusive in their proper application.

2. *If the use of E is such that we can properly say, "If it were the case that p, then E would be correct (incorrect)," then E has either a verifier or refuter.* That is, in view of the behavioral role embodied by *E*, if there are conceivable circumstances under which *E would* be either correct or incorrect, then, in that role, *E is* either correct or incorrect; for if none of the circumstances hold under which *E would* be correct (incorrect), then this is itself a sufficient condition for *E* to be incorrect (correct). The relation of entailment enters this picture in that

3. *If a fact f is a verifier (refuter) of E, then any fact which entails f is also a verifier (refuter) of E.* This may easily be seen by translating "fact f is a verifier (refuter) of *E*" as "since f is the case, *E* is necessarily correct (incorrect) in view of its behavioral properties"—i.e., "f and the facts about *E*'s behavioral role entail that *E* is correct (incorrect)." Then (3) follows by the transitivity of entailment. Such sentence forms as 'f entails g,' 'f verifies (refutes) E,' '*S* signifies a fact entailed by f,' and 'S_1 signifies f, S_2 signifies g, and S_2 is validly deducible from S_1' appear to be so tightly interrelated in meaning that they are undoubtedly grounded primarily on a common underlying concept. One apparently analytic consequence of this common ground is 3. Another is that

4. *If observation sentence S_2 is validly deducible from observation sentence S_1, then any verifier of S_1 is a verifier of S_2, and any refuter of S_2 is a refuter of S_1.*[14] Whenever S_2 is validly deducible from S_1, we say that S_1 "formally entails" S_2. More generally, whether S_2 is syntactically de-

[14] The restriction to observation sentences would seem to be necessary here, for a sentence can apparently be incorrect not only in virtue of signifying a fact falsely, but also in virtue of containing a nonlogical constant which has no referent (see [13]). Hence if S_2 contains a nonlogical constant 'a' not contained in S_1, it is possible for S_1 to be true and S_2 to be false, even though S_1 formally entails S_2, simply because 'a' has no

ducible from S_1 or not, if every verifier of S_1 is a verifier of S_2 and every refuter of S_2 is a refuter of S_1, we say simply that S_1 entails S_2, or, to emphasize that the relation is not necessarily syntactical, that S_1 "analytically" entails S_2. (Note that 'entails' may hence be used to describe either a relation between facts or a relation between sentences. Presumably, one usage is definable in terms of the other.)

We may now define the "factual content" of an entity as the set of facts which verify or refute it. We need to retain a distinction between content which verifies and content which refutes, however, or a sentence will have the same content as its negation.

Definition 3. The positive (negative) factual content of E $=_{def}$ *The set of verifiers (refuters) of E.*

Definition 4. The factual content of E $=_{def}$ *The partially ordered set composed of, first, the members of the positive factual content of E, then the numeral '0,' and, finally, the members of the negative factual content of E.*

Inclusion of '0' in the factual content of *E* is merely a formal device to give a simple distinction between the content of *E* and that of its contradiction. Since an entity cannot be both correct and incorrect, its factual content cannot contain both verifiers and refuters. Hence '0' is either the first or the last member of *E*'s factual content. If it is both, then *E* has neither a verifier nor a refuter, and we may simply say that *E* has no factual content. If E does have factual content, then E is correct or incorrect according to whether '0' is the last or the first member of its content.

In virtue of the relation of entailment, certain members of the factual content of an entity E stand in a special relation to the remainder; namely, those facts of minimal "strength" to verify or refute *E*.

Definition 5. f is a positive (negative) primary content of E $=_{def}$ *f is a verifier (refuter) of E, and any fact which verifies or refutes E entails f.*

Since any fact which entails a verifier or refuter of *E* is itself a verifier or refuter of *E*, a fact f which is a primary content of *E* exhaustively specifies the factual content of *E*, apart from the position of '0,' in that a necessary and sufficient condition for a fact g to be a member of the factual content of *E* is for g to entail f. Hence the factual content of *E* is com-

referent. For example, '(x) (x = x)' formally entails 'Pegasus = Pegasus,' even though the latter is false (see [13]). However, the definition of 'observation term' ensures that all nonlogical constants of L_o have designata, and it then follows by any acceptable theory of semantics that 4 obtains.

pletely described by giving a primary content of E and stating whether it verifies or refutes E. Whether or not every E which has factual content also has a primary content is an interesting question which will not be explored here. Neither are we here able to judge whether E can have more than one primary content. It does follow from Definition 5, however, that if two facts f and g are each a primary content of E, then f and g are analytically equivalent. While this still leaves open the question whether f and g can be analytically equivalent without being identical, the relation must in any case be so intimate that little harm will be done by speaking informally of *the* primary content of E, if it has one.

To what extent does the present definition of 'factual content' agree with intuitive understanding of this term as applied to sentences? By classical semantics, a cognitively meaningful sentence signifies, truly or falsely, exactly one fact, and it is this fact which intuitively is its factual content. But if f is the only fact truly signified by S, then a fact g is a sufficient condition for S to be correct if and only if g entails f; hence by the present definitions, f is the positive primary content of S. Similarly by classical semantics, the negative primary content of a false sentence is the fact falsely signified by it. But as we have seen, the primary content of an entity determines its factual content. Hence the present definition of 'factual content' agrees in logical essentials with intuitive notions insofar as the latter are well formed.

The main purpose of this paper is to clarify the "factual content," "factual commitments," etc., of a scientific theory, and to this end, the first of these expressions has been explicitly defined in terms of conditions of verification and refutation. We have yet to say, however, what kind of an *answer* can be given to a question about factual content. For example, suppose we are asked for the factual content of the sentence Q: 'No crows are pink.' Such queries arise frequently during scientific and philosophical pursuits in the guise "What does Q mean?" or "What are the truth conditions of Q?" It is rather unhelpful to answer, "The factual content of Q is the set of its verifiers, followed by '0,' followed by its refuters," even though by Definition 4 this is literally correct. Rather, what one wants to know is *what* are the verifiers or refuters of Q. But this cannot be answered without first determining what the facts of the case are. We cannot correctly say, for example, that the nonexistence of a pink crow is the primary content of Q unless it is a *fact* that no pink crows exist. It would seem that so long as we restrict "factual content" to include only that which exists and abjure appeal to a nebulous realm of "possibility," we cannot identify the factual content of a sentence or theory without at the same time determining whether it is correct or incorrect. Thus we cannot expect to give *this* kind of an answer in general explication of the factual contents of theories.

The fact remains, however, that in actual practice we do manage effectively to discuss the truth conditions of a sentence without attempting to judge its correctness. We simply resort to considerations of *meaning*, rather than factual content and say that "state-of-affairs" *p* is a necessary and sufficient condition for sentence *S* to be true—thereby indicating that *if* it is a fact that *p*, then *p* verifies *S*, and *if* $\sim p$ is a fact, then $\sim p$ refutes *S*. What is involved here is discovery in our metalanguage of a sentence '*p*' which is related in meaning to sentence *S* in such a way that we can recognize that any verifier or refuter of '*p*' must also be a verifier or refuter of *S*, and conversely. Hence to say (truthfully) that *p* is a necessary and sufficient condition for *S* to be correct, even though it is not known whether it is the case that *p*, is to use a sentence '*p*,' with the same factual content as *S*, in the analysis of *S*. In this way we alleviate our uncertainties about the truth conditions of *S* by reducing the problem of its content to the equivalent problem for another sentence whose meaning, presumably, arouses no puzzlement. In like manner, if we are able to reason, "Theory *T* is verified if it is the case that p, and is refuted if it is not the case that p," then we may conclude that theory *T* has the same factual content as 'p.'

What are the conditions under which a fact may properly be said to verify or to refute a theory? To determine this we can only appeal to the way in which theories are actually used. It will be profitable to concentrate first on the conditions of refutation, for while the history of science is littered with abandoned theories, seldom if ever is a theory judged to be unconditionally verified. Thus while it might be difficult to determine when we would consider a theory correct, it is little trouble to discover conditions of incorrectness.

We saw earlier that to accept a theory *T* involves commitment to the O-consequences of *T*. But if one of the latter is false, then to accept *T* is to be led into error. That is, having a false O-consequence is a sufficient condition for *T* to be incorrect, a conclusion which is amply substantiated by actual practice, in which we feel compelled to revise or discard a theory whenever it leads to an erroneous conclusion that cannot be written off as an "error of measurement," "approximation error," or the like. Thus the refuters of a theory T include all facts which refute an O-consequence of T.

What can we say about the relation of a fact f to theory T if T does *not* have an O-consequence which is refuted by f? One is tempted to answer that at least in the case where f is an *observational* fact (i.e., signifiable in L_o), f refutes T only if it also refutes an O-consequence of T. For do not the O-consequences of a theory constitute its observational force? Unfortunately, this claim prejudges the *semantic* status of the theory. For if T may itself be an assertion, as the realist insists, we must consider the possibility that T falsely signifies a fact which refutes no O-consequence of T, but which is entailed by, or is itself, an observational fact. For example, if investigation of certain known entities $t_1, \ldots, t_n$ eventually reveals it to be the case that $T(t_1, \ldots, t_n)$, the realist may wish to claim that $T(t_1, \ldots, t_n)$ is what the theory '$T(\tau_1, \ldots, \tau_n)$' signified all along, even though nothing signified by an O-consequence of T entails that $T(t_1, \ldots, t_n)$; hence we may also wonder, if it had turned out that $\sim T(t_1, \ldots, t_n)$, whether this would not have refuted T.

This point is so important that I will try to clarify it a bit further. When the realist insists that a theory may itself make an assertion which goes beyond what is asserted by its O-consequences, he has raised the possibility that the theory may be false even though all its O-consequences are true—i.e., that it may be possible for a theory to signify falsely, and hence be refuted by, a fact f even though f refutes none of its O-consequences. But if theories can themselves signify facts, they are surely not limited to signifying only facts which *cannot* be signified in L_o, for then a theory which signifies a certain fact could be deprived of its ability to do so simply by enriching L_o suitably. The mechanism by which a theory is able to signify a fact would indeed be peculiar if it could work only when the observation language is sufficiently impoverished. Hence if a theory can signify a fact at all, we must also suspect that it can signify an observational fact, and the realistic interpretation of theories hence carries with it, at least prior to further analysis, the possibility that a theory may be refuted by an observational fact even when all O-consequences of the theory are true.

Now, there must be *some* principle according to which we can judge, in at least some cases, that an observational fact is *not* a refuter of a given theory, or theories would be utterly useless—we should then never have reason to doubt that a given theory had already been refuted by known facts. Nor is it legitimate, if the realistic interpretation of theories has any intuitive plausibility at all, simply to take for granted that a theory is refuted by an observational fact only if this refutes an O-consequence of the theory. (Oddly enough, this matter—circumscription of the conditions of refutation—has been almost totally overlooked in philosophical analyses of the meanings of theories, although it would now appear that this is actually the core of the problem.) On the other hand, to claim that theories have designative meaning while not also at least implicitly adducing some principle about the *limitations* of possible theoretical reference is sheer philosophic irresponsibility. For given any nontrivial theory '$T(\tau_1, \ldots, \tau_n)$,' there will almost certainly exist *some* set of entities $t_1, \ldots, t_n$ such that $\sim T(t_1, \ldots, t_n)$; hence if we are given no grounds upon which to deny that T asserts that $T(t_1, \ldots, t_n)$—i.e., to deny that T falsely signifies the fact $\sim T(t_1, \ldots, t_n)$—we should never have reason to doubt that T has a refuter.

Now, it seems to me indisputable that the way in which theories are actually used with the observation language *does* impose limits on the possible meanings of theoretical expressions. The fact, for example, that theory T does *not* have an O-consequence which is refuted by observational fact f, while perhaps insufficient grounds for deciding that f does not refute T, is nonetheless *relevant* for judging the factual content of T, whether T is itself an assertion or not. Unless we yield to an unbridled transcendentalism, it is difficult to see how the referential ability, if any, of theoretical terms (which, after all, are only signs with no intrinsic meanings, and whose semantic properties must hence be *acquired*) could derive from anything other than their effective observational import, which, in turn, is determined by the O-consequences of the theory in which they are contained. These rather vague conclusions may be given somewhat more coherent form as a tenet which might be called

The Thesis of Semantic Empiricism: The semantic properties, if any, of theoretical expressions derive, in a potentially useful and syntactically general manner, wholly from their use with the observation language.

That is, 'τ_1,' . . ., 'τ_n' and expressions which contain them have whatever ability to designate that they do have because the 'τ_i' occur in the (perhaps provisionally) accepted theory '$T(\tau_1, \ldots, \tau_n)$.'

The Thesis of Semantic Empiricism by no means claims that all designata of expressions containing theoretical terms are accessible to the observation language—i.e., that the 'τ_i' are definable in L_o—though neither is this possibility ruled out. It insists merely that a theoretical term has no meaning that it *brings to* the theory, so that two theories differing only in their theoretical symbols have the same factual content. As for the stipulation of generality, this is both important and intuitively inescapable: if theories are themselves able to make assertions, the mechanism by which this occurs must be a basic feature of linguistic processes which does not depend upon the theory's being of a special, restricted syntactical form.

The Thesis of Semantic Empiricism expresses the minimum restriction upon the possible meanings of theoretical expressions that can be demanded by anyone who seriously believes that knowledge is based, in some important sense, upon experience; indeed, in its loosely worded form, the Thesis scarcely seems a restriction at all. Yet properly exploited, it suffices to determine the limits of refutation, and hence the factual content, of a theory. In order to show this at all effectively, however, we need to make more explicit what until now has been left to intuitive understanding; namely, the manner in which cognitively meaningful sentences are given truth value through the semantic properties of their constituent terms. The factual content of a meaningful sentence cannot, in general, be determined wholly by its transformational force for other sentences; otherwise, any attempt to judge truth values would precipitate an infinite regress. Hence if we are to pass judgment on the factual content of a theory under the supposition that the theory might itself be meaningful, we can reach no conclusion without some explicit assumptions about the verifiers and refuters of meaningful sentences. The semantical assumptions which follow are very similar in underlying form to the Tarski-Carnap approach, except for expressing a relationship between sentences and their designata, i.e., facts, rather than between sentences and their meanings, i.e., propositions.

Definition 6. Fact f has the form $F(\phi_1, \ldots, \phi_n) =_{def}$ *There exist entities* $t_1, \ldots, t_n$ *such that f is the fact that* $F(t_1, \ldots, t_n)$.

(It will be noted that the "form" of a fact, as defined here, is not unique. For example, the fact signified by a true observation sentence '$P(a_1, a_2)$' has both the form $P(\phi_1, \phi_2)$ and the form $P(\phi_1, a_2)$. Whether a fact's *logical* form—i.e., a form whose description contains no extralogical terms—is in some sense unique is a question on which there is here no need to take a stand.)

Semantic Principle I[15] *(SP I). If '*$S(s_1, \ldots, s_n)$*' is a sentence formed by substituting the symbols '*s_1*,' . . ., '*s_n*' for the variables '*ϕ_1*,' . . ., '*ϕ_n*,' respectively, in the observation predicate '*$S(\phi_1, \ldots, \phi_n)$*,' and '*$S(s_1, \ldots, s_n)$*' is "semantically proper"—i.e., its meaning, or designative potentiality, if any, is governed by the syntax of* L_o*—then '*$S(s_1, \ldots, s_n)$*' signifies a fact f only if there exist entities* $t_1, \ldots, t_n$ *such that '*s_1*,' . . ., '*s_n*' designate* $t_1, \ldots, t_n$*, respectively, and f is either the fact that* $S(t_1, \ldots, t_n)$ *or the fact that* $\sim S(t_1, \ldots, t_n)$*. If '*$S(s_1, \ldots, s_n)$*' signifies f, it does so truly or falsely according to whether f has the form* $S(\phi_1, \ldots, \phi_n)$ *or* $\sim S(\phi_1, \ldots, \phi_n)$*, respectively.*

That is, in more conventional (and in some respects misleading) terms, a sentence '$S(s_1, \ldots, s_n)$' can assert only that $S(t_1, \ldots, t_n)$, where $t_1, \ldots, t_n$ are entities designated by 's_1,' . . ., 's_n.' By saying that the meaning of '$S(s_1, \ldots, s_n)$' is "governed by the syntax of L_o," we rule out the possibility that the meaning of '$S(s_1, \ldots, s_n)$' has been arbitrarily assigned without regard for the meaning of its constituents. It should be noted that SP I does not stipulate that the 's_i' are observation terms. The distinction between "observational" and "theoretical" concerns possible differences in the conditions under which terms acquire their referential powers, whereas SP I deals with the more basic relation between the designative properties of a sentence and those of its constituents. In justification of SP I, one should observe that it merely formalizes *part* of the classic semantical belief that a sentence S makes an assertion in virtue of S's attributing a certain property (in the broad sense) to a set of entities designated by the subject terms of *S*, and that if these entities exemplify the property, this is the fact signified by *S*, whereas if these entities do not exemplify the property, this is the fact in virtue of which S is false. Of course, the classical view may be in error in various respects; in fact, SP I is stated only as a conditional, rather than as the biconditional authorized by the classical view, because further developments will indicate that the latter may be in some respects too strong. But it is incumbent upon anyone who might wish to challenge SP I to provide an alternative statement

[15] Strictly speaking, SP I is not itself an assertion, but is only a schema which generates a set of assertions. We could obtain a single assertion by stating that every sentence (in L_M) obtained by proper substitutions for '$S(\phi_1, \ldots, \phi_n)$' in SP I is true. However, this statement would have to be constructed with care, and need not detain us.

about the conditions under which sentences stand in semantical relations to extralinguistic reality.

There is, to be sure, one objection to SP I which might be raised on more or less orthodox grounds. SP I places no restrictions on the logical forms of facts—indeed, it allows us to invoke facts as needed, corresponding to the true statements in our metalanguage—and this will undoubtedly give offense to one whose tolerance for "facts" does not extend, say, to general or molecular facts. This is not the occasion, even if I were prepared to do so, to take inventory of reality's ingredients, and so I shall abstain from reciting the familiar difficulties which arise when only "atomic" facts are countenanced. Neither shall I explore the possibility that to be surprised that nature should so obligingly fit a fact to each true statement we can construct is very like being perplexed over why the world should be articulated in length units which precisely match our concepts of "inch," "centimeter," etc. Instead, I shall merely venture that if and when it becomes possible to account adequately for the factual content of true observation sentences without presupposing that each signifies a fact, it will not be difficult to accommodate the proof of Theorem 3 to whatever semantical assumptions replace the present ones.

Semantic Principle II (SP II). If a sentence S signifies a fact f truly, S is true and f is a verifier of S. If S signifies f falsely, S is false and f is a refuter of S.

This could also be put by saying that S is correct or incorrect, respectively, under the conditions stated, for when a sentence ascribes a predicate to a set of entities, it is correct or incorrect according to whether or not those entities satisfy the predicate. It should be observed that SP II does *not* say that a formula is false *only* when it signifies a fact falsely. As will be discussed further in Section IV, a descriptive term may be meaningful even when it does not designate anything. Since a sentence which contains such a term may thus be meaningful even though it does not signify any fact, we must allow for the possibility that such sentences should be called false (cf. footnote 14).

It will be noted that SP I, II partially define the conditions under which a sentence signifies a fact; namely, a semantically proper sentence '$\mathrm{S}(\mathrm{s}_1, \ldots, \mathrm{s}_n)$' signifies fact f truly (falsely) only if there exist entities $t_1, \ldots, t_n$ designated by 's_1,' . . ., 's_n,' respectively, such that f is the fact that $S(t_1, \ldots, t_n)$ (the fact that $\sim S(t_1, \ldots, t_n)$), and f verifies (refutes) '$S(s_1, \ldots, s_n)$.' Now, while the intuitive notion of the manner in which a sentence makes semantical contact with extralinguistic reality is unfortunately vague, it would seem to consist essentially in the idea that a sentence S designates, or signifies, a fact f when S and f are related in such a way that the descriptive terms in S designate the constituents of f, and S is true, or false as the case may be, in virtue of the fact that f. Hence the conditions of signification entailed by SP I, II should be sufficient as well as necessary, and we may further assume:

Semantic Principle III (SP III). If '$S(s_1, \ldots, s_n)$' is a semantically proper sentence formed from the observational predicate '$S(\phi_1, \ldots, \phi_n)$' and '$s_1$,' . . ., '$s_n$' designate entities $t_1, \ldots, t_n$, respectively, then: (a) if it is a fact that $S(t_1, \ldots, t_n)$ and this verifies '$S(s_1, \ldots, s_n)$,' then '$S(s_1, \ldots, s_n)$' signifies $S(t_1, \ldots, t_n)$ truly; (b) if it is a fact that $\sim S(t_1, \ldots, t_n)$ and this refutes '$S(s_1, \ldots, s_n)$,' then '$S(s_1, \ldots, s_n)$' signifies $\sim S(t_1, \ldots, t_n)$ falsely.

Since SP I–III together give necessary and sufficient conditions for a semantically proper sentence to "signify" a fact, they may, then, be construed essentially as a definition of this concept. (It would be simple to extend the definition to semantically *improper* sentences such as coded abbreviations, but this is irrelevant for present purposes.) It might seem that the clauses ". . . *and this verifies* [*refutes*] '$S(s_1, \ldots, s_n)$' . . ." in SP III are redundant, for by classical semantics, if 's_1,' . . ., 's_n' designate $t_1, \ldots, t_n$, respectively, the fact that $S(t_1, \ldots, t_n)$ (the fact that $\sim S(t_1, \ldots, t_n)$) is a sufficient condition for '$S(s_1, \ldots, s_n)$' to be true (false) and hence verifies (refutes) the sentence. However, we shall see in Section III that the classical view may not be wholly correct in this assumption, whereas if the classical view *is* correct, the redundancy does not hurt.

Semantic Principle IV (SP IV). No sentence can be both true and false.

This merely makes explicit for sentences the point made earlier, that as we normally use the notion of "correctness," an entity cannot be both correct and incorrect. This does not, of course, preclude the possibility that a sign-design may change its truth value if the meanings of its constituent terms change. But the present analysis presupposes throughout that all linguistic entities concerned remain stable in their various behavioral roles so long as the accepted theory is not altered, and the terms "sentence," "symbol," "expression," etc., here denote not bare sign-designs (which are neither correct nor incorrect as such) but sign-designs cum roles. That an entity cannot be both true and false, or both correct and incorrect, within a fixed context of usage would seem to be an analytic truth about these concepts.

Let us now examine the Thesis of Semantic Empiricism in the light of these principles. According to the Thesis, the semantic properties of a theoretical expression, E, admitted into use through adoption of a theory T are determined wholly by the use of E with respect to the observation language. That is, for every such E, its use with L_o provides a criterion by which can be determined whether or not it designates a given entity. In particular, whether or not T itself signifies a given fact must be so determinable.

Now, the use of an accepted normal syntactic theory 'T$(\tau_1, \ldots, \tau_n)$' with L_o would seem to be comprehensively (albeit schematically) described by saying that T is *formed from* the observational predicate 'T$(\phi_1, \ldots, \phi_n)$' and *used* to generate its O-consequences, where both formation of T and deductions from T *conform to* the syntax of L_o. But that T has the particular O-consequences it does have is a logical result of its formation from 'T$(\phi_1, \ldots, \phi_n)$' and use in conformity with the syntax of L_o. Hence whether or not T signifies a fact f must be determined wholly by whether or not f stands in a certain determinate relation to 'T$(\phi_1, \ldots, \phi_n)$' and the syntax of L_o. One restriction of the possible significata of T on these grounds is obvious: since the semantic properties of T derive from its use with L_o, and this use conforms to the syntax of L_o, the designative potentialities of T, if any, must be governed by the syntax of L_o. Hence by SP I, if T signifies a fact, this must either be a fact of form $T(\phi_1, \ldots, \phi_n)$ truly signified by T, or a fact of form $\sim T(\phi_1, \ldots, \phi_n)$ falsely signified by T. That is, if T itself makes an assertion, it must do so by ascribing the predicate 'T$(\phi_1, \ldots, \phi_n)$' to some set of entities $t_1, \ldots, t_n$ designated, respectively, by 'τ_1,' . . ., 'τ_n,' a conclusion which should be intuitively evident even without appeal to the Thesis. Being of either of these two forms cannot be a *sufficient* condition for a fact to be signified by theory T, however, or T would simultaneously signify truly all facts of form $T(\phi_1, \ldots, \phi_n)$ and falsely all facts of the form $\sim T(\phi_1, \ldots, \phi_n)$, in flagrant violation of SP IV. Hence the predicate 'T$(\phi_1, \ldots, \phi_n)$' and the syntax of L_o must impose on the possible significata of T an additional condition which excludes this possibility. And since it is 'T$(\phi_1, \ldots, \phi_n)$' which gives the theoretical terms their particular character, contrasted to the semantic properties, if any, which would have been imparted by their normal syntactic use with a semantically different predicate of L_o, it is clear that this additional condition must stem essentially from the linguistic force of '$\mathrm{T}(\phi_1, \ldots, \phi_n)$.' More precisely, since any one fact of the form $T(\phi_1, \ldots, \phi_n)$ or $\sim T(\phi_1, \ldots, \phi_n)$ is syntactically as legitimate a significatum of T as any other, then if T signifies the one but not the other, it must be because the one is related to '$\mathrm{T}(\phi_1, \ldots, \phi_n)$' in virtue of the latter's linguistic properties—i.e., is "semantically" related to '$\mathrm{T}(\phi_1, \ldots, \phi_n)$'—in a manner in which the other is not.

Now, there are a great number of possible ways in which a fact may be semantically related to the predicate '$\mathrm{T}(\phi_1, \ldots, \phi_n)$' that might be considered as potential criteria for whether or not a given fact is signified by theory T. For example, there are an indefinite number which depend on the predicate's being of some special syntactic structure, such as the relation that obtains between a predicate of the form '$\mathrm{P}(\phi) \cdot \mathrm{Q}(\phi)$' and a fact of form $P(\phi)$, or the relation between a fact $P(a)$ and a complex predicate '$\mathrm{P}(\phi)$' when the latter contains a constant which refers to a (e.g., the fact that $2 \leqq 2$, and the predicate '$2 \leqq \phi$'). It would be excessively tedious to examine these various and for the most part artificial possibilities in any detail. Fortunately, most are immediately dismissed by the demand of the Thesis that theoretical expressions derive their meanings, if any, in a syntactically *general* manner. This rules out the possibility that if theoretical expressions can designate at all, only those formed from predicates of special syntactical characteristics can do so. Hence an acceptable criterion for whether or not a fact is signified by T cannot specify any special syntactical structure for '$\mathrm{T}(\phi_1, \ldots, \phi_n)$.'

The next point to be considered is that in (presumed) contrast to relations between facts and observation *sentences*, it does not seem possible to find any *one-many* semantic relations between facts and observation *predicates* which do not constrain the predicate to special characteristics. That is, if the Thesis of Semantic Empiricism is correct, there appears to be no way in which '$\mathrm{T}(\tau_1, \ldots, \tau_n)$' can be assured a unique significatum.[16] Any criterion for what is signified by T, based only on the rela-

[16] An example of a potential criterion which singles out at most a unique significatum for T by violating the condition of syntactical generality is the following: *A theory T signifies a fact f if and only if T is of syntactical form* '$\mathrm{T}(\tau_1, \ldots, \tau_n) \cdot (c_1 = c_1) \cdot \ldots \cdot (c_n = c_n)$' *and there exist entities $t_1, \ldots, t_n$ such that* 'c_1,' $\ldots$, 'c_n' *designate* $t_1, \ldots, t_n$, *respectively, while f is either the fact that* $T(t_1, \ldots, t_n) \cdot (t_1 = t_1) \cdot \ldots \cdot (t_n = t_n)$ *or the fact that* $\sim[T(t_1, \ldots, t_n) \cdot (t_1 = t_1) \cdot \ldots \cdot (t_n = t_n)]$. This prospective criterion guarantees T a unique significatum so long as T is of the necessary syntactical form even though there may be many facts of the form $T(\phi_1, \ldots,$

tion of a fact to the predicate '$T(\phi_1, \ldots, \phi_n)$' in virtue of the linguistic properties of the latter and without regard for any special form the predicate may have, will in principle be satisfiable by an indefinite number of facts. Now, we have already seen that in order for SP IV to be satisfied, our criterion must exclude the possibility that it is simultaneously satisfied both by a fact of the form $T(\phi_1, \ldots, \phi_n)$ and a fact of the form $\sim T(\phi_1, \ldots, \phi_n)$, since if this possibility were realized, T would signify the one fact truly and the other falsely. But since the criterion allows T in principle to signify more than one fact, the only apparent way to ensure against the significata of T including both facts of the form $T(\phi_1, \ldots, \phi_n)$ and facts of the form $\sim T(\phi_1, \ldots, \phi_n)$ and still have the criterion concern *only* the semantic relation of a fact to '$T(\phi_1, \ldots, \phi_n)$'[17] is for it to be impossible for facts of one of the two forms to satisfy the criterion. That is, it must be either that facts of the form $T(\phi_1, \ldots, \phi_n)$ or that facts of the form $\sim T(\phi_1, \ldots, \phi_n)$ are excluded as possible significata of T; and since a fact must be of one of these forms in order to be signified by T, it must be that either T can signify only facts of the form $T(\phi_1, \ldots, \phi_n)$ or only facts of the form $\sim T(\phi_1, \ldots, \phi_n)$. Now, if a fact had to be of the latter form in order to be signified by T, then, since T could signify such a fact only falsely, a theory could itself signify a fact only if the theory were false. But it would be absurd to propose seriously that despite a theory user's patent effort to make a true assertion thereby, or at the very least to make correct commitments as to what is the case, that the only way a theory can acquire cognitive meaning is through be-

$\phi_n) \cdot (c_1 = c_1) \cdot \ldots \cdot (c_n = c_n)$ or its negation; however, it would be absurd to propose this seriously as the criterion for what a theory can assert. For another illustration, consider a potential criterion which limits T to the form '$T(\tau_1, \ldots, \tau_n) \cdot (\tau_1 = c_1) \cdot \ldots \cdot (\tau_n = c_n)$' if T is to have a significatum. Again, this assures a unique significatum to any theory which meets the special syntactic requirement, but would limit semantically meaningful theories to those which are formally equivalent to an observation sentence—namely, '$T(c_1, \ldots, c_n)$'—and hence categorically denies the possibility of a realistic interpretation of theories in the more general case.

[17] One can always get around SP IV by adding a safety clause which brings in extraneous material. For example, if '$\Gamma(f)$' is a predicate describing a potential criterion for what is signified by T but which is simultaneously satisfiable by facts of both forms in question, the revised criterion '$\Gamma(f) \cdot [(\phi_1, \ldots, \phi_n)T(\phi_1, \ldots, \phi_n) \vee \sim(\exists \phi_1, \ldots, \phi_n)T(\phi_1, \ldots, \phi_n)]$' no longer violates SP IV, but neither does its satisfaction depend only on a semantic relation of the fact f to the predicate '$T(\phi_1, \ldots, \phi_n)$.' It should be clear that introducing safety clauses of this sort into potential criteria for what is expressed by a theory violates the intent of the Thesis of Semantic Empiricism when the latter stipulates that the meaning of a theory depends *only* on the theory's use with L_o, and not, in addition, upon whether or not a number of other conditions, which have nothing to do with this usage, obtain.

coming false. For this reason, the Thesis of Semantic Empiricism as formulated above rejects this possibility by including the stipulation that a theory derives its meaning, if any, in a potentially useful manner. Hence facts of the form $\sim T(\phi_1, \ldots, \phi_n)$ are excluded from what can be signified by T, and we conclude that SP I–IV and the Thesis of Semantic Empiricism imply that a theory '$T(\tau_1, \ldots, \tau_n)$' can itself signify a fact only if the fact is of the form $T(\phi_1, \ldots, \phi_n)$. That is,

Postulate 1 (P 1). If '$T(\tau_1, \ldots, \tau_n)$' *is an accepted normal syntactic theory in which* '$T(\phi_1, \ldots, \phi_n)$' *is an observational predicate and* 'τ_1,' . . ., 'τ_n' *are theoretical terms, then* '$T(\tau_1, \ldots, \tau_n)$' *signifies a fact f only if there exist entities* $t_1, \ldots, t_n$ *designated by* 'τ_1,' . . ., 'τ_n,' *respectively, such that f is the fact that* $T(t_1, \ldots, t_n)$.

It is an immediate consequence of SP I and P 1 that *a theory can itself signify a fact only truly.*[18] This conclusion might at first seem somewhat counterintuitive; and since Theorems 3–5 depend crucially upon it (more accurately, upon the slightly weaker premise that a theory can falsely signify only a fact which refutes one of its O-consequences), I shall try to show that it claims no more than lies implicit in our normal use of theories. As pointed out earlier, if a theory can signify a fact falsely, then we must expect that a theory may also be refuted by an observational fact even when all its O-consequences are true. Now, while we modify or abandon theories for a variety of reasons (excessive formal complexity, too many *ad hoc* accretions, probabilistic disconfirmation of an O-consequence, etc.), it is, I think, an indisputable fact about natural usage that we never regard a known fact to be an actual refuter of a theory unless the fact disproves an O-consequence of the theory. Thus we behave *as though* it were impossible for a theory to be refuted by an observational fact which does not refute an O-consequence of the theory, and hence as though it were impossible for a theory itself to signify a fact falsely (or, at any rate, impossible for the theory to signify falsely a fact which does not refute one of its O-consequences).

Again, suppose a person who accepts theory '$T(\tau_1, \ldots, \tau_n)$' is shown a set of entities $t_1, \ldots, t_n$ and asked whether by proposing this theory, he means to claim that $t_1, \ldots, t_n$ satisfy '$T(\phi_1, \ldots, \phi_n)$.' If T is really the totality of that person's theoretical assumptions, so that he has no

[18] This by no means implies that a theory cannot be incorrect, or even semantically false (see the comment following SP II). The conclusion is only that if a theory is incorrect, it must be so in a way other than by signifying a fact falsely—as, for example, by having an O-consequence which signifies falsely.

commitments to the identities of the theoretical entities other than what is made explicit in T, the fact that $T(t_1, \ldots, t_n)$ is not the case suffices for him to reject the supposition that T falsely signifies the fact that $\sim T(t_1, \ldots, t_n)$—otherwise, we should be able to make him doubt that T is correct merely by finding a set of entities (of appropriate types) which did *not* satisfy '$T(\phi_1, \ldots, \phi_n)$.' Now, the fact that in practice we *recognize* as refuters of a theory only those facts which refute one of its O-consequences does not in itself prove that the theory *has* no observational refuters beyond this class, since we are not entitled to take for granted that our use of theories gives them no factual content beyond that which we normally recognize. What natural usage does show, however, is that *if* actual practice is correct, a theory can signify a fact only truly. Hence what P 1 amounts to is an assertion that natural usage is correct.

Returning to the Thesis of Semantic Empiricism, our analysis so far has left us only a step from an important conclusion which, though unnecessary for Theorems 3–5, will be needed in Section III. We have seen that in order to be signified by theory T, a fact must be of form $T(\phi_1, \ldots, \phi_n)$. But it should also be apparent that the use of T with L_o is unable to differentiate further among facts of this form unless, contrary to the Thesis, one takes into consideration special characteristics of '$T(\phi_1, \ldots, \phi_n)$.' For as already seen, the predicate '$T(\phi_1, \ldots, \phi_n)$' (together with the syntax of L_o) fully characterizes the observational use of T, and it does not seem possible to find a difference in the way two facts, both of form $T(\phi_1, \ldots, \phi_n)$, could be related to '$T(\phi_1, \ldots, \phi_n)$' which does not draw upon special features of the predicate.[19] That is, there are no plausible grounds upon which a person who had accepted T could say, after extending the range of facts with which he is observationally acquainted, "Although I now know both that $T(t_1, \ldots, t_n)$ *and that* $T(t'_1, \ldots, t'_n)$, it was the former, rather than the latter, that I intended to signify by '$T(\tau_1, \ldots, \tau_n)$.'" Similarly, he could not legitimately claim, "Although I now know it is the case that $T(t_1, \ldots, t_n)$, this isn't what I meant by '$T(\tau_1, \ldots, \tau_n)$,'" unless he would make the same judgment about *any* fact of the form $T(\phi_1, \ldots, \phi_n)$. The conclusion seems inescap-

[19] For example, two facts $T(t_1, \ldots, t_n)$ and $T(t^*_1, \ldots, t^*_n)$ could differ in whether or not the predicate '$T(\phi_1, \ldots, \phi_n)$' contains a constant which designates one of the t_i or t^*_i. If this kind of difference could matter with respect to what is signified by a theory, then certain theories would be barred from the possibility of signification simply because they fail to exhibit certain features of formal complexity.

able that all facts of the form $T(\phi_1, \ldots, \phi_n)$ must equally satisfy any acceptable criterion for what is signified by T, and hence

Postulate 2. (Not to be used until Section III.) *If an accepted normal syntactic theory* 'T$(\tau_1, \ldots, \tau_n)$' *is able to signify any fact at all, it signifies all facts of form* $T(\phi_1, \ldots, \phi_n)$.

Although the detailed arguments of the past few pages may seem a bit rarified in spots, the basic development so far is actually quite robust: If the nonobservational components of an accepted theory acquire their meanings, if any, in virtue of this usage, then the observational predicate which characterizes the theory must provide the criterion for its factual significance, and there are presently no intelligible grounds on which to suspect that such a criterion would admit as a designatum of the theory a fact which did *not* comply with the form specified therein. But now the argument becomes more delicate, for while P 1 takes an important step toward determining the refuters of a theory by excluding the possibility that the theory signifies falsely, limiting conditions yet remain to be imposed on facts which refute without being signified. Such limits will now be developed by a four-stage postulate which commands assent by drawing on our intuitive feel for the meanings of certain ill-defined but basic semantical notions. If the considerations which follow appear esoteric, they are no more so than the traditional gambits of philosophical speculation which they are designed to forestall.

Postulate 3a. *If a sentencelike formula is not itself cognitively meaningful when accepted, then the only way in which its acceptance can be pragmatically inappropriate is for the formula to be inappropriate as a transformation mechanism. That is, an accepted sentencelike but cognitively meaningless formula has a refuter only if it transforms a true cognitively meaningful sentence in its accepter's language into a false one.*

While we shall not here attempt a formal definition of "cognitively meaningful," we may roughly interpret it as "having pragmatic import in virtue of its semantic properties." And if a formula is not given pragmatic import by its semantic properties, if any, it is difficult to imagine what could be relevant to the appropriateness of its acceptance except its efficiency at transforming expressions which are cognitively meaningful. It must be confessed that the second sentence in P 3a is not altogether a mere clarification of the one which precedes it. For it may fairly be asked whether a transformation technique which does not lead to error when operating only upon its user's present language might not still be ultimately incorrect through having false consequences in an augmented language whose resources *T*'s user may eventually attain through subsequent language enrichments (e.g., by addition to his stock of observation terms as a result of new experiences). The extent to which acceptance of a transformation technique involves commitments in languages not actually in use at the time of the adoption, and whether this entails that the acceptability of a cognitively meaningless transformation procedure cannot be decided on the basis of its import for the existent language alone, is an issue calling for somewhat more extensive analysis than seems profitable on this occasion. It can be shown with the help of Lemma 1, however, that so long as L_0 is as powerful syntactically as stipulated here, P 3a is not disturbed by such considerations.

According to classical views on semantics, P 3a plus a simple extension of P 1 to cover theoretical consequences of *T* in addition to *T* itself are sufficient to determine the refuters of a theory. For in the classic tradition, a sentencelike formula '$S(s_1, \ldots, s_n)$' is held to be cognitively meaningful only if there exist entities $t_1, \ldots, t_n$ designated by 's_1,' . . ., 's_n,' respectively, and through which '$S(s_1, \ldots, s_n)$' is then verified or refuted according to whether or not it is the case that $S(t_1, \ldots, t_n)$. However, a persistent, though less clearly conceptualized, alternative interpretation of the conditions of meaningfulness has been that a sentence which is cognitively meaningful can nonetheless fail at factual reference through lack of designata for some of its descriptive components—in fact, such a view will later be proposed in Section III, below. Hence we should also consider the possibility that a sentencelike formula may fail to signify a fact and yet escape meaninglessness. But what could refute such a formula? It seems inescapable that if the linguistic attributes of a cognitively meaningful sentence S do not establish some immediate semantical correspondence between *S* and an external reality which determines its correctness, the pragmatic import of S can come only through its meaning connections with other cognitive states which establish eventual reference to the conditions of verification or refutation. That is, to employ notions which have not previously been introduced here and which will appear only briefly as an intuitive justification for Postulate 3b, if *S* expresses a proposition which does not itself signify a fact, *S* has a refuter only if *S* commits its believer to some false proposition *p* which does so signify. Now if *S* makes commitment to p, this is accomplished through some relation of syntax and meaning between S and p. But if S is an accepted normal syntactic theory T, the Thesis of Semantic Empiricism implies that the theoretical terms in '$T(\tau_1, \ldots, \tau_n)$' contribute no extrasyntactical implicative force beyond what is already contained in the predicate '$T(\phi_1, \ldots, \phi_n)$.' Hence if T makes commitment to p, so for like reason should any other sentence formed by instantiation of '$T(\phi_1, \ldots, \phi_n)$.' But then any disjunction of instantiations of '$T(\phi_1, \ldots, \phi_n)$' and hence also, as a limiting case, '$(\exists\phi_1, \ldots, \phi_n)T(\phi_1, \ldots, \phi_n)$,' should likewise make commitment to p; so any refuter of p must also refute T's prime consequence, R_T. Accordingly,

Postulate 3b. If an accepted normal syntactic theory T is cognitively meaningful but does not signify a fact, then T has a refuter only if its prime consequence has a refuter.

There is yet one further possibility to be reckoned with, namely, that some of the theoretical consequences of T—i.e., sentencelike formulas containing theoretical terms and deducible from T—are cognitively meaningful when T is accepted, even though T itself is not. To be sure, if the notion of "cognitive meaningfulness" is already vague, speculation about the circumstances under which T could have meaningful theoretical consequences while itself remaining meaningless is vagueness compounded. Nonetheless, this is one of the logical alternatives which must be accounted for by any exhaustive attempt to delimit the factual content of theories, and fortunately, by groping cautiously, we can work our way across to firm ground. To begin with, we may safely assume that a syntactically complex, well-formed expression is cognitively meaningful if and only if its formal constituents are cognitively meaningful. Hence,

Postulate 3c. If E_1 is an expression containing theoretical terms 'τ_1,' . . ., 'τ_n' which is given cognitive meaning by acceptance of a normal syntactic theory T, and E_2 is a well-formed expression (of any type) containing no theoretical terms other than 'τ_1,' . . ., 'τ_n,' then E_2 is also cognitively meaningful when T is accepted.

Let a meaningful consequence M of accepted theory T such that all meaningful consequences of T are also deducible from M be called a "meaning abstract" of T. More precisely,

Definition 7. M is a meaning abstract of normal syntactic theory $T =_{def}$ M is a sentencelike formula which is cognitively meaningful when T is accepted as a normal syntactic theory, and any sentencelike formula S which is cognitively meaningful when T is accepted is deducible from T if and *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 '$T(\tau_1, \ldots, \tau_n)$' 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 'τ_1,' . . ., 'τ_m,' $R^m{}_T =_{def}$ '$(\exists\phi_{m+1}, \ldots, \phi_n)T(\tau_1, \ldots, \tau_m, \phi_{m+1}, \ldots, \phi_n)$.' 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 $R^m{}_T$

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.[20] 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 construing 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.

[20] 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.

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

[21] If the sense of this claim is not readily apparent, it would be advisable to review the discussion on p. 296, including fn. 13, of how the verifiers or refuters of an expression are in principle to be identified. Formalizing this identification procedure would have raised the logical complexity of the present arguments, already difficult enough, to a prohibitive level. Without a grasp of the unformalized procedure, however, the reader will be unable to appreciate how, in what ensues, assumption that fact *f* is a verifier or refuter of one linguistic entity leads to a conclusion that *f* is also a verifier or refuter of another.

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 discoordinated 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 incompatible when and only when they jointly make a self-contradictory assertion—i.e., their conjunction is necessarily false and hence refuted by every fact. But the analogous claim for theories, namely, that theories T_1 and T_2 are incompatible when and only when their conjunction is refuted by every fact, will not do at all. For according to the views developed here, a theory acquires factual commitments only through being accepted, and the force of a given sentencelike formula may be expected to be influenced by whatever additional theoretical postulates have also been adopted.

Thus to ask whether theories T_1 and T_2 are incompatible is, more precisely, to ask whether T_1, when accepted as a totality of theoretical assumptions, makes factual commitments which are incompatible with those made by T_2 when the latter is accepted as an alternative to T_1. Consequently, we cannot test the incompatibility of T_1 and T_2 merely by examining the force of simultaneous acceptance of the formulas comprising T_1 and T_2, for $T_1 \cdot T_2$ is still *another* (alternatively acceptable) theory in which the factual content of sentencelike constituent T_1 or T_2 may not be altogether the same as its content when accepted as a theory by itself. In particular, *it must not be presupposed that two sentencelike formulas which formally entail a contradiction when accepted jointly are necessarily contradictory in their factual commitments when (alternatively) accepted singly*. For felicity of expression, we shall frequently speak, henceforth, of the factual content of a *theory* T without explicitly stipulating that T is normal syntactic and accepted. It should be understood, however, that what is thereby meant is not the factual content, if any, that *sign-design* T may actually have at the moment, but, subjunctively, the content that T *would* have if T *were* the totality of theoretical postulates accepted for normal syntactic use with L_0. This does not preclude the possibility that sign-design T might have the same factual content when part of a more inclusive accepted theory as it does when accepted by itself, but it forbids such an assumption being made without special argument (such as, for example, the one which supported P 3d).

The preceding considerations, however, have been deliberately roundabout in order to warn against seductive pitfalls. A perfectly natural way to explain "incompatibility" is to say that two sentences are incompatible if and only if they cannot both be true. To extend this notion to any pair of expressions E_1 and E_2 which have factual content (or which *would* have, if used in a stipulated way), we have but to substitute "have verifiers" for "be true"—i.e., E_1 and E_2 are incompatible if and only if they cannot 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 definiens 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) $=_{def}$ 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 tautologous 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_1 "under U" is meant, of course, the verifiers or refuters that E_1 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 $f \vee g$ 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

S_1 and S_3 are incompatible if and only if S_2 and S_3 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^i' 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^i,' and (b) a purely logical variable, 'x^j,' whose range includes that of 'x^i.'[22] But in such a language, for every sentence S_i containing nonlogical variables, there is an analytically equivalent sentence, S^*_i, containing descriptive constants corresponding to the ranges of the nonlogical variables in S_i and whose variables are only logical. Let such an S^*_i be called the "L(ogically)-normal form" of S_i. Then if *T* is a theory, containing theoretical variables as well as constants, whose factual content we wish to analyze, we simply find the L-normal form of *T* and proceed as before. In short, our assumption that the theoretical terms of a normal syntactic theory are constants places a restriction on the applicability of Theorems 2–5 only for an artificially weak language, just as Theorem 2 may fail for a language which does not permit existential quantification over all primitive descriptive constants.

[22] 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^j can be designated by a symbol of type i,' this is equivalent in force to 'x^j is a member of C,' where 'C' designates the class corresponding to formal type *i*. Moreover, if 'x^j' is also a nonlogical variable, we would have to determine that *t* belongs to the range of 'x^j' before using this predicate to inquire whether or not *t* belongs to the range of 'x^i'; 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.

III

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 $T(\tau) =_{\mathrm{def}} S$—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(τ)' 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(τ)' not merely carries factual content but also *signifies* a fact, whether it could do this in any way other than by 'τ' standing in a relation of reference to some entity of appropriate type. It would be *possible*, for example, to stipulate that 'T(τ)' is to mean that *p*, where 'p' is some observation sentence of syntactical form different from 'T(τ).' 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(τ)' is not governed by the syntax of

L_o. 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 $T(\tau) =_{def} (\exists\phi) T(\phi)$.

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

1. If 'T(τ)' 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(τ).' But if 'T(τ)' asserts that $(\exists\phi)T(\phi)$, then applying formal inference rules to the sentence form 'T(τ)' *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 $T(\tau) =_{def} (\exists\phi) T(\phi)$, we could *vindicate* taking a sentence in L_o, syntactically deducible from 'T(τ),' to be a logical consequence of '$(\exists\phi)$T(ϕ)' and hence of 'T(τ),' 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 $T(\tau) =_{def} (\exists\phi)T(\phi)$, 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 inacceptable to grant semantic status to 'T(τ)' and withhold it from its theoretical consequences, especially since in the *de facto* use of theories, the full conjunction of theoretical postulates, 'T(τ),' 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 $S_1(\tau) \cdot S_2(\tau) =_{def} (\exists\phi)[S_1(\phi) \cdot S_2(\phi)]$, 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 $T(\tau) =_{def} (\exists\phi)T(\phi)$, 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(τ)' unless S has the same syntactical structure as 'T(τ).' Since in this case, S must be of the form 'T(d),' where 'd' is a

referring expression in L_o, setting $T(\tau) =_{def} T(d)$ is the same as setting $\tau =_{def} d$. We may therefore conclude that if 'T(τ)' signifies a fact, it does so because 'τ' 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 'τ' might not be equivalent to some descriptive expression lying wholly within L_o. (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(τ)' is a cognitively meaningful statement in which 'τ' functions referentially so long as he were allowed to argue that there existed an expression 'd' in L_o with the same designative force as 'τ.' He would then contend that 'T(τ)' is equivalent to the observation sentence 'T(d),' and similarly, that any other expression 'E(τ)' is equivalent to 'E(d).'

What conditions must obtain if 'τ' 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(τ)' 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(τ).' Only two alternatives seem available to him:

1. He might assert that the refuters of 'T(τ)' are exactly the refuters of 'T(d),' where 'd' is that observational expression to which he claims 'τ' is equivalent. But to select 'd' for this purpose, rather than some other of the many expressions of the same formal type as 'τ' available in L_o, is simply to take the alleged equivalence of 'τ' and 'd' as a *premise* from which to analyze the force of 'T(τ),' which would be justified only if 'τ' had been explicitly introduced to have the same meaning as 'd.' But in this case, 'τ' is merely an ordinary defined term in the observation language, whereas we are here concerned with a class of terms which have *not* been introduced in this way. Hence this alternative is inacceptable.

2. He may hold that 'T(τ)' 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 'τ' 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 '$T(\tau)$'—must then also refute '$(\exists\phi)T(\phi)$.' 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 '$T(\tau)$' unless the theory has an O-consequence which is refuted by $\sim T(d)$; otherwise, by parity of reasoning, any observational fact $\sim T(d_1)$ should disprove the theory. Hence, unless $(\exists\phi)T(\phi)$ entails that $T(d)$, the force of 'T(d)' is stronger than that of '$T(\tau)$,' and 'τ' and 'd' cannot be equivalent. Consequently, a prerequisite of the positivist's thesis is that whenever a theory '$T(\tau)$' is itself a cognitively meaningful sentence, *there exists a descriptive expression* 'd' *in* L_o *such that* '$(\exists\phi)T(\phi)$' 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 '$(x)[\tau(x) \supset P(x)] \cdot \tau(a)$' is a theory in which 'P' and 'a' are observational constants, then '$(\exists\phi)[(x)[\phi(x) \supset P(x)] \cdot \phi(a)]$' is formally equivalent to '$(x)[P(x) \supset P(x)] \cdot P(a)$,' and a positivist could contend that 'τ' is equivalent to 'P.' However, it is obviously *not* the case that for *any* predicate '$F(\phi)$,' a descriptive expression 'd' can be found such that 'F(d)' has the same force as '$(\exists\phi)F(\phi)$'—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, '$T(\tau)$,' 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 '$T(\tau)$' 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 of 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 extraobservational 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 '$T(\tau_1, \ldots, \tau_n)$' *is an accepted normal syntactic theory in which* '$T(\phi_1, \ldots, \phi_n)$' *is an observational predicate and* 'τ_1,' ..., 'τ_n' *are theoretical terms, then:* (a) '$T(\tau_1, \ldots, \tau_n)$' *signifies a fact f if and only if there exist entities* $t_1, \ldots, t_n$ *such that f is the fact that* $T(t_1, \ldots, t_n)$. (*b*) *If t is an entity such that* $(\exists\phi_1, \ldots, \phi_{i-1}, \phi_{i+1}, \ldots, \phi_n)T(\phi_1, \ldots, \phi_{i-1}, t, \phi_{i+1}, \ldots, \phi_n)$, *then* '$\tau_i$' *designates t.*

That is, reverting to the simplest case, accepted theory '$T(\tau)$' signifies every fact of form $T(\phi)$, and 'τ' designates every entity which satisfies '$T(\phi)$.' It should be noticed, however, that Theorem 6 supplies sufficient but not necessary conditions for 'τ_i' 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 '$T(\phi)$,' we could regard '$T(\tau)$' as assigning a referent to 'τ' by means of *description*—i.e., we could assume that $\tau =_{def} (\iota\phi)T(\phi)$.[23] 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

[23] Owing to limitations of the Linotype font the regular instead of the inverted iota is used.

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,*[24] 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 $\sim P(t)$. 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 t_1 and t_2. If t_1 and t_2 are different entities, there must be some property P such that $P(t_1)$ and $\sim P(t_2)$. 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 t_1 and t_2, it is entirely possible that $F(t_1)$ while not $F(t_2)$, 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

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

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 jeopardous 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* $=_{def}$ '$E(\tau_1, \ldots, \tau_m)$' *is a sentencelike formula deducible from T in which* '$E(\phi_1, \ldots, \phi_m)$' *is an observational predicate and* 'τ_1,' ..., 'τ_m' *are m of the n* $(0 < m \leq n)$ *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 form $C \cdot S_E$, 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 $=_{def}$ *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* 'τ_1,' ..., 'τ_m' *are theoretical terms and* '$E(\phi_1, \ldots, \phi_m)$' *is an observational predicate, then* 'τ_i' *designates an entity t if and only if it is the case that* $(\exists \phi_1, \ldots, \phi_{i-1}, \phi_{i+1}, \ldots, \phi_m) E(\phi_1, \ldots, \phi_{i-1}, t, \phi_{i+1}, \ldots, \phi_m)$.

Since a set of entities $t_1, \ldots, t_n$ will satisfy '$T(\phi_1, \ldots, \phi_n)$' only if its subset $t_1, \ldots, t_m$ satisfies '$E(\phi_1, \ldots, \phi_m)$,' an entity *t* will qualify as a designatum of 'τ_i' under Theorem 6 only if it also qualifies under

[25] While these hypotheses concern only the acquisition of designata 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).

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* 'τ_1,' ..., 'τ_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* 'τ_1,' ..., 'τ_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 'τ' (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 'τ' designates an entity t and $\sim E(t)$ obtains, then '$E(\tau)$' is false. Rather, for '$E(\tau)$' to be false, *every* designatum of 'τ' 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.[26] 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 '$E(\tau)$,' not deducible from accepted theory T, falsely signifies a fact $\sim E(t)$, even though 'τ,' as introduced by theory '$T(\tau)$,' designates t. Suppose that there exists a t_1 such that $T(t_1)$ and $\sim E(t_1)$, and also a t_2 such that $T(t_2)$ and $E(t_2)$. Then by the present interpretation, 'τ,' introduced by '$T(\tau)$,' designates both t_1 and t_2 (cf. Theorem 6), and one might argue on classical grounds that if this is granted, then we should have to say that '$E(\tau)$' falsely signifies the fact that $\sim E(t_1)$ as well as truly signifying the fact that $E(t_2)$. Now, the concept of "incorrectness," of which "falsehood" is a special case, is pragmatical—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 '$E(\tau)$' *in addition* to accepting the theory '$T(\tau)$' is to accept the *enriched* theory '$T(\tau) \cdot E(\tau)$.' Since by hypothesis it is the case that $T(t_2) \cdot E(t_2)$, it follows that '$E(\tau)$' is then a consequence of the unambiguously true theory '$T(\tau) \cdot E(\tau)$,' and so

[26] When classical semantics analyzes the linguistic properties of the observation language through sentences of the form "'$S(s)$,' in L_o, asserts that $P(t)$," what is *primarily* being indicated is a relationship among the *meanings* of sentences in L_o and L_M; 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_o to their *designata*, since not all meaningful expressions have designata, even when their syntactic role is that of a descriptive term.

by Hypothesis A and SP I does *not* falsely signify $\sim E(t_1)$; hence there are then no grounds on which to argue that 'E(τ)' is incorrect. That is, when 'τ' is introduced by theory 'T(τ),' so long as it is the case that $(\exists\phi)[T(\phi) \cdot E(\phi)]$, the correctness of 'E($\tau$)' is uncontaminated by the existence of a t such that 'τ' designates t and it is the case that $\sim E(t)$, for 'E(τ)' can lead one into error only in the course of adopting a new, improved theory 'T(τ) · E(τ),' and with respect to the latter theory under the conditions stipulated, 'E(τ)' is in no way incorrect. But by SP II, if 'E(τ)' is not incorrect, it does not falsely signify a fact. Hence 'E(τ)' 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(τ)' falsely signifies $\sim E(t_1)$ depends only on the relation of $\sim E(t_1)$ to the use of 'E(τ),' and not, in addition, on whether or not some other entity t_2 satisfies 'E(ϕ).' To be sure, 'E(τ)' is false when there is no joint satisfier of 'T(ϕ)' and 'E(ϕ),' but only because 'E(τ)' 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(τ)' is not entailed by the theory which gives 'τ' 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 E_1 and E_2 may each be true separately, their conjunction may be false.[27] For example, if two distinct entities t_1 and t_2 each satisfy the observational predicate 'T(ϕ)' of accepted theory 'T(τ),' both '$\tau = t_1$' and '$\tau = t_2$' are separately true under Hypotheses A–C, yet '$(\tau = t_1) \cdot$

[27] It might hence be thought that Hypotheses A–C constitute a departure from classical logic, as well as from classical semantics, in that we cannot always correctly infer $S_1 \cdot S_2$ from S_1 and S_2. This difficulty is spurious, however, since Hypotheses B and C should properly be generalized to give the simultaneous significata and truth values of a *set* of sentences, instead of applying merely to the conjunction of the set (since one can accept a theory in the form of a set of postulates as well as in the form of their conjunction). We can then conjoin and separate sentences at will without change in their significata during a deduction from a given set of assumption formulas. On the

($\tau = t_2$),' which entails that $t_1 = t_2$, is false. But while this violates classical semantics, it makes a certain amount of intuitive sense upon reflection that while either '$T(\tau) \cdot (\tau = t_1)$' or '$T(\tau) \cdot (\tau = t_2)$' is a perfectly good (i.e., correct) enrichment of '$T(\tau)$' under the conditions stipulated, the stronger enrichment '$T(\tau) \cdot (\tau = t_1) \cdot (\tau = t_2)$' 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 t_1 and t_2 such that $T(t_1)$ and $T(t_2)$, but that $E(t_1)$ and not $E(t_2)$. Then theory '$T(\tau)$,' if accepted, is true, and 'τ' designates both t_1 and t_2. When it comes to enriching the theory in regard to whether or not τ satisfies '$E(\phi)$,' however, it would appear that we can have it *both* ways; the theory '$T(\tau) \cdot E(\tau)$' and the theory '$T(\tau) \cdot \sim E(\tau)$' 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 '$T(\tau) \cdot E(\tau) \cdot \sim E(\tau)$'? Or even if we do not try both directions at once, is not '$T(\tau) \cdot E(\tau)$' incompatible with '$T(\tau) \cdot \sim E(\tau)$'? The answer, of course, is that 'τ' has a *different meaning* in '$T(\tau) \cdot E(\tau)$' than it has in '$T(\tau) \cdot \sim E(\tau)$.' 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 '$T(\tau) \cdot E(\tau)$' is quite compatible with the enrichment '$T(\mu) \cdot \sim E(\mu)$.' 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.

directions at once, only this must be done, not by assertion of a logical absurdity, but by multiplication of theoretical terms; namely, '$T(\tau) \cdot T(\mu) \cdot E(\tau) \cdot \sim E(\mu)$.' The moral, here, is that as a theory becomes explored, tested, accepted, and elaborated, not only do we find the meanings of theoretical terms enriched, we may also expect to find—and, in fact, do find—at any stage of development that a theoretical concept has suddenly fissioned into a set of concepts which share a common core of meaning but which have now become free to evolve in their own individual ways.

If there is any important distinction between theoretical and observational terms (ignoring for the present that observational terms are themselves for the most part theoretical terms whose credentials we have come to accept at face value), it must lie in the dynamic aspects of the former. It is an intrinsic part of their usage that theoretical terms are to-be-enriched terms. A theoretical concept is not merely a "promissory note" in the sense of *permitting* future elaboration, it carries theoretical problems the resolution of which *demands* meaning enrichment, if only a provisional one. At any stage in the development of a theory, the theoretical terms then in use are enveloped in a penumbra of possible extensions and multiplications. Similarly, within the harder core of meaning imparted by the theoretical postulates actually believed to be true, lie the "lines of retreat,"[28] the order in which meaning components would be relinquished as the theoretical postulates were abandoned one by one under the press of disconfirming evidence. *Theoretical terms are concepts in the act of formation.*

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

[28] I am indebted to W. Sellars for this adroit phrase.

aspect[29] 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,

[29] 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., those linguistic properties 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.

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 '($\exists\phi$)T(ϕ)' do *not* signify the same fact; 'T(τ)' signifies each member of a (possibly empty) set of (in the main) nonobservational facts, each of which entails the observational fact, if any, signified by '($\exists\phi$)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 'τ' could be infused with just that degree of meaning which would make the behavioral force of 'T(τ)' essentially the same as that of '($\exists\phi$)T(ϕ)' when the latter already exists in the organism's behavioral repertoire. Presumably, this could be accomplished simply by using the same symbol 'τ' 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 'τ,' the combined effect should be the same as if 'τ' 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 '($\exists\phi$)T(ϕ),' if 'E(τ)' is entailed by 'T(τ)' but not conversely, the meaning of 'E(τ)' is *richer* than that of '($\exists\phi$)E(ϕ).' It will be recalled that one difficulty in regarding 'T(τ)' as a peculiar way of asserting that ($\exists\phi$)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 'τ' 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.

IV

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 'τ' 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 τ is!"), the analysis of which case is essentially that of (1); or (3) contains other theoretical terms,'μ_1,' . . ., 'μ_n,' in which case 'τ' has been "reduced" to 'μ_1,' . . ., 'μ_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 'τ' is a theoretical term whose meaning is defined by the theory 'T(τ),' and that 'd' is a descriptive term in the observation language. Under what circumstances would we be willing to say that d is the identity of τ—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 'τ' 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 τ? 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 'τ' 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 'τ' designates d does not necessitate the conclusion '$\tau = d$' is that the latter adds a further restriction on the designata of 'τ' beyond that imposed by the theory '$T(\tau)$.' *Assertion that d is the identity of τ 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 τ 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

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 'τ.' Under such circumstances, we should be tempted to identity τ 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 τ with $d\mu$ are as strong and as legitimate as identifying τ with some observational entity, d_o when it is known that $T_1(d_o)$.

But there is an important difference between observational identification and theoretical reduction. In the former case, we considered the legitimacy of asserting '$\tau = d_o$,' *given knowledge that* $T_1(d_o)$. 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 $T_1(d_o)$, 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 $\sim(\exists\phi)T_2(\phi)$ does not entail that $\sim(\exists\phi)T_1(\phi)$ unless T_1 and T_2 are analytically equivalent). That is, to identify τ 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 '$D(\alpha)$' (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 'α.' When a stipulative definition is of the form '$\alpha = d$' (or a conjunction of sentences of the form '$\alpha_i = d_i$'), it is known as an "explicit" definition.[30] When '$D(\alpha)$' is of a form *other* than '$\alpha = d$,' it is known as an "implicit" definition. Since any analysis of implicit definition sufficiently broad to cover all forms of '$D(\alpha)$' other than '$\alpha = d$' 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 'α' is determined (solely) by its usage in the sentence '$D(\alpha)$,' then 'α' designates every entity t such that $D(t)$. This account does not explain *how*

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

'α' acquires this meaning—such an explanation lies within the province of the psychology of language. What is of philosophical relevance is *that* 'α' designates in this way. Thus whatever role the symbol complex '$D(\alpha)$' may play in the acquisition by 'α' of meaning, the *semantical* status of '$D(\alpha)$' is simply that of a descriptive sentence. What needs to be spelled out in greater detail, however, are the truth conditions of '$D(\alpha)$': (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 '$D(\alpha)$' is true by virtue of its logical form only if the expression formed by replacing 'α' with any descriptive constant of the same formal type is necessarily true. This obtains only when the property $D(\phi)$ is necessarily possessed by every entity of the appropriate type, in which case (as will be elaborated below) '$D(\alpha)$' 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 'α' 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 '$D(\alpha)$.' It is in the meaning of 'α' that any entity designated by 'α' satisfies '$D(\alpha)$.'[31] On the other hand, it is *not* the case that '$D(\alpha)$' has no factual content, or that '$D(\alpha)$' 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 'α' designates any entity which satisfies '$D(\phi)$,' it by no means follows that there *is* any such entity. Hence, '$D(\alpha)$' is empirically true or false according to whether or not there exists an entity designated by 'α'—i.e., according to whether or not $(\exists\phi)D(\phi)$.

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

'$\alpha =_{\text{dat}} d$' (and some analogous expression in the case of implicit definition) should be treated as the "definition" proper, with '$\alpha = d$' (or '$D(\alpha)$') as a consequence of the definition, does not matter here so long as it is agreed that it is legitimate to analyze the force of an explicit definition in terms of a symbol 'α' being *given* meaning through its use in the sentence '$\alpha = d$.'

[31] Some complications may arise here if '$D(\alpha)$' 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

'The A is a B' appears to be of the logical form '$\psi(x)$,' in which 'x' designates the (only) entity which possesses a property ϕ, Russell contends that the genuine logical form is '$(\exists x)(\psi[x] \cdot (y)[\phi(y) \equiv y = x])$,' *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 '$(\exists x)(B[x] \cdot (y)[A(y) \equiv y = x])$' 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 '$\phi(x)$.'

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 '$T(\tau)$' is a theory, then while '$T(\tau)$' has the same *truth conditions* as '$(\exists\phi)T(\phi)$,' '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,[32] 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

[32] 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.

particular, if 'the A' is regarded as a theoretical term introduced by the unified theory '$(x)[A(x) \equiv x = \text{the A}]$,' 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 '$A(\text{an A})$.' Then by Hypotheses A–C, 'an A' designates every satisfier of '$A(x)$'; while the sentence 'An A is a B' is true if and only if $(\exists x)[A(x) \cdot B(x)]$, yet is of the logical form '$\phi(x)$.'

The meaning criterion. One of the dominant themes of modern analytic philosophy—certainly a guiding motive of the logical empiristic 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 'α' is imparted to it by its use in a set of (perhaps provisionally) accepted sentences, the conjunction of which is '$D(\alpha)$.' That is, '$D(\alpha)$' is the implicit definition of 'α.' Under what circumstances would we say that 'α' 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 '$D(\alpha)$' is false if and only if 'α' has no designatum—i.e., if and only if it is the case that $\sim(\exists\phi)D(\phi)$. Then if 'α' were meaningless when it has no designatum, (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 designata. 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 designata 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-design 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 'α' is defined by '$\alpha = \alpha$.' We should scarcely feel that 'α' 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 '$D(\alpha)$' contributes no useful meaning to 'α.' 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 'ϕ' is a purely logical variable, '$D(\alpha)$' gives 'α' effective meaning if and only if, when '$D(\alpha)$' is adopted, it is not the case that 'α' necessarily designates every entity in the range of 'ϕ'—i.e., if and only if '$(\phi)D(\phi)$' is not necessarily true. The reason for stipulating that 'ϕ' must be a logical variable is that if its range were a nonlogical category, it would not be logically decidable whether

'$D(\phi)$' is applicable to a given entity t even though the empirical fact that t is in the range of 'ϕ' logically entails that $D(t)$. Thus if 'x^h' is a variable ranging over humans, '$(x^h)(x^h = x^h)$' is logically true, but the implicit definition '$\alpha^h = \alpha^h$' gives 'α^h' effective meaning, since its designata 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 '$D(\alpha_1, \ldots, \alpha_n)$,' 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 '$D_1(\alpha_1)$' effective-meaningfully defines 'α_1,' that '$D_2(\alpha_2)$' is tautologous, and that '$D(\alpha_1, \alpha_2)$' is equivalent to '$D_1(\alpha_1) \cdot D_2(\alpha_2)$.' Then '$D(\alpha_1, \alpha_2)$' gives effective meaning to 'α_1' but not to 'α_2.' Here, as in the monadic case, 'α_2' is effectively meaningless since its designata remains unrestricted; yet '$D(\alpha_1, \alpha_2)$' is not necessarily true, and, if unified (cf. Definition 10), would give a designatum to 'α_2' only if it also provides one for 'α_1.' To be sure, '$D_1(\alpha_1)$' and '$D_2(\alpha_2)$' are undoubtedly autonomous subdefinitions (cf. Definition 9) of '$D(\alpha_1, \alpha_2)$' in this instance, and if so, permit '$D(\alpha_1, \alpha_2)$' to assign designata to 'α_2' whether '$D_1(\phi)$' 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 '$D(\alpha_1, \alpha_2)$' is unified or not, if there exists one pair of entities t_1, t_2 such that $D(t_1, t_2)$, then 'α_2' designates all entities in the range of its formal type, since for any entity t_i in that range, it is the case that $D(t_1, t_i)$. 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 designata 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.

Postulate 7. The term 'α_i' is given effective meaning by a logically consistent implicit definition 'D($\alpha_1, \ldots, \alpha_n$),' *where none of the definienda,* 'α_1,' ..., 'α_n,' *occur in the predicate* 'D($\phi_1, \ldots, \phi_n$)' *and* 'ϕ_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($\phi_1, \ldots, \phi_n$).'

The postulate may be applied to the case where 'ϕ_i' has extralogical restrictions on its range by first putting 'D($\alpha_1, \ldots, \alpha_n$)' 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 'α' is meaningless when defined by 'D(α),' and 'D(α)' is formally equivalent to 'D*(α),' then 'α' must also be meaningless when defined by 'D*(α)'; 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] '$D(\alpha) \equiv D^*(\alpha)$' 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 '$D(\alpha)$' formally entails '$(\exists\phi)D(\phi)$,' it follows from [a] that '$D^*(\alpha) \supset (\exists\phi)D(\phi)$' 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 '$F(\ \ )$' 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 meaninglessness 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 'τ,' '$T(\tau)$' 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, '$\alpha =_{def} d$,' is a perfectly legitimate way to confer meaning upon 'α.' 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 '$\alpha = d$' 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 '$D(\phi)$' contains meaningful descriptive constants, the sentence '$(\exists\phi)D(\phi)$,' even when analytic, contains effective meaning components carried by its descriptive terms which may then be mobilized to give effective meaning to 'α' when defined by '$D(\alpha)$.'

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

1. Bergmann, G. "A Note on Ontology," *Philosophical Studies*, 1:89–92 (1950). Reprinted in G. Bergmann, *The Metaphysics of Logical Positivism*. New York: Longmans, Green and Co., 1954.
2. Bergmann, G. "Sense Data, Linguistic Conventions, and Existence," *Philosophy of Science*, 14:152–163 (1947). Reprinted in G. Bergmann, *The Metaphysics of Logical Positivism*.
3. Carnap, R. *Introduction to Semantics*. Cambridge, Mass.: Harvard University Press, 1942.
4. Carnap, R. "The Methodological Character of Theoretical Concepts," in *Minnesota Studies in the Philosophy of Science*, Vol. I, H. Feigl and M. Scriven, eds. Minneapolis: University of Minnesota Press, 1956.
5. Feigl, H. "Existential Hypotheses," *Philosophy of Science*, 17:35–62 (1950).
6. Feigl, H. "The 'Mental' and the 'Physical,'" in *Minnesota Studies in the Philosophy of Science*, Vol. II, H. Feigl, M. Scriven, and G. Maxwell, eds. Minneapolis: University of Minnesota Press, 1958.
7. Hempel, C. "Problems and Changes in the Empiricist Criterion of Meaning," *Revue Internationale de Philosophie*, 11:41–63 (1950). Reprinted in *Semantics and the Philosophy of Language*, L. Linsky, ed. Urbana: University of Illinois Press, 1952.
8. Hempel, C. "The Theoretician's Dilemma," in *Minnesota Studies in the Philosophy of Science*, Vol. II, H. Feigl, M. Scriven, and G. Maxwell, eds. Minneapolis: University of Minnesota Press, 1957.
9. Henkin, L. "Completeness in the Theory of Types," *Journal of Symbolic Logic*, 15:81–91 (1950).
10. Quine, W. V. "Designation and Existence," *Journal of Philosophy*, 36:701–709 (1939). Reprinted in *Readings in Philosophical Analysis*, H. Feigl and W. Sellars, eds. New York: Appleton-Century-Crofts, 1949.
11. Ramsey, F. P. *The Foundations of Mathematics, and Other Essays*. London: Paul, Trench, Trubner and Co., 1931.

12. Rozeboom, W. W. "Do Stimuli Elicit Behavior?—A Study in the Logical Foundations of Behavioristics," *Philosophy of Science*, 27:159–170 (1960).
13. Rozeboom, W. W. "Intentionality and Existence," *Mind*, 71:15–38 (1962).
14. Rozeboom, W. W. "Studies in the Empiricist Theory of Scientific Meaning. Part I—Empirical Realism and Classical Semantics: A Parting of the Ways. Part II—On the Equivalence of Scientific Theories," *Philosophy of Science*, 27:359–373 (1960).
15. Russell, B. "On Denoting," *Mind*, 14:479–493 (1905). Reprinted in *Readings in Philosophical Analysis*, H. Feigl and W. Sellars, eds. New York: Appleton-Century-Crofts, 1949.
16. Russell, B. *The Problems of Philosophy*. New York: Henry Holt, 1912.
17. Scriven, M. "Definitions, Explanations, and Theories," in *Minnesota Studies in the Philosophy of Science*, Vol. II, H. Feigl, M. Scriven, and G. Maxwell, eds. Minneapolis: University of Minnesota Press, 1958.
18. Strawson, P. F. *Introduction to Logical Theory*. London: Methuen, 1952.
19. Tarski, A. "The Semantic Conception of Truth," *Philosophical and Phenomenological Research*, 4:341–375 (1944). Reprinted in *Readings in Philosophical Analysis*, H. Feigl and W. Sellars, eds. New York: Appleton-Century-Crofts, 1949.