The Semantic Conception of Theories and Scientific Realism

FREDERICK SUPPE

UNIVERSITY OF ILLINOIS PRESS
Urbana and Chicago

Publication of this work was supported in part by a grant from the Office of Graduate Studies and Research, University of Maryland.

Manufactured in the United States of America
C 5 4 3 2 1

This book is printed on acid-free paper.

Library of Congress Cataloging-in-Publication Data

Suppe, Frederick.
The semantic conception of theories and scientific realism.

Bibliography: p.
Includes index.
1. Science—Philosophy. 2. Science—Methodology.
3. Semantics (Philosophy). 4. Realism. I. Title.
Q175.S9393 1989 501 88-27878
ISBN 0-252-01605-X (alk. paper)

2

What's Wrong with the Received View on the Structure of Scientific Theories?

For many years the *Received View on Scientific Theories* has been that theories are to be construed as axiomatic calculi in which theoretical terms are given a partial observational interpretation by means of correspondence rules. Underlying this analysis is a strict bifurcation of the nonlogical terms of the theory into an observational vocabulary and a theoretical vocabulary. Putnam, Achinstein, and others have urged the rejection of the Received View because (i) the notion of partial interpretation it employs cannot be given a precise formulation adequate for the purposes of the Received View, and (ii) the observational/theoretical distinction cannot be drawn satisfactorily.[1] It is my contention that the Received View is unsatisfactory and ought to have been rejected, but not for these reasons. Section I of this chapter (based on my 1971 article, "On Partial Interpretation") argues that reason (i) is false. Section II goes on to argue that it is virtually impossible to establish reason (ii). Section III attempts to show that the Received View nonetheless ought to be rejected because its reliance on the observational/theoretical distinction causes it to obscure a number of epistemologically important and revealing features of the structure of scientific theories. In the process of arguing for this latter claim, a more adequate account of the epistemological structure of scientific theories is presented—it is a version of the *Semantic Conception of Theories.*

I. ON PARTIAL INTERPRETATION

Achinstein and Putnam argue in support of reason (i) for rejecting the Received View by observing that its advocates have not made the notion of partial interpretation clear. They consider a number of possible explications of the notion, then show that they are inadequate for the purposes of the Received View (Achinstein 1968, 85–91; Putnam 1962, 244–48).

In this section I attempt to present an analysis of partial interpretation compatible with the rest of the Received View. My interests in presenting this analysis do not lie in the direction of attempting to rescue the Received View from its critics—for I am convinced that ultimately it must be rejected as unsatisfactory—but rather are grounded in the belief that a clear understanding of the notion of partial interpretation employed by the Received View will bring to light a number of important characteristics about scientific theories and scientific meaning.

A. An Explication of the Received View

Explication of the notion of partial interpretation requires a precise formulation of the Received View, and since Carnap and Hempel have given its most extensive and sophisticated development, their formulation will be used.[2] Since their position has undergone considerable revision over the years, and because they have published no single comprehensive account of their ultimate formulation, I have reconstructed it from their various recent writings on the subject.

The reconstruction is as follows: Scientific theories are such that they can be given a canonical reformulation which satisfies the following conditions:

(1) There is a first-order language L in terms of which the theory is formulated, and a calculus K defined in terms of L.

(2) The nonlogical or descriptive primitive constants (i.e., the "terms") of L are bifurcated into two disjoint classes:
   (a) V_O, which contains just the observation terms, and
   (b) V_T, which contains the nonobservation or theoretical terms.
   V_O must contain at least one individual constant.

(3) The language L is divided into the following sublanguages, and the calculus K is divided into the following subcalculi:
   (a) The *observation language*, L_O, is a sublanguage of L which contains no quantifiers or modalities and contains the terms of V_O, but none from V_T. The associated calculus K_O is the restriction of K to L_O and must be such that any non-V_O-terms (i.e., nonprimitive terms) in L_O are explicitly defined in K_O. Furthermore, K_O must admit of at least one finite model.

(b) The *logically extended observation language, L_O'*, contains no V_T-terms. It may be regarded as being formed from L_O by adding the quantifiers, modalities, etc., of L to L_O. Its associated calculus, K_O', is the restriction of K to L_O'.

(c) The *theoretical language, L_T*, is that sublanguage of L which does not contain V_O-terms; its associated calculus, K_T, is the restriction of K to L_T.

These sublanguages together do not exhaust L, for L also contains *mixed sentences*, i.e., those in which at least one V_T- and one V_O-term occur. In addition, it is assumed that each of the sublanguages above has its own stock of predicate or functional variables and the L_O and L_O' have the same stock, which is distinct from that of L_T.

(4) L_O and its associated calculi are given a *semantic interpretation* which meets the following conditions:

(a) The domain of interpretation consists of concrete observable entities such as observable events, things, or thing-moments; the relations and properties of the interpretation must be directly observable.

(b) Every value of any variable in L_O must be designated by an expression in L_O.

It follows that any such interpretation of L_O and K_O, when augmented by appropriate additional rules of truth, will become an interpretation of L_O' and K_O'. We may construe interpretations of L_O and K_O as being *partial semantic interpretations of* L and K, and we require that L and K be given no empirical semantic interpretation other than that provided by such partial semantic interpretations.

(5) A *partial interpretation* of the theoretical terms and of the sentences of L containing them is provided by the following two kinds of postulates: the *theoretical postulates* T (i.e., the axioms of the theory) in which only terms of V_T occur, and the *correspondence rules* or postulates C which are mixed sentences. The correspondence rules C must satisfy the following conditions:

(a) The set of rules C must be finite.

(b) The set of rules C must be logically compatible with T.

(c) C contains no extralogical term that does not belong to V_O or V_T.

(d) Each rule in C must contain at least one V_O-term and at least one V_T-term essentially or nonvacuously.[3]

Let T be the conjunction of the theoretical postulates and C be the conjunction of the correspondence rules. Then the scientific theory based on L, T, and C consists of the conjunction of T and C and is designated by 'TC'.

Note that condition (4) allows the possibility of alternative semantical systems (or interpretations) for L_O which may differ in the designata of V_O-terms. The Received View intends, however, that there be a fixed set of designata for the terms of V_O, and so restrictions must be imposed on the class of admissible interpretations. Let us assume that a fixed set of rules of designation has been specified for these V_O-terms; then let us say that the class of semantical systems that use these rules are *permissible semantical systems* for L_O, and the class of interpretations they specify are *permissible interpretations* for L_O and K_O. Notice that different permissible interpretations are possible, since the rules of designation for predicate variables may differ. The classes of permissible semantical systems and interpretations for L'_O and K'_O are defined analogously.

B. A Formal Analysis of Partial Interpretation

A central claim of the Received View is that TC together with the specification of a permissible semantical system for L_O and L'_O provides L (and hence the V_T-terms and L_T-sentences) with a partial interpretation. As Carnap puts it:

> All the interpretation (in the strict sense of this term, i.e., observational interpretation) that can be given for L_T is given in the C-rules, and their function is essentially the interpretation of certain sentences containing descriptive terms, and thereby the descriptive terms of V_T. . . .
>
> For L_T we do not claim to have a complete interpretation, but only the indirect and partial interpretation given by the correspondence rules. . . .
>
> . . . Before the C-rules are given, L_T, with the postulates T and the rules of deduction, is an uninterpreted calculus. . . . Then the C-rules are added. All they do is, in effect, to permit the derivation of certain sentences of L_O from certain sentences of L_T or vice versa. They serve indirectly for derivations of conclusions in L_O, e.g. predictions of observable events, from given premises in L_O, e.g. reports of results found by observation, or the determination of the probability of a conclusion in L_O on the basis of given premises in L_O. [1956, 46–47]

The crucial notion here, partial interpretation, plays a central role in the Received View analysis, and as many authors have pointed out, the notion is far from clear.[4] For, as Putnam points out, nowhere have Carnap, Hempel, and the other proponents of the Received View defined what they mean by partial interpretation.

Hempel (1963) has given what is perhaps the most detailed explication of partial interpretation advanced by any proponent of the Received View. He raises the following question: In what sense and to what extent does an interpretative system (rules of correspondence) specify an interpretation for L_T? He points out that for a given term in V_T, the interpretative system C "may establish a necessary and sufficient condition" in terms of V_O, which he illustrates with an explicit definition. But he adds that this need not be the case, since for some terms, C will establish (1) "only a necessary and a different sufficient condition" in terms of V_O, (2) just a necessary condition, (3) a sufficient condition, or (4) "neither a necessary nor a sufficient condition" in terms of V_O (693). This, however, tells us little, since he never stipulates what these conditions are necessary or sufficient for—are they truth conditions, derivability conditions, or what?

Although Hempel doesn't explicitly address these questions, it is possible to construe his subsequent discussion, which focuses on possible ways in which partially interpreted theoretical sentences are significant, as constituting a somewhat oblique partial answer to this question (694–95). He distinguishes three concepts of significance: (a) pragmatic intelligibility, (b) empirical significance, in the sense of being relevant to potential empirical evidence expressible in terms of V_O, and (c) semantical significance, in the sense of being true or false. Further, he claims that partially interpreted theoretical sentences are significant in all three senses. With respect to (a) he says that scientists understand how to employ partially interpreted theoretical sentences in the sense that they know how "to use" them correctly, and that in the Received View reconstruction, this is, in essential respects, equivalent to knowing the rules of TC. As to (b), this condition is met, since TC has derivable sentences of L_O that may be used for prediction and are subject to empirical test. Thus a necessary condition for the existence of a partial interpretation seems to be that the C-rules must enable one to derive sentences in L_O from T that could not be derived otherwise. Hempel's argument that theoretical statements are semantically significant in the sense of being true or false, simply put, is that they can be stated in a suitable metalanguage, and so it is possible to specify truth criteria for theoretical statements. Although this tells us nothing about the way C-rules provide a partial interpretation, it does provide a certain indirect support for my analysis, which follows.

Although Hempel's discussion of significance sheds some light on what is meant by partial interpretation—telling us, at most, that partial interpretations must supply theories with observational or testable consequences—it does not provide an adequate analysis. Since it would appear that little more can be said syntactically in the way of characterizing partial interpretation, if we are to find an adequate analysis of the concept, we must turn to semantic considerations. As a first step, let us consider the semantics of explicit definition.

Consider a language L^* which is that sublangauge of L'_O whose only nonlogical symbols are the individual variables and the symbols in V_O, and let K^* be the restriction of K to L^*. Let S be a true interpretation of K^*—that is, a semantical system S for L^* such that every sentence of L^* that is a theorem of K^* is true under S. Suppose now that we augment the alphabet of L^* with the one-place predicate constant 'P', and we introduce a definition of 'P' in terms of V_O as an additional axiom for K^*. Call this new language and this new calculus L^{**} and K^{**} respectively. Under what circumstances will this definition qualify as an explicit definition?

Suppose that our definition is of the form

$$(1) \qquad (x)(\mathrm{P}x \equiv \phi(x))$$

where ϕ is a formula whose only predicates are from V_O, and 'x' is the only free variable in ϕ. Then, for (1) to qualify as an explicit definition of 'P', for any formula θ such that 'P' occurs in θ, if θ is a theorem of K^{**}, then the result ψ of replacing all occurrences of 'P' in θ by ϕ (with appropriate changes of variables) must be a theorem of K^*. That is, for 'P' to be explicitly definable, the introduction of 'P' must be noncreative and theoretically eliminable.[5] But this, in essence, is nothing other than the requirement that 'P' be definable in accord with the modern theory of definition in logic as developed by Padoa and Beth.[6] The key semantic features of this theory are summarized in Padoa's principle and its converse, the Beth definability theorem. Stated heuristically, for 'P' to be *explictily definable in terms of V_O in K^{**}*, it is necessary and sufficient that the following conditions be met: Let S_1 and S_2 be any two true interpretations of L^{**} such that S_1 and S_2 have the some domain and assign the same designata to terms in V_O. Let P_1 be the designatum of 'P' under S_1 and let P_2 be its designatum under S_2. Then 'P' is explicitly definable in terms of V_O in K^{**} if and only if for each member a of the domain (of both S_1 and S_2), a has property P_1 if and only if a has property P_2.[7]

This result admits of the following interpretation: Let S' be a true permissible interpretation of L'_O. Call the result S^* of deleting all relations, properties, functions, and so on, in S' that are not the designata of terms in V_O *the restriction of* S' *to* L^*. The class of *permissible interpretations for* L^* is the class of restrictions of S' to L^*, where S' is an arbitrary permissible interpretation for L'_O. Let S be a permissible interpretation for L^*. The result of adding a new property P to S and a new rule of designation that 'P' designates P is said to be a *permissible extension of* S *from* L^* *to* L^{**}. Then the Beth-Padoa result says that 'P' must have the same extension (be "true of" the same entities) in every true interpretation of K^{**} that is a permissible extension of S from L^* to L^{**}. This simply means that whenever one explicitly defines a term 'P' on the basis of V_O, if one assumes both that all the theorems of K^* are true and that the sentence defining 'P' is true, then the property designated by 'P' in any given permissible extension of a true interpretation will be extensionally equivalent to the property designated by 'P' under any other such interpretation. Taken in conjunction with the various assumptions made about permissible interpretations for L'_O, this result essentially says that the assumed truth of the definition of 'P' and K^{**} necessitates that the true permissible extensions to K^{**} of true permissible interpretations of K^* be indistinguishable on extensional grounds.

Now recall from Hempel's discussion, referred to earlier, that terms introduced by interpretative systems of C-rules generally fail to meet all the requirements imposed on explicit definitions. Thus, such terms generally will *not* be explicitly definable on the basis of V_O in the Beth-Padoa sense of definability. Since Padoa's principle and the Beth definability theorem supply us with necessary and sufficient conditions for the explicit definability of terms on the basis of V_O, they yield the following characteristic of partial interpretations: The designata of 'P' in true permissible extensions to K^{**} of true permissible interpretations of K^* will *not* be extensionally equivalent.

Despite this fact, however, the assumed truth of the sentence partially defining 'P' does in general impose considerable restrictions upon the class of true permissible extensions. For example, suppose the definition sentence is of the form

$$(\exists x)(\mathrm{P}x \lor (\phi x \;\&\; \psi x)) \tag{2}$$

where ϕ and ψ are different one-place predicates in V_O. Then any permissible semantical system for L'_O whose restriction to L^* is a true interpretation of K^* such that '$(\exists x)\mathrm{P}x$' or '$(\exists x)(\phi x \;\&\; \psi x)$' are also true, will qualify as a true permissible extension to K^{**} of a true permissible interpretation of K^*. But suppose that the definition sentence is of the form

$$(x)(\mathrm{P}x \lor (\phi x \;\&\; \psi x)) \tag{3}$$

Then only those extensions to K^{**} such that every individual either has P or else has both ϕ and ψ will be true, and the class of such extensions will be a proper subclass of those such that '$(\exists x)Px$' is true or '$(\exists x)(\phi x \;\&\; \psi x)$' is true. Accordingly, the class of defining sentences will impose differing restrictions on the class of true permissible extensions to K^{**} of true permissible interpretations of K^*.

Thus far this discussion has considered what happens under partial and explicit definition when a single-place predicate constant from L'_O is defined. The same basic treatment will generalize to the case where a finite set of predicate symbols is defined sequentially via introductive chains,[8] provided the following assumption is made (which has been tacit in the discussion thus far): the only permissible definitions are those which do not make K^{**} inconsistent. This treatment also extends more or less straightforwardly to the case where terms of V_T are given either explicit or partial definitions via C-rules on the basis of V_O. Here we consider those permissible extensions of K^* to K, where we treat the conjunction TC (of the sentences in the set T of theoretical postulates and the set C of sentences in the interpretative system, or C-rules) as being the defining sentence. The primary difference is that the entire set of terms in V_T is being defined simultaneously by TC.

For permissible extensions S of S^* for K^* to K we do not require that S and S^* have the same domain, but rather only that the domain of S contain the domain of S^*. This allows the possibility that the domain of S may contain both theoretical entities and observable entities. As long as K (whose nonlogical axioms are the sentences in T and C) is consistent, the assumed truth of TC will impose restrictions upon the class of true permissible extensions to K of true permissible interpretations of K^*. This in turn will impose restrictions on the class of relations, and so on, which qualify as designata for terms in V_T in such extensions. But from a different perspective, this simply means that the assumption that TC is true imposes restrictions on the class of permissible models of K. This, then, suggests that the sense in which the interpretative system C supplies L_T (and its associated calculus) with a partial interpretation is that it imposes restrictions on the class of permissible models for it. Furthermore, the C-rules must enable sentences of L_O (or possible L'_O) to be derivable in K.

C. The Analysis Evaluated and Defended

How adequate an analysis of partial interpretation is this? Since Carnap and Hempel have not revealed sufficiently what they mean by partial interpretation, one cannot be sure whether this is what they have in mind, but as far as I have been able to determine, the analysis just suggested is wholly compatible with what they have to say on the matter. However, Achinstein and Putnam have considered possible analyses of partial interpretation that are somewhat similar to this and have rejected them as inadequate for the purposes of the Received View. It will be illustrative to consider these to see whether the objections to them present a challenge to my own analysis and also to display further features of partial interpretation.

Achinstein (1968) suggests the following analysis which, so far as it goes, has some affinity with mine: "To speak of a term 'X' as partially interpreted might be to say that although the term has a meaning, in the sense of a semantical rule or explicit observational definition, only part of that meaning has been given."[9] This suggested analysis is rejected by Achinstein as not being consistent with the Received View, since it presupposes that the term in question has a meaning, in the sense that there is either an explicit semantical rule or explicit observational definition for it, and for theoretical terms this is denied. If the only kind of meaning a theoretical term can have is that supplied by explicit semantical rules or explicit observational definitions, then Achinstein is correct in rejecting this interpretation. But why can't the meaning partially given or captured by the *C*-rules be the pre-analytic meaning for the theoretical term or its translation in the ordinary or metalanguage of science (of which the *C*-rules provide a partial explication)? Viewed in this manner, the analysis Achinstein considers becomes the claim that the partial interpretation captures a part, but not all, of the pre-analytic meaning.

Now it might be objected that the proponents of the Received View would reject this suggestion on the grounds that it presupposes the sorts of meanings their program has tried to eschew—and once they would have made such an objection. Later it became unlikely that they would so object, for consider what Hempel (1963) has written:

> To turn, finally, to the question of semantic significance: Let T be interpreted by a system C which does not furnish every V_T-sentence an equivalent in terms of V_O. Then it is nevertheless quite possible to provide a necessary and sufficient condition of truth for every sentence expressible in terms of the theoretical vocabulary. All that is needed for the purpose is a suitable metalanguage. If we are willing to use a metalanguage which contains V_O, V_T, and C, or translations thereof, then indeed each V_T-sentence has a truth criterion in it, namely simply its restatement in, or its translation into, that metalanguage.[10]

Thus it would appear that my suggestion is quite compatible with Carnap's and Hempel's published views.

Achinstein's proposal is of considerable value, because when altered and taken in conjunction with Hempel's statement just quoted, it will allow a more perspicuous treatment of partial interpretation. But first we must consider Hempel's statement further. At first blush it is an extremely surprising statement which seems to go contrary to the basic tenets of the Received View. For does it not seem to give ontological status to the very abstract entities the Received View regards with such suspicion that it wishes to avoid them? Indeed, one doubts that the assertion is even consistent with the remainder of the Received View analysis.

This first-blush reaction is, I think, quite understandable but cannot be sustained under analysis. To see this we need to remember that the Received View is offered as a reconstructive account of scientific theories and, in particular, that its notion of partial interpretation is intended to provide a reconstructive account of the *empirical* or *observational meaning* of theoretical terms. If we look at theoretical terms such as 'electron' in the ordinary scientific language, we find that much of the content of the concepts these terms embody is related in no explicitly discernible manner to the observational or empirical manifestations of, say, electrons, but does concern extra-empirical associations. For electrons, these might include various features of the billiard-ball model, various classical intuitions about macroscopic point-masses, and so on.

Such associations contribute to the meanings of the theoretical terms in ordinary scientific language, and it is quite likely that without them little scientific progress would be made. But despite their legitimate place as meaning constituents of theoretical terms, they need not—and usually do not—have empirical or observable or testable consequences. The proponents of the Received View do not deny this point, and they need not, for all they are committed to doing is presenting a reconstructive analysis of theoretical terms so as to explicate the ways in which they have empirical content, and the ways in which the empirical aspects of their meanings interrelate with observation reports or statements and with empirical laws. They need not be committed to the position that their analysis exhausts all the conceptual or meaning content of the corresponding ordinary scientific-language terms, or to the position that statements about electrons, for example, are neither true nor false. They certainly are not committed to the position that when scientists talk about electrons, they are speaking metaphysical gibberish, but only to the position that when scientists talk about electrons, only part of the meaning of 'electron' is empirical.[11]

Accordingly, although committed to including only the empirical aspects of meaning of these theoretical terms in their treatment of meaning specification for theoretical terms via *C*-rules or interpretative systems, the Received View can tolerate nonempirical meaning components. The resolution of the seemingly paradoxical or contradictory claims made in the Hempel quotation thus lies in the fact that by using the ordinary scientific-language meanings of theoretical terms, we can specify a semantic interpretation for *TC*—stipulating that, for example, the domain includes neutrons, electrons, protons, and so on, and that the term in V_T that corresponds to 'is an electron' denotes the class of objects in the domain that are electrons. At the same time we can refuse to allow the terms of V_T that are partially interpreted to be given an independent empirical semantic interpretation. That is, the semantic interpretations allowed in the Hempel quotation are allowed simply because they are not empirical interpretations.

But does not Hempel's claim now reduce to the trivial assertion that it is possible to give an unofficial extra-empirical semantic interpretation of the terms of V_T, but that officially such interpretations are banned? If so, have we not misinterpreted Hempel's claim? Possibly, but I think not—for Hempel continues by saying:

> Let us note here with Carnap that the semantical criteria of truth and references which can be given for the sentences and for the terms, or "constructs," of a partially interpreted theory offer little help towards an understanding of these expressions. For the criteria will be intelligible only to those who understand the metalanguage in which they are expressed; and the metalanguage must contain either the theoretical expressions themselves or their translations; hence, the latter must be antecedently understood if the semantical criteria are to be intelligible. [1963, 696]

Although this passage seems to me to contribute something to the plausibility of the interpretation I have advanced, it still is quite possible that I am misinterpreting Hempel. But if so, I find it impossible to see a reading of his position which does not deviate fundamentally from what I have just suggested and which does not at the same time make Hempel's assertions violently contradict the prohibition against providing *TC* with a direct empirical semantic interpretation (condition [4] of sec. I-A).

Assuming that my interpretation of Hempel is more or less accurate, does it not follow that Hempel's claim that partially interpreted theoretical statements are significant in the sense of being true or false is now beside the point? For is not the real question whether partially interpreted theoretical sentences are true or false when reconstructed in accord with the defining conditions (1) through (5) of the Received View (sec. I-A)? In deciding this it will be helpful to recall the restatement made earlier of Achinstein's possible analysis of partial interpretation: For a term in V_T, the *C*-rules specify part, but not all, of the meaning of the corresponding theoretical term in ordinary scientific language. Thus advocates of the Received View refuse to commit themselves to the existence of entities that have *all* the attributes specified by the meanings of ordinary scientific-language theoretical terms. However, so long as any *TC* in which the corresponding terms in V_T occur has some observable consequences, it seems highly unlikely that they would want to maintain that one can simultaneously assert *TC* and refuse to be committed to the existence of *some* entities that have the properties specified by *TC* and also lead to these observational consequences.[12] Admittedly, even then they would likely refuse to say anything more about these entities other than that they have these observable consequences. If this is the case, and if we accept—as I think we must—this analysis of partial interpretation as telling at least an important part of the story about partial interpretation, then the obvious conclusion is that in ordinary scientific discourse about electrons, and so on, ontological commitments are being made that commit one to terms such as 'electron' having referents. But this commitment is limited to the extent determined by the partial interpretation supplied for the V_T-terms by *TC;* that is, one is committed to the existence of *something* that has the specified observational consequences.

This last observation provides a basis for supplying Hempel's quoted assertion with an interesting and nontrivial construal. To assert *TC* and then give it a semantic interpretation in the manner he specifies is to commit oneself to the existence of a true permissible interpretation of *TC;* that is, to the assertion that one of the interpretations within the class circumscribed by the partial interpretation of *TC* is in fact true. But at the same time, one is *not* thereby committed to being able to specify on empirical grounds which of these interpretations one is committed to. Differently put, this commitment requires that one be able to specify empirically a semantic interpretation modulo-equivalence by virtue of being within the class of true permissible interpretations, but does not commit one to being able to specify it further on empirical grounds (except possibly by the addition of further *C*-rules). Viewed in this way, I do not find Hempel's assertion trivial, since it illuminates the way in which ordinary scientific language carries ontological commitment and the extent to which it can be used to specify an empirical semantic interpretation for the sentences of *TC*, and further, it illustrates how it is that partial interpretation via the observational consequences of a theory modulates the extent and specificity of that ontological commitment.

Putnam considers a suggested analysis of partial interpretation which is a less specific version of mine and claims that it is inadequate for the purposes of the Received View. This serious challenge to my analysis is worth looking at in detail. Putnam suggests that to partially interpret L_T is to specify a nonempty class of intended models and that if the class contains just one member, the interpretation is complete; otherwise it is partial (1962, 245). He then rejects this analysis as inadequate, arguing roughly as follows: To specify the class of intended models, one will have to employ theoretical terms (presumably from the metalanguage), and to rule out flagrantly unintended interpretations, one must use such notions as 'physical object' and 'physical magnitude' to specify the domains and designata of V_T-terms. For example, to interpret 'mass' as a real-valued function, the function's values must be restricted to physical magnitudes if those interpretations which a realistically minded scientist would reject are to be ruled out. And terms such as 'physical magnitude' are neither theoretical nor observational terms, but rather "broad-spectrum terms" which are *not* defined in advance, science itself telling us what "physical magnitudes" are. In short, although these are not theoretical terms, they eventually tend to acquire technical senses via theoretical definitions (Putnam 1962, 246).

Although Putnam takes these observations as demonstrating that his proposal is inadequate for the Received View's purposes, he never explains how it is that they bear upon the adequacy of his proposal—and it is neither clear nor immediately obvious that they do demonstrate its inadequacy. Presumably the inadequacy lies in the fact that terms such as 'physical magnitude' are broad-spectrum terms not defined in advance, their scope determined by science, with many changes of opinion along the way. But if science can tell us at a given time what the scope of such a term is, why cannot science use these terms to specify the intended class of models or interpretations for a theory? Surely not because the scope of a term may later change, for presumably such a change will be reflected in a change in theory which may alter the class of intended interpretations. Perhaps Putnam's point here is that what is meant by, for example, 'physical magnitude' will be determined in part by the theory itself, and so if we specify the models for L_T in terms of such notions, this will be rather unenlightening to anyone who does not understand the theory beforehand. But if this is so, it will not be taken as a telling criticism by proponents of the

Received View, for as evidenced by the Hempel quotation interpreted earlier, this is precisely what they themselves claim.

In what way might Putnam's observation lead to a telling objection against his proposed analysis? It will be helpful to use my own suggested analysis of partial interpretation as a foil, for Putnam's comments do suggest a way in which it may be inadequate. Recall that any possible extension of an S for L^* such that TC is true under it is allowed as a true permissible interpretation. Putnam's discussion suggests that this will allow the inclusion of flagrantly unintended interpretations that any realistically minded scientist would reject. Should we not impose restrictions on this class of permissible interpretations so as to exclude these? Perhaps, but it is not clear that this actually can be done—for although we (who understand the theory) have some idea of what we would count as reasonable and unreasonable interpretations, often we have no idea whether the interpretation in question is reasonable or not. Furthermore, in that case we are unwilling to decide by fiat whether it is or not, perhaps basing our refusal on the grounds that this is an open question, and so we must remain undecided. It follows, then, that we cannot appeal to the theory or the scientific metalanguage in order to specify what will count as a reasonable interpretation in any way that will make the class of intended interpretations well defined. Accordingly, it is impossible to characterize the class of intended interpretations in any acceptable way that rules out the flagrantly unintended interpretations, since we never will be certain which they are.

These observations suggest what may be the brunt of Putnam's observations: The reason we are not always able to specify whether certain interpretations are intended or not is that any specification of intended interpretations must employ broad-spectrum terms that are open-textured. If this is so, then presumably Putnam's charge is that the Received View cannot adequately specify the class of intended interpretations; hence the proponents of the Received View cannot mean by partial interpretation the analysis he has suggested. Read in this fashion, I think that Putnam's objections tell.

If Putnam's objections against his suggested analysis are telling in this way, what consequences do they have for my own proposed analysis of partial interpretation? First, my analysis differs from Putnam's in one crucial respect: It does not demand that all the permissible interpretations be reasonable in the sense of capturing what realistically minded scientists would intend. It purports to specify only a class of interpretations which, while it includes the class of intended interpretations, also includes any interpretation that is empirically compatible with the assumed truth of *TC;* and it is quite likely that some of these interpretations will be flagrantly unintended. Or, differently put, my analysis purports to specify the class of intended interpretations only insofar as they can be specified in terms of empirical or observational consequences—and of course, this is all that partial interpretation is supposed to do.[13] What Putnam's analysis (under my construal of it) shows is *not* that my analysis is inadequate; but rather, given the limited interpretative objectives of partial interpretation, my analysis would be inadequate if partial interpretation were to be analyzed so as to completely specify the class of intended interpretations, since such a precise specification is impossible. Thus, it partially vindicates my analysis by showing that it is as strong as possible.

Nevertheless, does it not seem that Putnam's objections show that the notion of partial interpretation is bankrupt? I think not, for as I suggested earlier, the Received View is concerned only with using partial interpretation to specify the extent to which theoretical terms have empirical meaning content and the ways this content relates to L_O-statements. Given this limited objective, there is no reason why partial interpretation should do any more than specify how empirical meaning content imposes restrictions on the class of possible interpretations; in particular, these restrictions need not rule out all the unintended interpretations, since whether an interpretation is intended or not depends in part on considerations that are not of an empirical sort.[14]

Putnam raises one further objection to his analysis of partial interpretation which requires consideration. He makes the claim that theories with false observation consequences have *no* model "standard" with respect to the observation terms. This difficulty is unacceptable, since we normally would say under such circumstances that such a theory is wrong, not senseless (1962, 247).

This objection is serious but can be avoided. In general, we do not have sufficient observational data to give a completely explicit semantic interpretation to L_O. Rather, we proceed roughly as follows: First we specify rules of designation by stating that the predicate constant ψ will designate, say, the property of being red. Second, we specify the rules of truth for sentences ϕ of L_O as follows: The sentence ϕ is true under the interpretation if and only if the situation, event, or whatever, that ϕ describes under the specified rules of designation actually obtains. No more precise a statement of the rules of truth can be given, since in general we do not have sufficient information abut what situations do or do not obtain. But notice that, employing the same rules of designation, we could specify other interpretations where situations obtain that are different from those in the real world. For example, we could specify a world in which objects that are red in the real world are green. In particular, given the specified rules of designation, we could specify a world such that a situation obtains in that world if and only if it is described under the rules of designation by a provable L_O-sentence of TC. Then we can use this interpretation to obtain the interpretation of K^* and, hence, to define the class of true permissible interpretations of TC. So the set of permissible models specified by partial interpretation will always be nonempty, and so the possibility Putnam considers can never occur.[15]

Under this proposal we employ the following characterization of what it is for a theory TC to be *empirically true*. The L_O-theorems are interpreted as being true statements about some possible world which is similar to the real world, and this world is specified jointly by TC and the rules of designation. Using the same rules of designation, the sentences of L_O are interpreted in terms of the real world, and the truth conditions are specified in terms of what situations actually obtain there. Thus we have two interpretations of L_O: one, determined by the theory TC, and the other, determined empirically. Then TC is empirically true if and only if these two worlds are identical—that is, TC is empirically true just in case the following condition is met: For each wff ϕ of L_O, ϕ is true under the one interpretation if and only if ϕ is true under the other interpretation.

This characterization of empirical truth for theories seems to me to be in close accord with what actually happens in the exact or formalized sciences, and so I find it an acceptable formulation of the notion. (A more sophisticated version of it will be developed in chapter 3.) This construal in turn points out some important differences between semantic truth and empirical truth for formalized theories. Since it is possible that TC can be semantically true without being empirically true, it follows that the theory TC can have empirically false consequences, yet be meaningful under my analysis of partial interpretation. In such cases, the possible state of affairs truly described by TC will not be the state of affairs that actually will be observed. Putnam's objection thus fails to show that my analysis is unsatisfactory.

One might object to the way I have avoided Putnam's objection with the argument that since TC in general will be undecidable, my specification of the interpretation determined by TC will be nonconstructive. This is, of course, true; but I think it is an unsatisfactory objection, since in practice, the specification of the interpretation of L_O in terms of situations that actually obtain will be equally nonconstructive. In general, the specification of truth conditions will cover situations where we do not yet know what actually obtains, and since the truth conditions are specified in terms of what in fact does obtain there, the specification will be no more constructive than that determined by *TC*.

D. *Partial Interpretation Summarized*

This discussion of partial interpretation concludes with a summarization of my main lines of argument and my findings. First, recall that the *C*-rules together with *T* partially interpret L_T in the sense that their assumed truth imposes restrictions on the class of permissible extensions of true interpretations of K^* by imposing extensional restrictions on the relations, and so forth, that are admissible designata for V_T-terms. These restrictions are not sufficient to circumscribe the class of intended interpretations for *TC;* rather, they only specify the class of intended interpretations to the extent possible on the basis of empirical or observational considerations. The complete specification of the class of intended interpretations would require an appeal to extra-empirical considerations, and in general, a complete specification cannot be given, since many of these additional considerations constitute open questions for science. In addition, although the specification of permissible interpretations for *TC* captures the empirical meaning components of theoretical concepts, these concepts also involve extra-empirical meaning components.

Further, partial interpretations also reveal the extent of one's ontological commitment to theoretical entities: that entities exist that have the observable manifestations specified by *TC*. Although in asserting a theory one is committed to some permissible interpretations being true, one is not committed to being able to specify which ones they are. One must only be able to specify which are true to the extent made possible by *TC* and perhaps also on the basis of the extra-empirical meaning components. Even if recourse is made to such extra-empirical considerations, it will not be possible to specify the true interpretation of the theory to the exclusion of all other permissible interpretations.

This construal of partial interpretation suffices to refute Achinstein's and Putnam's reason (i) for rejecting the Received View (stated at the outset of this chapter): that the notion of partial interpretation cannot be given an adequately precise formulation.

II. THE OBSERVATIONAL/THEORETICAL DISTINCTION

Achinstein's and Putnam's second main reason (ii) for rejecting the Received View is that the observational/theoretical distinction cannot be drawn satisfactorily. The arguments Achinstein (1968, chs. 5 and 6) and Putnam (1962, 240–44) advance in support of reason (ii) attempt to show that

(a) the observational/theoretical distinction cannot be drawn on the basis of the ordinary usage of scientific terms.

Of course, (ii) follows from (a) only if the further assumption is made that

(b) to be tenable for the purposes of the Received View, the observational/theoretical distinction must be drawn on the basis of the ordinary usage of scientific terms.

This latter assumption is neither made explicit nor argued for in either work; as such, Achinstein and Putnam have not made their case. But I wish to establish something stronger—namely, that (a) is true, whereas it is virtually impossible to establish (b). However, I do not want to base my claim that (a) is true on their arguments, for I do not find them wholly satisfactory: Achinstein's arguments only show that the observational/theoretical distinction cannot be drawn on the basis of ordinary usage in the ways Carnap and others have suggested, and so they establish a conclusion that is weaker than (a); Putnam's arguments contain numerous lacunae. Rather, I will refine the sorts of considerations they raise into a much tighter and stronger argument for (a); then I will use features of that argument to argue that (b) is virtually impossible to establish.

Condition (2) of my reconstruction of the Received View stipulates that the nonlogical terms of L be bifurcated into two disjoint classes—observation terms and theoretical terms. Since this distinction lies at the heart of the Received View analysis, one would expect that in advancing the Received View, its proponents would have extensively discussed the nature of this bifurcation and the basis upon which it is drawn. In fact, however, all one usually finds in the literature is discussion of a very few examples of what would count as observation terms and what would count as theoretical terms. The most extensive discussion I have found of the observational/theoretical distinction by a proponent of the Received View is in Carnap's *Philosophical Foundations of Physics* (1966).

Carnap begins by stating that "the term 'observable' is often used for any phenomenon that can be *directly observed*" (225, emphasis added). He then observes that this use of 'observable' is not that of the scientist, and that he intends to use the term in a very narrow sense "to apply to such properties as 'blue', 'hard', 'hot'," and so on, which are "properties directly perceived by the senses" (ibid.). In defending his somewhat special sense of 'observable' he says:

> There is no question of who [the philosopher or the scientist] is using the term "observable" in a right or proper way. There is a continuum which starts with direct sensory observations and proceeds to enormously complex, indirect methods of observation. Obviously no sharp line can be drawn across this continuum; it is a matter of degree. . . . In general the physicist speaks of observables in a very wide sense compared with the narrow sense of the philosopher, but, in both cases, the line separating observables from non-observables is highly arbitrary. [ibid., 226]

Thus far Carnap is discussing the use of the term 'observable' and its application to attributes, things, events, objects, and so on. He claims that he is using the terms to apply to those such attributes and entities that can be directly perceived by the senses. From this it follows that attributes, entities, events, objects, and so on, are to be divided into two classes—the observable and the nonobservable. In terms of this distinction Carnap bifurcates the nonlogical constants of L: "The terms of V_O are predicates designating observable properties of events or things (e.g. 'blue,' 'hot,' 'large,' etc.) or observable relationships between them (e.g. 'x is warmer than y,' 'x is contiguous to y,' etc.)" (1956, 40). On the other hand V_T contains theoretical terms often called 'theoretical constructs' or 'hypothetical constructs', which are intended to refer to such entities as electrons and their attributes. The vocabularies V_O and V_T constitute an exhaustive bifurcation of the nonlogical constants of L into the class of those which refer to observable attributes or entities, and the class of those which refer to nonobservable or theoretical entities or attributes.

Carnap apparently believes that the bifurcation into V_O and V_T can be drawn on the basis of the standard usages of nonlogical terms in, e.g., scientific English. For example, Carnap (1966) writes:

> For many years it has been found useful to divide *the terms of a scientific language* into three main groups.
>
> 1. Logical terms, including all of mathematics.
> 2. Observational terms, or *O*-terms.
> 3. Theoretical terms, or *T*-terms (sometimes called "constructs").
>
> It is true, of course, . . . that no sharp boundary separates the *O*-terms from the *T*-terms. The choice of an exact line is somewhat arbitrary. From a practical point of view, however, the distinction

> is usually evident. Everyone would agree that words for properties, such as "blue," "hard," "cold," and words for relations, such as "warmer, "heavier," "brighter," are *O*-terms, whereas "electric charge," "proton," "electromagnetic field" are *T*-terms referring to entities that cannot be observed in a relatively simple, direct way. [259; emphasis added]

Thus V_O will contain all those terms of a natural scientific language, such as scientific English, which in their normal usage refer to observables, and V_T will contain all the nonlogical terms of that language which refer to nonobservables in their normal usage; moreover, V_O and V_T are jointly exhaustive of the nonlogical terms of the language *L* (see condition [2] in sec. I-A).[16] The Received View thus seems to presuppose that a bifurcation of the nonlogical terms of a natural scientific language (such as scientific English) into theoretical and observational terms can be drawn on the basis of ordinary usage. Of course, it remains to be seen whether it is necessary for the Received View to make that presupposition.

Although Carnap usually does not make it explicit, it is obvious from this discussion that the observational/theoretical bifurcation is a dual dichotomy. First, there is a bifurcation of entities, properties, and so on, into those which are capable of direct observation and those which are not. Second, the terms in natural languages of science (such as scientific English) are bifurcated into two disjoint classes—the observational terms and the theoretical terms. These two bifurcations must parallel each other in the sense that a term may be included among the observational terms just in case it is used only in reference to directly observable attributes or entities. Or, differently put, the bifurcation of terms is drawn on the basis of the bifurcation of attributes and entities into the directly observable and the nondirectly observable. (In case an artificial language *L* is used, in setting up *L* we divide the nonlogical constants of *L* into those terms which are allowed to stand for, abbreviate, or correspond to observational terms of, say, scientific English, and those which are not.)

Is the observational/theoretical dichotomy a viable one? The answer to this question turns on what answers can be given to the following three subsidiary questions: (A) Is it possible to dichotomize entities and attributes on the basis of whether they are directly observable or not, and if so, what will be the nature of the dichotomy? (B) Is it the case that terms of, for example, scientific English under normal scientific usage can be bifurcated into the observational and the theoretical? (C) If the answers to the first two questions are affirmative, then are the two bifurcations coextensive?

A. The Entity and Attribute Dichotomy

Question (A). Carnap suggests that the property of being blue is a paradigmatic observable property, in the sense that its presence is directly ascertainable without recourse to complicated apparatus. But this is too imprecise. Is he claiming that a property is observable if its presence *sometimes* is ascertainable by direct observation? Or must it *always* be so ascertainable?

If he intends the latter, then the property of being blue is not directly observable: Although in some cases I can directly ascertain whether things are blue, when objects are too small it becomes impossible. Similarly, consider another of Carnap's paradigm examples, 'being warmer than'. While in some circumstances I can directly ascertain that something is warmer than something else (e.g., the water in the shower before and after adjusting it), there are numerous other circumstances in which I cannot because my sensory apparatus will not function at the temperatures involved (e.g., for an object at −250°C, which is warmer than an object at −273°C).

To take another example, I cannot directly observe that one part of the sun is warmer than another. Since direct observation precludes recourse to elaborate instrumentation, reliance on spectrographic evidence, and so on, in order to directly observe that one part of the sun is warmer than another I would have to be at those parts of the sun and compare their warmth. But this is humanly impossible in the following sense: Minimally, I would have to wear protective clothing, but in that case I would be directly observing that the air inside my space suit was warmer when I was at one part of the sun than when I was at another—not that one part of the sun is warmer than another. The latter could be determined only indirectly—for example, by using known heat transfer properties of my protective dress. To conclude: Even though it may be possible to determine by direct observation whether a particular attribute obtains under certain circumstances, the same attribute often will obtain in circumstances where it is impossible in principle to determine whether it does or does not obtain.

Since Carnap takes attributes such as being blue and being warmer than as paradigmatic examples of directly observable attributes, it follows that it is not necessary that one should be able in principle to ascertain by direct observation whether a directly observable attribute obtains in *every* circumstance in which it could obtain; rather, it is required that there be *some* circumstances in which it is possible in principle to ascertain by direct observation whether the attribute obtains. Since there is a strict bifurcation of attributes into the directly observable and the nondirectly observable (which for convenience we call nonobservables), it follows that an attribute is nonobservable if for every circumstance in which it could obtain, it is in principle impossible to ascertain by direct observation whether it obtains. Thus, for example, the property of being a gas must be directly observable, since it is possible to directly observe the presence of certain gases under certain circumstances (e.g., I can smell sulfur gas). And being electrically charged is directly observable, since by sticking my finger into a socket I can directly observe the presence of the electrical charge. Similarly, static electricity, forces, acceleration, gravitational attractions, and so on, would qualify as directly observable since we sometimes can directly observe their presence.[17] But this is clearly unsatisfactory, since we now are forced to count as directly observable various attributes which, according to Carnap, clearly should count as nonobservable.

To summarize: If we require that an attribute's presence *always* must be ascertainable in principle by direct observation in order for it to qualify as directly observable, then the paradigmatic ones (such as the property of being blue) fail to qualify. Further, if we require only that their presence *sometimes* be so ascertainable, then paradigmatic non-directly observables (such as the property of being a gas) become directly observables.

The problems encountered in attempting to draw a line between observable and nonobservable properties, and so forth, stem from the fact that many attributes of scientific relevance have both directly observable and nondirectly observable occurrences, which makes any natural division into the observable and the nonobservable impossible. If an observational/nonobservational distinction is to be drawn, perhaps it ought to be drawn on the basis of occurrences of attributes rather than on the basis of attributes simpliciter. Then, perhaps on the basis of the limits in discrimination of human sensory apparatus, we could count this or that attribute-occurrence observable or not. For example, we might say that for objects between such and such dimensions, the attribute of one being longer than the other is a directly observable attribute-occurrence; but if the objects are of larger or smaller size, then in that instance the attribute of being longer than is not a directly observable attribute-occurrence. Assuming this can be done in a sufficiently precise and general manner (which is by no means obvious), it would then be possible to distinguish observable occurrences and nonobservable occurrences of attributes and entities.

What we are doing here, in effect, is defining two new attributes (e.g., *observable-red* and *nonobservable-red* in terms of the old attribute (e.g., *red*) and replacing the old one with the two new ones. Thus we would say that the barn has the property of being *O*-red (observable-red), whereas the microscopic blood speck is *N*-red (nonobservable red). This, of course, has rather unusual consequences. If I take an *O*-red object of minimal area and smash it to pieces, the pieces will not be *O*-red, but rather *N*-red. And if I combine together a number of *N*-red blood specks, I will obtain an *O*-red blood patch. More complicated situations are encountered in the case of relations. If I heat an object at t' to a certain degree, it will be *O*-warmer than it was at time t, but if I heat the object still more at t'', it may be that the object is too hot for '*O*-warmer than' to apply. In such a case the object presumably will be *N*-warmer at t'' than it was at t. These examples show why some provision will have to be made for allowing comparisons between nonobservable and observable occurrences of properties and also for comparative relations whose applications straddle the observable/nonobservable boundaries.

It is not clear whether this proposal is workable, but it is clear that it will be rather complicated if it is. It is equally clear from the considerations raised above that some such division of attribute-occurrences into the observable and the nonobservable is required if we are to obtain an observational/nonobservational dichotomy for attributes which at all resembles the one Carnap requires.

B. The Term Dichotomy

Question (B). Since the observational/theoretical bifurcation of terms can be drawn along Carnap's lines only if a bifurcation of properties, and so on, into the observable and nonobservable can be drawn satisfactorily, let us assume that the dichotomy has been drawn along the rough lines previously suggested. Is it then the case that a natural bifurcation of terms can be drawn on this basis?

The linguistic analogues to the problems raised above now confront us. For we can use paradigmatic observation terms such as 'blue' or 'is warmer than' to refer to both observable and nonobservable occurrences of properties. (In the previous section's discussion [sec. II-A] we used such terms in precisely this way.) Thus we are faced with two choices: We may employ the terms in their natural uses (in which case observational terms sometimes have nonobservable referents and theoretical terms sometimes have observable referents) or we may adopt special uses (say 'red$_O$' and 'red$_T$') together with the rule of usage that the former may be used to refer only to observable occurrences of red and the latter only to nonobservable occurrences. The latter option will require introducing rather complicated semantic rules into the language, including rules which enable us to use 'red$_O$' and 'red$_T$'

comparatively. Whether sufficiently precise and general rules can be specified is not clear.

C. *Are the Dichotomies Coextensive?*

Question (C). Turning to the question whether the two bifurcations discussed in sections II-A and II-B are coextensive, the previous discussions lead immediately to the following conclusion: On the basis of ordinary linguistic usage, there is no natural bifurcation of terms into the observational and the theoretical which is coextensive with any reasonable distinction between either observable and nonobservable attributes and entities, or observable and nonobservable occurrences of attributes and entities. Only in an artificial or reconstructed language *L* could the distinction be drawn naturally. Hence the truth of claim (a) (see above) follows, that the observational/theoretical distinction cannot be drawn on the basis of the ordinary usage of scientific terms.

It is notorious that Carnap and most other proponents of the Received View have had little respect for ordinary usage as an instrument of precision in philosophical analysis, and so it would seem that the truth of claim (a) should not bother them, so long as a viable observational/theoretical distinction can be drawn in some other way. For if this is possible, their mistake in supposing that it could be drawn on the basis of ordinary usage does not in itself seriously jeopardize the tenability of the Received View. Can the observational/theoretical distinction be drawn in some other way? That is, what is the status of claim (b) (i.e., that to be tenable for the purposes of the Received View, this distinction must be drawn on the basis of the ordinary usage of scientific terms)?

The underlying motivation for the observational/theoretical distinction is the idea that statements which describe what can be directly observed are relatively nonproblematic as to truth, whereas those which describe what cannot be directly observed are more problematic as to truth. Moreover, the verification of scientific theories must ultimately rest on the nonproblematic evidence supplied by the senses. Accordingly, any observational/theoretical distinction which reflects the division of nonlogical statements into those which can be directly verified on the basis of the senses and those which cannot should be acceptable to proponents of the Received View. In particular, a dichotomy of terms which parallels what we intuitively would accept as a bifurcation of occurrences of attributes and entities into the observable and the nonobservable should be satisfactory for the purposes of the Received View. The discussion so far makes it clear that such a distinction will have to proceed roughly along the lines sketched earlier in sections II-A and II-B. Thus, demonstrating the falsity of claim (b) amounts to showing that no such distinction can be drawn on the basis of occurrences of attributes and entities. This in turn amounts to showing that the sort of division of occurrences of attributes and entities proposed in section II-A is impossible.

How would one show that such a division is impossible? To show that any such division will be an "artificial" convention will not do, since Carnap and others admit this. To discuss problems about borderline cases also will not do, since someone like Carnap could admit these and make conventional conservative decisions about how to handle these cases. And to consider various proposed divisions and attack them does not demonstrate the impossibility of drawing such a distinction. In fact, it seems that the only way to show that such a division is impossible is to demonstrate either that no finite characterization of the division is possible, or else that any possible division which clearly makes observable occurrences directly observable will result in such an impoverished stock of observable occurrences that most of science could not be confirmed.

The chances of successfully demonstrating either of these contentions seems quite remote. Accordingly, it appears virtually impossible to establish claim (b). And since claim (a) leads to reason (ii) (that the observational/theoretical distinction cannot be drawn satisfactorily) only if claim (b) can be established, it follows that reason (ii) for rejecting the Received View has not been established. This, together with the fact that reason (i) (that partial interpretation cannot be formulated adequately) is false, is sufficient warrant to conclude that Achinstein and Putnam are urging the rejection of the Received View for the wrong reasons.

III. THE RECEIVED VIEW VERSUS THE SEMANTIC CONCEPTION

Our consideration of the observational/theoretical distinction makes it clear that if the distinction can be drawn in a manner satisfactory for the purposes of the Received View, things will be exceedingly complex. The fact that science manages to go about its business without involving itself in such complexities suggests that the distinction is not really required or presupposed by science, and so is extraneous to an adequate analysis of scientific theories. The question, then, is whether the observational/theoretical distinction is required for an adequate analysis of the epistemological structure of theories. More specifically, is it possible to give an analysis of the structure of theories which does not employ the observational/theoretical distinction, yet is epistemologically more revealing than the Received View? If such an analysis can be shown possible, then I think we have sufficient reason for rejecting the Received View.

Those who claim that the observational/theoretical distinction is an essential ingredient of an adequate analysis of scientific theories apparently justify their contention with the following implicit line of argument:

> Scientific theories are developed to explain or predict events which can be observed; however, for reasons of simplicity, scope, and economy, such theories typically must employ theoretical entities or constructs in providing these explanations or predictions; these theoretical constructs are not directly observable. Accordingly, in any theoretical explanation or prediction one finds two sorts of sentences: (a) various premises the truth of which is nonproblematic in virtue of their being confirmed by direct observation; (b) various laws the truth of which is problematic since they cannot be confirmed by direct observation. And the observational-theoretical distinction is needed to keep distinct the different statuses of these two kinds of sentences.[18]

This picture is partially correct. Evidently it is the case that in theoretical explanation and prediction the truth of the laws used often is problematic, (especially when predictions are made in order to test the theory), whereas the truth of the evidential premises used in conjunction with the laws is assumed to be nonproblematic. Thus far the dichotomist's argument is satisfactory; but to infer from this that the premises are nonproblematic by virtue of being observational statements and that the laws are problematic by virtue of being nonobservational is unwarranted, for it amounts to assuming an additional premise in the argument—that to be nonproblematic is to be an observational statement.

Not only does this premise beg the question, it also seems false. For the purposes of explanation and prediction all that is required is that the data premises used with the theory be considered nonproblematic relative to the theory or law which provides the prediction or explanation. This is, in applying a theory (or law) to phenomena, what we do is collect data about phenomena; the process of collecting the data often involves recourse to rather sophisticated bodies of theory. If accepted standards of experimental design, control, instrumentation—and possibly involved reliability checks—are carried out, a body of "hard" data is obtained from experimentation and is taken to be relatively nonproblematic; sometimes generally accepted laws or theories are also employed in obtaining these "hard" data.[19] It is to this body of "hard" data that the theory is applied.

If the purpose of the application of a theory is explanation, then the theory explains the event under the description provided by this "hard" data by relating it to other "hard" data which function as descriptions of other features which were the cause of the event so described.[20] If the point of the application of the theory is prediction, then the initial "hard" data are used as premises from which to obtain predictions as to the "hard" data one subsequently would obtain. And these "hard" data may be quite theory-laden, hence nondirectly observable. In addition, what counts as "hard" or nonproblematic data is relative—for should the theory's predictions fail, we may come to treat the data as problematic again.[21] Thus the relevant distinction is between "hard" data and the more problematic theories, and not between the directly observable and the nondirectly observable. Accordingly, the correspondence rules for a theory should not correlate direct-observation statements with theoretical statements, but rather should correlate "hard" data with theoretical statements. Thus it seems that the observational/theoretical distinction is not essential to an adequate analysis of the structure of scientific theories.

Suggestive as this may be, this line of argument does not establish the inadequacy of the Received View. For an advocate of it could accept such an argument and still deny the conclusion, arguing as follows: "It is true that in actual scientific practice theories are pitted against 'hard' data. But what makes them 'hard' is that they ultimately rest on directly observable evidence; and in the Received View reconstruction of theories, that dependence of 'hard' data on the direct evidence of the senses is reflected in the correspondence rules. In fact, even the relativity of 'hard' data can be accommodated in terms of changes in the correspondence rules." There is little doubt that this can be built into the correspondence rules, but the relevant question is whether this can be done without obscuring important epistemological features of scientific theorizing. However, when one reflects that the theory's reliance on the results and procedures of related branches of science, the design of experiments, the interpretation of theories, calibration procedures, and so on, are all being lumped into the correspondence rules, there is reason to suspect that a number of epistemologically important and revealing aspects of scientific theorizing are being obscured.

I maintain that this is indeed so: Because of its reliance on the observational/theoretical distinction, the Received View's account of correspondence rules must combine together a number of widely disparate aspects of the scientific enterprise in such a manner as to obscure a number of epistemologically important and revealing aspects of scientific theorizing. To support this contention it will be necessary to sketch a more adequate alternative account of scientific theories, the Semantic Conception, which reveals what the Received View's treatment of correspondence rules obscures.

The notion of a *physical system* provides us with a convenient starting point for sketching and motivating this alternative account. A science does not deal with phenomena in all of their complexity; rather, it is concerned with certain kinds of phenomena only insofar as their behavior is determined by, or characteristic of, a small number of parameters abstracted from those phenomena.[22] Thus in characterizing falling bodies, classical particle mechanics is concerned with only those aspects of falling-body behavior which depend upon mass, velocity, distance traveled over time, and so on. The color of the object and such are aspects of the phenomena that are ignored; but the process of abstraction from the phenomena goes one step further: We are not concerned with, say, actual velocities, but with velocity under idealized conditions (e.g., in a frictionless environment, with the mass the object would have if it were concentrated at an extensionless point). Thus, for example, classical particle mechanics is concerned with the behavior of isolated systems of extensionless point-masses which interact in a vacuum, where the behavior of these point-masses depends only on their positions and momenta at a given time. A physical system for classical particle mechanics consists of such a system of point-masses undergoing a particular behavior over time. Physical systems, then, are highly abstract and idealized replicas of phenomena, being characterizations of how the phenomena *would have* behaved *had* the idealized conditions been met. Looking at classical particle mechanics again for an illustration, the phenomena within its scope are characterized in terms of the physical systems corresponding to the phenomena.

In arguing that scientific theories are concerned with characterizing the behavior of physical systems, and not phenomena, I may seem to be making the case too easy for myself by using the example of classical particle mechanics—which is what Quine has called a "limit myth" and thus is particularly susceptible of my treatment. However, a brief consideration of some examples will indicate that this is not so and will display the generality of my treatment.

First, consider classical thermodynamics, statistical mechanics, and quantum mechanics. These embody essentially the same "limit myth" and easily can be shown susceptible of my treatment (see Suppe 1967, ch. 3, for details). Second, observe that the gas laws (e.g., Boyle's law and Charles' law) describe the behavior of ideal gases, not real gases; yet they are used in work with actual gases. Here, the ideal gases described by the laws are the physical systems. Subject to appropriate experimental design, and so on, they correspond to actual gases as idealized replicas.

The third example, the valence theory of chemical reactions, is similar. It describes the way theoretically pure chemical substances react together. However, such pure substances are fictional ideals, and the substances in actual chemical reactions are always only approximations of them. The theory describes physical systems, which are chemical reactions theoretically pure substances undergo in this case, and with appropriate experimental and quality controls we can approximate the fiction that our actual substances are pure substances and thereby treat the actual chemical reactions (phenomena) as if they were idealized reactions between pure substances (physical systems).

The fourth example concerns the genetic theory of natural selection which characterizes evolutionary phenomena in terms of changes in the distributions of genotypes in populations as a function of reproductive rates, reproductive barriers, crossover frequencies, and so on. As such, the theory treats populations of individuals (phenomena) as if they were idealized populations of genotypes (physical systems) whose changes in genotypic distributions are functions of only a few selected factors.[23]

A fifth example is the body of stimulus-response behavioral theories which attempts to characterize various kinds of behavior as functions of selected stimulus and response parameters. Such theories describe the behavior of populations of idealized individuals whose behavior is only a function of the specified stimulus and response patterns, reinforcement schedules, and so on (physical systems). On the contrary, the behavior of individuals in actual populations of, say, rats or humans (phenomena) is not simply a function of these selected parameters, and only under the most strictly controlled laboratory conditions can this fiction be approximated. The stimulus and response theories thus describe the behavior of physical systems, not phenomena.

In addition to the above examples, one may include grammatical theories of linguistic competence, kinship system theories, theories in animal physiology, and so on, which also describe the behavior of idealized systems or mechanisms, whose actual systems or mechanisms are, to varying degrees, only idealized approximations. Although brief and sketchy, the examples suffice to illustrate the variety of theories susceptible of my treatment. Further, their variety strongly suggests that scientific theories in general describe the behavior of physical systems, which are idealized replicas of actual phenomena. I will defend this stronger claim in chapter 11.

In general, a scientific theory has the task of describing, predicting, and (possibly) explaining a class of phenomena. It does so by selecting and abstracting certain idealized parameters from the phenomena, then characterizing a class of abstract replicas of the phenomena which are characterized in terms of the selected idealized parameters (see note 22). These abstract replicas are physical systems. The theory thus provides a comprehensive characterization of the behavior of phenomena under the idealized conditions characteristic of the physical systems corresponding to the phenomena; typically, this characterization enables one to predict the behavior of physical systems over time.[24] When coupled with an appropriate experimental methodology, the theory can also predict or explain phenomena which do not meet these idealized conditions by displaying how these phenomena *would have* behaved *had* the idealized conditions been met.

A central task of a theory, then, is to present descriptive, predictive, and possibly explanatory accounts of the behavior of physical systems which correspond to phenomena. The theory is not concerned merely with providing such an account for just the phenomena we do in fact observe, but also with providing one for any phenomena of the sort we *might* encounter in *any* causally possible universe.[25] Further, it must provide a predictive, and possibly explanatory, characterization of all those physical systems which correspond (as abstract replicas) to phenomena of the latter sort. Let us call this class of physical systems the class of *causally possible physical systems*. A central task of any scientific theory is to provide a precise characterization of the set of causally possible physical systems for the theory.

How does the theory provide such a characterization? Once the relevant parameters for the theory have been abstracted and selected from the phenomena, the physical systems for the theory can be specified in terms of these parameters (a physical system being a possible behavior pattern specifiable in terms of these parameters). For example, in classical particle mechanics we might specify a particular state of a physical system in terms of the values of position and momentum parameters at a given time, and then characterize a physical system as a possible sequence of states over time. Of all logically possible physical systems capable of being specified in terms of the chosen parameters, only some will be empirically possible. For example, some of them will not be compatible with accepted existing bodies of theory. Of those which are, only some will be causally possible—in the sense that they correspond (as abstract idealized replicas) to phenomena which could be observed in some causally possible universe.

The theory, then, must specify which logically possible physical systems are causally possible—typically, by providing general laws which are claimed to describe the behavior patterns characteristic of just the causally possible physical systems. These laws are designed to yield predictions of subsequent system states when used together with specifications of initial states and boundary conditions. For example, in classical particle mechanics the equations of motion provide a general description of the class of causally possible physical systems. The characterization of a particular causally possible physical system can be obtained by solving the equations of motion relative to specified boundary conditions and an initial state; the solution can then be manipulated to yield predictions of subsequent system states.[26]

The account of theories just sketched seems to cohere closely with the actual formulations of many theories in the physical sciences. If it is substantially correct, then an observational/theoretical distinction is not required in an adequate analysis of the structure of scientific theories; this is so because theories are not concerned primarily with applying laws directly to phenomena, but rather with using laws to predict and explain the behavior of physical systems abstracted from phenomena in such a manner that their behavior can be correlated with phenomena. These conclusions obviously have important implications for the Received View's notion of a correspondence rule.

We now explore these implications, beginning with a look at how the "hard" data relate to physical systems and their corresponding phenomena. The observation reports or "hard" data to which the theory is applied are partial descriptions of the behavior of some physical system, the physical system being an abstract replica of the phenomena from which the data were collected. Data collection not only involves performing measurements upon the phenomena, which determine the "actual" values of the chosen parameters at different times, but it also involves employing various correction procedures (such as using friction coefficients, and the like) to alter the observed data into data representing the measurement results which *would have been* obtained *had* the defining features of the idealized parameters of the physical system been met by the phenomena.

Thus, in classical particle mechanics our data do not represent, for example, the velocity with which the milk bottle actually fell, but rather the velocity with which it *would have* fallen *had* it fallen in a vacuum, *had* it been a point-mass, and so on. That is, in a typical predictive or explanatory application of a theory, the "hard" data employed are data about the behavior of a physical system at certain times rather than about the actual behavior of the corresponding phenomena. As such, the "hard" data will be expressed in terms of the basic parameters common to the physical system and the theory—which is to say, in terms of what might be called the "theoretical" vocabulary.

Once these "hard" data are obtained, perhaps together with "hard" data about boundary conditions, and so forth, they are used in conjunction with the laws of the theory to deduce various predictions or explanations about the physical system. These deductions typically are "calculational" in nature. For example, in classical particle mechanics they might encompass solving the basic equations of motion for special case solutions, and then "plugging in" values of the parameters to calculate subsequent states of the physical system. Typically the predicted data about these subsequent states of the physical system are then converted into data about the corresponding phenomena by reversing the procedures used originally to convert the data about the phenomena to data about their corresponding physical system.

What we have here, then, is a two-stage move from raw phenomena to statements of the theory—first a move from phenomena to "hard" data about the physical system in question, and then a second move from the physical system to the postulates, and so on, of the theory.[27] The two sorts of moves are qualitatively quite different, the former being essentially empirical or experimental (being, in effect, a "translation" from the phenomena to an idealized description of it in the vocabulary of the theory's formalism), and the latter being essentially mathematical or computational in nature.

This perspective—together with the observation that theories have "hard"-data reports as their primary subject matter rather than direct-observation reports—invites reassessment of the Received View's account of the correspondence rules. For the rules of correspondence lump together the two sorts of moves just discussed so as to eliminate the physical system. It is tempting to reject the Received View's treatment of correspondence rules on the ground that most paradigmatic exact theories in physics and chemistry do work in terms of physical systems in the manner just explained, and then conclude that the Received View is inadequate since it fails to take them into account. While this is a somewhat appealing line, given the explicative character of the Received View analysis, it is not clear how far the criticism cuts. However, if important epistemological features of scientific theorizing are obscured by the failure to countenance physical systems, then it is justifiable to insist that the Received View is defective and epistemologically misleading by failing to include them.

The second-stage movement from data about the physical system to the theory (e.g., the various predictions, etc., about subsequent behavior of the physical system calculated on the basis of these data and the laws or postulates of the theory) is essentially computational in nature. If the theory is quantitative, the theory will be essentially mathematical, involving the solution of equations of motion, various auxiliary definitions and hypotheses, and so on;[28] and at no time are counterfactual inferences involved. On the other hand, the transition from phenomena to a physical system (or vice versa) involves processes of measurement, equipment design, experimental techniques, interpretation and correction of raw data, the employment of theory from other branches of science, inter alia. And the transition from phenomena to physical system is, as I said before, fundamentally counterfactual—being a characterization of what the phenomena *would have been* under idealized circumstances.

From these characteristics it follows that the ways a transition from a physical system to theory can go wrong will be quite different from the ways that the transition from phenomena to a physical system can go wrong. And in the case of a disconfirming experiment, if the source of the difficulty can be isolated as occurring in the transition from phenomena to physical system (i.e., the data proved to be less "hard" than thought), the resolution of the disconfirmation does not require alteration of the theory. In this case, the theory was not at fault; rather, poor experimental procedure was followed (e.g., the instrumentation was miscalibrated, the wrong corrective factors were applied to the raw data, etc.). Only when the disconfirmation cannot be attributed to the transition from the phenomena to a physical system (i.e., the data are as "hard" as we had first supposed), will resolution of the defects require alteration or modification of the theory itself.[29]

It seems amply clear from these observations that there is considerable epistemic difference between the two transitions, and that attention to these differences exposes some rather characteristic features of the relations holding between theory and phenomena. The correspondence rules of the Received View obscure these differences by agglomerating all these various aspects of the relations holding between theory and phenomena into the one correspondence-rule transition. This, in particular, means that experimental errors, and so on, which result in disconfirming instances of a theory will require modification of the correspondence rules and hence of the theory itself, for the correspondence rules are part of the theory and embody a complete specification of all allowable experimental procedures. Another related problem is that the Received View's treatment of correspondence rules gives one little reason to suppose that an exhaustive explicit specification of allowable experimental procedures of the sort required can be given for most theories.[30]

It seems quite obvious, then, that the Received View's characterization of the correspondence rules gives a quite misleading account of the ways in which theories correlate with phenomena, thus obscuring a number of characteristic and important epistemic features of scientific theorizing. Using my characterization of physical systems and the two- (or more) stage transition between phenomena and theory, we obtain an epistemologically more revealing picture of scientific theorizing. Indeed, the need for an observational/theoretical dichotomy disappears, for at no point in that picture is such a dichotomy needed. Replacing it is the distinction between nonproblematic "hard" data about physical systems and boundary conditions, and so on, and the more problematic theoretically obtained assertions about these systems.[31]

And in place of the correspondence rules providing a bridge between theory and phenomena, we now have a two-stage transition: (a) the transition from phenomena to physical systems (which reduces to problems of measurement, experimental design, counterfactuals, and the like) and (b) the connection between the theory and physical systems, which are deductively determined by the (often mathematical) apparatus of the theory without requiring additional correspondence rules or postulates other than boundary conditions and data about the initial state of the physical system. The former transition is not part of the theoretical apparatus of the theory, but rather belongs to the experimental procedures used in applying the theory to phenomena; the latter transition is essentially computational in nature.

This suggested alternative account of the structure of scientific theories enables us to see another flaw in the Received View. If it is correct that the subject matter of a theory is the behavior of physical systems and that the "hard" data include experimental data about the behavior of physical systems, then the central distinction between the nonproblematic "hard" data and the more problematic theoretical assertions about physical systems cannot be drawn on the basis of language. This is because the defining parameters of the physical systems (e.g., position and momentum coordinates in classical particle mechanics) are the basic parameters of the theory, and so the same "theoretical" terms will be used to provide linguistic characterizations of both the theory and the "hard" data. That is, the relevant distinction here is not a linguistic one, but rather an epistemological one. The fact that the key distinction here is not a linguistic one indicates that a number of epistemologically revealing features of the structure of scientific theories are not reflected in their linguistic formulations, and so they cannot be characterized adequately by an analysis of the language of theories—herein lies the ultimate inadequacy of the Received View.

IV. SUMMARY AND CONCLUSION

To summarize, in section I of this chapter my primary aim has been to make the nature of partial interpretation as employed in the Received View reasonably precise for two reasons. First, I feel that this analysis can lead to a number of interesting criticisms of the Received View. In particular, it would appear that the sort of semantic interpretation provided by partial interpretation is insufficient to make Carnap's treatment of inductive logic and degree of confirmation applicable to *TC*. Second, although ultimately the Received View is unsatisfactory and must be rejected, a number of facts about the meaning of scientific terms revealed in the discussion of partial interpretation will prove useful in developing an alternative analysis to the Received View. In particular I suggest that an adequate analysis of theories in the exact or formalized sciences will have to treat the relationships between ordinary scientific language and mathematical formalism along the lines presented here, albeit in a manner that does not make recourse to an observable/theoretical distinction.

I also have tried to show that the sort of criticisms against the Received View raised by Achinstein and Putnam do not succeed in showing its inadequacy. Nonetheless, the Received View is unsatisfactory, since its reliance on the observational/theoretical distinction obscures much that is epistemologically important and revealing about how theories relate to, or connect with, phenomena. To demonstrate this, I have sketched an alternative analysis of the structure of theories and used it to show the following: how the Received View obscures the role of physical systems, the way in which extratheoretical postulates provide nonexhaustive characterizations of the admissible transitions between phenomena and physical systems, and wherein lies the role of counterfactuals in connecting theories with phenomena. These epistemic revelations do not exhaust the potential of this alternative account. To indicate just some of its potential, further development of the analysis (e.g., along the lines of Suppes 1962) will reveal much more about the experimental relations holding between phenomena and physical systems. In addition, the isolation of the counterfactual component of scientific theorizing in the transition between phenomena and physical systems provides a perspective which conceivably could advance us toward a breakthrough on the problem of laws and counterfactuals. For the exact sciences, there is ample evidence that this sort of account can be expanded and developed so as to give a particularly revealing account of exact theories (e.g., the sorts of revelations about phase spaces, the connection between deterministic and indeterministic theories, and so on, found in van Fraassen 1970; Suppe 1967, ch. 2; and parts II and III of this volume).

What's wrong with the Received View? It obscures much of epistemic importance other analyses can reveal. For this reason it should be rejected in favor of such an alternative analysis, which I have tried to sketch. In part II of this volume, the Semantic Conception analysis will be developed in much further detail.

NOTES

1. Cf. Putnam 1962 and Achinstein 1968, 85–91, 157–58, 197–202. Achinstein's 1968 book incorporates his earlier writings on the subject with minor changes. Putnam 1962 also urges that the observational/theoretical distinction is untenable. He argues that it is misleading both to label the class of nonobservational terms 'theoretical terms' and to characterize sentences formulated solely in terms of the observational vocabulary as observational sentences, and those formulated solely in terms of the theoretical vocabulary as theoretical sentences. However, while this is true, it hardly necessitates rejection of the Received View.

Although Achinstein (1968, 199–201) suggests that it would be epistemologically more revealing if we avoided reliance on an observational/theoretical distinction in our analysis of theories, this is only a corollary to his main arguments against the observational/theoretical distinction. The strength of Putnam's and Achinstein's contention that the Received View should be rejected lies in the establishment of (i) and (ii), and I shall confine my attention to those arguments.

2. Versions of the Received View have been advanced by a number of authors, including Braithwaite (1953, 22 ff.), Campbell (1920, ch. 6), Carnap, (e.g., 1956, 43), Duhem (1954, 19), Hempel (1952, 1958), Hesse (1962, 1966), Kaplan (1964, 298–99), Margenau (1950), Nagel (1961, 90), Northrop (1947), Ramsay (1931, 212–36), and Reichenbach (1962, ch. 8). Although a number of differences exist (some significant) between these various versions of the Received View, there is a substantial core of agreement among them.

The primary disagreements among them are about the form of the correspondence rules: Various authors refer to the rules as coordinating definitions, dictionaries, interpretative systems, operational definitions, epistemic correlations, and rules of interpretation. Campbell, Nagel, Hesse, and Kaplan maintain that (in addition to satisfying conditions [1] through [6] of my reformulation of Carnap's and Hempel's version) the theory also must possess realizable or concrete models. Kaplan deviates from the others in that he claims that the analysis only works for one type of theory. Hempel (1974) no longer adheres to the Received View and later adopted a similar position in which the observational/theoretical distinction was replaced by a different bifurcation of terms.

My discussion of the Received View and of partial interpretation will require recourse to a fair amount of symbolic logic. I assume that the reader is familiar with first-order languages, their alphabets, predicate calculi based on first-order languages, theories formulated in a first-order predicate calculus, semantical systems for languages, rules of designation, rules of truth, interpretation of first-order theories via semantical systems, true interpretations or models of first-order theories, and validity. Those readers requiring further details on these notions may skip the more technical portions of section I-B, relying on the informational summaries of their technical results. For further comprehension, one can consult Carnap 1942 and also the relevant portions of Mates 1965.

3. This formulation of the Received View is extracted from Carnap 1956, 1959, 1963, and 1966 and Hempel 1958 and 1963. Conditions (2), (3), and (4) are more explicit in certain respects than either Carnap or Hempel specifically requires, but this is necessary if Carnap's restrictions on the sublanguages L_O and L'_O (Carnap 1956, 41–42) are to be satisfied. To meet these restrictions different additional conditions could have been imposed, but I have selected the most conservative ones. My particular choice in no way affects the analysis of partial interpretation given here.

Carnap and Hempel disagree as to the requirements to be imposed on the rules of correspondence. Hempel would replace clause (5)(d) with the following: "C contains every element V_O and V_T essentially—i.e., C is not logically equivalent to some set of sentences in which at least one term of V_O or V_T does not occur at all" (1963, 692). His version thus is more restrictive than Carnap's. Carnap also would require that the theory be cognitively significant, whereas Hempel doubts that a satisfactory criterion of cognitive significance can be given. These minor differences between Carnap's and Hempel's formulations need not concern us here.

4. Cf., e.g., Achinstein 1968, 85 ff., and 1963, and Putnam 1962.

5. See Hempel's (1952) discussion of explicit definition for such a characterization.

6. See Suppes 1957, ch. VIII, for a very lucid and not excessively technical discussion of this theory.

7. Selected for the sake of simplicity and intelligibility, this formulation is only a very special case of the Beth-Padoa result and takes a number of heuristic liberties that are of no consequence here. For a rigorous treatment of the result in general form, see Shoenfield 1968, 81.

8. For a discussion and characterization of introductive chains, see Carnap 1936–37, sec. 6. The characterization there is for terms partially interpreted

via reduction sentences, but extends straightforwardly to C-rules as formulated here.

9. See pages 85–86. Achinstein also considers two other possible interpretations which need not concern us here (despite the fact that the second one, on p. 86, at first seems quite similar to mine and thus possibly relevant to my proposal). This is because his criticism turns on various features of the peculiar semantical rules he introduces, and since my rules do not have these features, his discussion in no way affects my proposal.

10. See p. 695. (Minor notational changes have been made in this and other quotations to bring the notation into agreement with that adopted in this chapter.) Carnap makes essentially the same observations (1939, 62).

11. Such a position was, of course, not theirs initially, but seems to have gradually developed since Carnap (1936–37); most adherents of the Received View were slow in coming to such a realization, however, and some (e.g., Bergmann) apparently never did. But the position seems to be either implicit or semi-explicit in the more recent writings of Hempel and Carnap.

12. See, e.g., Carnap 1966, 256, for a statement in this vein.

13. This seems to be the point of Carnap's (1956) claim that "all the interpretation *(in the strict sense of this term, i.e., observational interpretation)* that can be given for L_T is given in the C-rules" (46; emphasis added).

14. These are, e.g., the sort of considerations Kuhn's (1962) paradigms are supposed to handle.

15. This last set of claims may be overstated; it may be that for a certain *TC* no such interpretation is conceivable given the specified rules of designation. Whether this is possible requires further investigation, and I leave it as an open question. I would conjecture, however, that if such a situation did occur, the *TC* in question would be so bizarre that a good case could be made for treating the theory as being meaningless, given the observational phenomena to which it was supposed to be applied. Unfortunately, I don't know how to demonstrate this.

16. In case the L used in the Received View canonical formulation is a symbolic language of the sort used by Carnap (1956), then V_O would contain predicates which correspond to, for example, English observational terms, and V_T would contain predicates which correspond to, e.g., English theoretical terms. Regardless of whether the L used in the Received View is a natural or an artificial language, then, the V_O–V_T distinction apparently would be drawn on the basis of standard usages in some natural language.

17. I can imagine that some proponents of the Received View would protest here that I do not directly observe, for example, that something is a gas, but rather that I observe certain manifestations of the presence of the gas; accordingly, the property of being a gas never can be directly observed. But this argument fails; for if it is legitimate, then it seems equally fair to argue that I do not directly observe that something is hard, but instead merely observe certain manifestations of the thing being hard. In this sense one can never directly observe the property of being hard, and so it is not a directly observable property—contrary to the fact that Carnap advances it as a

paradigmatic example of a directly observable property. It should also be noted that the Received View does not limit direct observation to visual perception; direct observation can be made by any sense, as Carnap's own examples in the quotations above make clear. Finally, although the argument here proceeds in terms of attributes, it is clear that analogous arguments could be given, and similar conclusions drawn, for entities. For simplicity of exposition I present the arguments just for attributes.

18. At one time proponents of the Received View also might have justified introducing the dichotomy by appealing to considerations of cognitive significance and using a thesis about language acquisition. The apparent failure of the positivistic account of cognitive significance and the falsity of the thesis about language acquisition make it both unlikely and undesirable that Received View advocates would argue it on these grounds.

19. This discussion has benefited from my conversations with Professor Don E. Dulany. See Suppe 1974, 424–33, for Putnam's treatment of the use of such auxiliary hypotheses; see also van Fraassen 1970 for a related discussion.

20. This rough characterization of the role of data in explanation turns on an observation—insufficiently considered in the literature on explanation—that explanations do not explain events simpliciter, but rather explain *events under a particular description* (see p. 189). While it is beyond the scope of this chapter to argue it, this observation apparently can be exploited to show that the alleged symmetry between explanation and prediction collapses.

21. See the introduction to Quine 1959 for a discussion of this point.

22. I use the term 'parameter' rather than 'variable' to mark the fact that the state "variables" need not be measurable on my quasi-realistic version of the Semantic Conception. This fact will be crucial in my treatment of measurement in ch. 4. I am not using 'parameter' in the statistical sense, where it means a variable set to a fixed or constant value.

23. In sec. XI of ch. 7, the applicability of the analysis of theories presented here to the genetic theory of natural selection is worked out in some detail. See also sec. III-C of ch. 6. In ch. 5, sec. II, microeconomic examples are considered.

24. For brevity, I confine my attention here only to theories which describe the behavior of physical systems in terms of changes in state over time. In addition to working for such theories with laws of succession, the analysis also will work for theories with laws of coexistence, laws of interaction, functional laws, and laws of quasi-succession. Also, it makes no difference whether the laws are deterministic or statistical. See chapter 5.

25. The problems of characterizing causally possible universes are many, but they can be viewed roughly as the class of universes in which all the laws assumed nonproblematic relative to the theory in question hold. For a detailed characterization of the notion of a causally possible universe, see Burks 1977, ch. 10. My purpose in employing the notion is to use it to introduce the concept of a causally possible physical system later on. Since the rough characterization given meets the limited purposes set for it here, the difficult

problem of providing an adequate characterization of causally possible universes can be avoided for the time being. I offer a detailed analysis of the notion in a book in progress, *Facts, Theories, and Scientific Observation* (see also p. 300 below).

26. On this account, a theory may be construed as defining a class of theoretically possible physical systems; the theory will be empirically true just in case this class is identical with the class of causally possible physical systems. The account of empirical truth just specified is essentially a generalization of that introduced in section I of this chapter. In both cases the idea is that there exists a class of systems determined theoretically and a class determined empirically, the theory being empirically true just in case the classes are coextensive. Thus, intuitively, the class of causally possible physical systems is the class of physical systems which are empirically possible. Further consideration of this key notion of a causally possible physical system can be found in my dissertation (1967, chs. 1, 2) and in ch. 3 below.

27. Actually this is still an oversimplification, the former move involving many more steps. (See ch. 4 and also Suppes' discussion in his 1962 and 1967 works.)

28. For an illuminating discussion of what is involved in this sort of move, see Putnam's discussion in Suppe, 1974, 424–33.

29. For a detailed discussion of this point, see Suppe 1967, ch. 3.

30. For a more detailed discussion of this last point, see Kuhn 1974, where he discusses the role of exemplars in the application of theories to phenomena; see also my commentary (1974a) on Kuhn 1974.

31. Hempel now rejects the Received View, and in his 1969, 1970, and 1974 works he advances an analysis based on a distinction similar to this. He distinguishes between a *theoretical vocabulary* and *an antecedently available vocabulary*, where the latter may include theoretical terms from generally accepted theories. His proposal differs from mine in that he thinks the relevant distinction can be drawn on linguistic grounds, whereas I explicitly deny that it can. His analysis differs in other respects as well—especially on the nature of the transition between the "hard" data and the theory (his so-called *bridge principles*).

The "hard" data notion proves to be an overly simple heuristic notion. It will be supplanted by the analysis of experimental methodology in ch. 3, sec. VIII, and ch. 4 below.

PART II

The Semantic Conception of Theories

Section III of chapter 2 presented a heuristic account of the Semantic Conception of Theories. The next three chapters provide a detailed development of a quasi-realistic version of the Semantic Conception. Chapter 3 sets forth the basic structural account and examines the various relations holding between theories and their linguistic formulations. Chapter 4 focuses on the physical interpretation of theories, their empirical truth conditions, and the ways observation, measurement, and experimental design relate theories to actual phenomena. Chapter 5 explores the structure of scientific laws on the Semantic Conception.

In addition to helping to flesh out the development of the Semantic Conception, these chapters also use the quasi-realism of the Semantic Conception to contribute to the resolution of other philosophical issues. Chapter 3 addresses positivistic instrumentalism-versus-realism debates and argues against the operational imperative (and it provides background for the discussion of scientific realism in chapter 11). Chapter 4 develops a realistic approach to measurement and argues that so construed, measurement is a species of observation. Chapter 5 argues against Scriven's claim that theories are inaccurate and against the associated idea that approximate truth notions must be invoked in interpreting laws or theories.

3

Theories, Their Formulations, and the Operational Imperative

The *operational imperative* demands that all theoretical concepts be defined or introduced into theories operationally.[1] Its tenability can be judged only relative to an analysis of the nature or structure of scientific theories. For the most part, philosophical assessment of the operational imperative has been relative to the positivistic Received View. When viewed from this perspective, the operational imperative is interpreted as the requirement that theoretical terms be defined as explicit definitions (or possibly as reduction sentences) in terms of an observation vocabulary which specifies various operations and possible outcomes resulting from such operations being performed: "An operational definition of a term is conceived as a rule to the effect that the term is to apply to a particular case if the performance of specified operations in that case yields a certain characteristic result. . . . The operations are to be intersubjective in the sense that different observers must be able to perform 'the same operations' with reasonable agreement in their results."[2]

So interpreted, the operational imperative becomes the requirement that the correspondence rules in the Received View be restricted to a species of explicit definition (or, possibly, reduction sentences). Since the positivistic account of theories is untenable when correspondence rules are restricted to explicit definitions or reduction sentences,[3] it follows that the operational imperative is untenable.[4] On the basis of such arguments most philosophers of science rejected the operational imperative, although a number of working scientists (especially in the social and biological sciences) continued to swear by it.

It is the contention of this chapter that these arguments do not settle the philosophical question of the operational imperative's tenability. Strictly speaking, such arguments only show that the operational imperative is incompatible with the Received View, and the untenability of the operational imperative follows only if the Received View of theories is itself an adequate analysis. However, the arguments of the previous chapter and a number of other papers make it clear that the Received View is untenable and must be rejected,[5] so the incompatibility of the operational imperative with the positivistic analysis of theories does not show that the operational imperative is untenable. Accordingly, the operational imperative deserves a new philosophical hearing. It is one task of this chapter to provide that hearing.

I. THEORIES

In order to reassess the operational imperative it will be necessary to presuppose some analysis of the nature of scientific theories. To my mind the most plausible and developed analysis currently available is the Semantic Conception, introduced in the previous chapter. This chapter will develop in further detail some of its main features.

As actually employed by working scientists, theories admit of a number of alternative linguistic formulations—for example, classical particle mechanics sometimes is given a Lagrangian formulation and other times a Hamiltonian formulation—but it is the same theory regardless which formulation is employed. As such, scientific theories cannot be identified with their linguistic formulations; rather, they are extralinguistic entities which are referred to and described by their various linguistic formulations. This suggests that theories be construed as propounded abstract *structures* serving as models for sets of interpreted sentences that constitute the linguistic formulations. These structures are *metamathematical models* of their linguistic formulations, where the same structure may be the model for a number of different, and possibly nonequivalent, sets of sentences or linguistic formulations of the theory.[6]

Theories are formulated to characterize a class of phenomena known as the *intended scope of the theory*, perhaps, say, the class of all mechanical phenomena of interacting bodies. The theory does not attempt to characterize the phenomena in all their complexity, but only attempts to do so in terms of a few parameters abstracted from the phenomena. For example, classical particle mechanics attempts to characterize mechanical phenomena *as if* they depended only on the abstracted position and momentum parameters.[7] In point of fact, however, other unselected parameters usually do influence the phenomena; so the theory does not characterize the actual phenomena, but rather characterizes the contribution of the selected parameters to the actual phenomena, describing what the phenomena *would have been had* the abstracted parameters been the only parameters influencing them. For example, classical particle mechanics does not describe actual inclined plane phenomena, but instead describes what inclined plane phenomena *would be* in frictionless environments. In effect, then, what the theory does is directly describe the behavior of abstract systems, known as *physical systems*, whose behaviors depend only on the selected parameters. However, these physical systems are abstract replicas of actual phenomena, being what the phenomena *would have been* if no other parameters exerted an influence. Thus by describing the physical systems, the theory indirectly gives a counterfactual characterization of the actual phenomena.

In abstracting from the phenomena, the physical systems also may idealize the phenomena in various ways. For example, in classical particle mechanics physical systems are isolated systems of dimensionless point-masses interacting in a vacuum. Such physical systems are abstract replicas of phenomena on which certain idealized conditions (e.g., being isolated systems of dimensionless point-masses) are imposed, which actual phenomena cannot ever meet. In such cases, the physical systems still are characterizations of what the phenomena would be were certain conditions met—only some of these conditions cannot possibly be met by any actual phenomenon.

The behavior of physical systems can be described wholly in terms of the selected parameters abstracted from the phenomena; these are the *defining parameters of the physical system*. The values of these parameters are *physical quantities*, which may be determinate or statistical.[8] A set of simultaneous values for the parameters of a physical system is a possible *state* of the physical system. At any given time, a physical system is in exactly one of its possible states, though the state it is in may change over time. The *behavior* of a physical system is its change in states over time, and this can be viewed as its history. Just as each phenomenon has a unique history, each physical system has a unique sequence of states it assumes over time.[9] Each physical system, then, can be characterized fully by a specification of the possible states it can assume (as a function of the defining parameters) and the sequence of states it assumes over time. The defining parameters of a physical system will be *basic parameters* of its associated theory, and so physical systems will be described in "theoretical language."

Corresponding to any *causally possible phenomenon* P within the theory's intended scope will be a physical system S such that S is what P *would have been were* the idealized conditions (if any) imposed by the theory met and the phenomenon P were influenced *only* by the selected parameters. Let the class of *causally possible physical systems for a theory* be the class of physical systems which correspond in the manner just indicated to causally possible phenomena within the theory's intended scope. Inter alia, it is the job of a scientific theory to exactly circumscribe the class of causally possible physical systems for the theory.[10] The theory does so by determining a class of physical systems known as *theory-induced physical systems*. In propounding the theory we are claiming that the class of theory-induced physical systems is identical with the class of causally possible physical systems for the theory. If the theory is *empirically true*, then these two classes are identical;[11] and if they are not identical, the theory is *empirically false*. In testing a theory it is the job of the experimental methodology to determine which physical systems correspond to which phenomena;[12] and if one subscribes to an inductive logic approach to confirmation, the confirmation of a theory essentially involves determining through (nonparametric) goodness-of-fit statistics or their informal analogues whether the class of theory-induced physical systems is identical with the class of causally possible physical systems for the theory.[13]

Our discussion thus far has indicated two important features of theories: (1) they are propounded extralinguistic structures which qualify as metamathematical models of their linguistic formulations, and (2) they determine a class of theory-induced physical systems. These features can most easily be accommodated if we analyze theories as *relational systems* (Tarski and Vaught 1957) consisting of a domain containing all (logically) possible states of all (logically) possible physical systems for the theory together with various attributes defined over that domain. These attributes, in effect, are the *laws* of the theory.[14]

If the theory has *laws of succession*, then the attributes will be relations of succession indicating which sequences of states various physical systems will assume over time; these relations may be such that the sequences are deterministic or statistically determined, continuous or discrete. Deterministic laws of succession are exemplified by the laws of classical particle mechanics, and statistical ones by the transition matrix for a finite Markov process. If the theory has *laws of coexistence*, then the attributes will be equivalence relations indicating which states are equivalent (if it is a deterministic law) or which states are equiprobable (if it is a statistical law). Deterministic laws of coexistence are exemplified by the ideal gas law, and statistical ones by the Boltzmann hypothesis that each microstate of a gas has equal probability. Finally, if a theory has *laws of interaction* (deterministic or statistical), the attributes will determine which states result from the interaction of several systems and will be composites of the kinds of attributes mentioned previously.[15]

Regardless which forms of laws a theory has, the laws do two things. First, they indicate which states are *physically possible* (these being the states which enter into the satisfaction of the theory's attributes). Second, they indicate which sequences of states a physical system can assume; as such, the laws determine the class of theory-induced physical systems. Deterministic laws of coexistence sanction all sequences whose constituent states belong to the same equivalence class; statistical laws of coexistence sanction those whose constituent states have an assigned probability measure; and laws of interaction sanction sequences analogously to the above cases or their admixture. Finally, sequences sanctioned by statistical laws have a probability measure assigned to them by the sanctioning laws; these measures determine the probability that a physical system in state s at time t will be in state s' at time t' ($> t$). The class of sequences so sanctioned constitutes the class of theory-induced physical systems. Usually this class is a proper subclass of the class of all logically possible sequences.[16]

Some states may enter into no sequence, and so are not physically possible. And certain logically possible sequences of states do not satisfy the relations determining sequences of states, and so do not qualify as behaviors of physical systems in the class of theory-induced physical systems. On the other hand, every sequence of states determined by the theory's relations will be the behavior of a physical system in the class of theory-induced physical systems. Thus the relations of the theory determine all and only those sequences which are the behaviors of physical systems in the class of theory-induced physical systems.

Although theories directly determine only the class of theory-induced physical systems, they can be used to predict phenomena in the following manner: Suppose the theory is one whose laws are deterministic laws of succession and that we wish to predict the subsequent behavior of some phenomenon at t'. By means of one's experimental methodology it is determined what physical system state corresponds to the phenomenon at some prior time t. Then, using some formulation of the theory, one determines which theory-induced physical system characterizes the behavior of a physical system in state s at time t; this may be done, for example, by solving the equations in the theory formulation to obtain special case equations. Determining the physical system in question indicates a sequence of states the physical system subsequently will assume, and from that sequence one determines (e.g., by solving the special case solution relative to initial, boundary, and time conditions) what state s' the physical system will be in at t'. If the theory is empirically true, then s' will correspond to the phenomenon in question at t'—that is, s' indicates what the phenomenon *would be* at t', if its parameters *were* the only ones affecting the phenomenon and the phenomenon *were* to meet the idealized conditions imposed by the theory. Then by the experimental methodology, one determines the actual phenomenon p' which should correspond to s', yielding the theory's predictions about the phenomenon at t'. If T has statistical laws of succession or laws of coexistence, the procedure is analogous except that we are only able to determine from the theory that one member of some restricted class of physical systems corresponds to P. Hence we are limited to predicting that P will be in a state at time t' which corresponds to one of a number of physical system states s'; if the laws in question are statistical, our predictions can assign a probability to each of these states s'.[17]

II. THE OPERATIONAL IMPERATIVE REFORMULATED

Viewed from the perspective of this analysis, it becomes clear that the operational imperative is about theory formulations and in effect stipulates that certain relationships do or should hold between theories and their linguistic formulations. This, together with the fact that theories admit of alternative and nonequivalent linguistic formulations, reveals that the operational imperative could be making either of the following demands:

(1) *Weak Operational Imperative.* Every scientific theory must admit of a full linguistic formulation in which theoretical terms are operationally defined.
(2) *Strong Operational Imperative.* The only theory formulations which may be employed are those in which theoretical terms are operationally defined.

The strong imperative clearly presupposes the weak imperative. The weak imperative can be construed either as a descriptive claim supposedly true of all possible theories or else as a prescription for how scientific research should be conducted. As a descriptive claim, the weak imperative is not obviously true of all theories (e.g., there is good reason to doubt whether it is true of quantum theory's ψ function), and it is difficult to see how one would show it true. Most likely the weak imperative is intended as a prescriptive claim. The strong imperative clearly is prescriptive.

Why should anyone insist on either the weak or the strong imperative as a prescription for doing science? Historically, the operational imperative was introduced to explain how theories legitimately could employ parameters which could not be directly observed or measured, and how theories describing phenomena in terms of such parameters could be tested and confirmed observationally. Following the operational imperative not only explained the legitimacy of such procedures, but it also afforded a method for testing such theories and also guaranteed that they did not invoke ad hoc fictitious theoretical entities in explaining phenomena.

Bridgman (1927) clearly had these ends in mind when he introduced the operational imperative. Underlying the operational imperative was the idea that what could be directly observed was epistemologically secure, whereas nonobservable parameters and assertions about nonobservable entities lacked this surety. However, if theoretical parameters could be specified (wholly or partially) in terms of observable conditions and operations whose performance could be observationally checked for correct execution, then the epistemic surety of observation could be passed on to theories. The point, then, of following the operational imperative was to put one's scientific theorizing on an epistemologically sound footing, and to explain how theories employing nonobservable parameters could be tested by observing phenomena.

If the operational imperative does not adequately capture the relations holding between observable phenomena and theoretical parameters which make the empirical testing of theories possible, and if the claimed epistemic advantages of following the operational imperative do not accrue, then there is little point to insisting on it. As such, we can evaluate the weak and strong operational imperatives by considering whether there is any epistemic advantage to following them and whether they adequately characterize the ways in which theories are empirically testable. In order to do so it will be necessary to consider in some detail the semantic relations which hold between theories and their formulations, between physical systems and theory formulations, and between phenomena and theory formulations. It also will be necessary to consider the epistemological or structural relations holding between phenomena, their corresponding physical systems, and theories. These relations will be investigated in the next five sections and then in section VIII, we will use our findings to assess the operational imperative.

III. SEMANTIC RELATIONS

A *formulation of a theory* is a collection of propositions which are true of the theory.[18] Typically, a formulation of a theory consists of a few specified propositions together with all of their deductive consequences under some logic. A *full formulation* of a theory is one which describes all the characteristic features of the theory, whereas a *partial formulation* of a theory describes some, but not all, of the theory's characteristic features. The propositions constituting a theory formulation are of some language, known as the *theory formulation language*, and typically constitute a proper subset of the propositions in that language. Propositions in a theory formulation language may be used not only to describe the theory, but also with reference to physical systems and phenomena within the theory's intended scope. We need to consider the semantic relations holding between theory-formulation-language propositions and theories, physical systems, and phenomena.

At this point it will be useful to introduce a distinction between *strict usage* and *amplified usage* of propositions, which originated with Evert Beth.[19] Often descriptive propositions are such that they can be asserted about any of a number of different systems. Under strict usage, one intends to describe a particular one of these systems and uses the proposition solely with reference to that intended system. Under amplified usage, the proposition is used indifferently to describe any or all of the different possible systems. Any proposition admitting of amplified usage also admits of strict usage, but the converse is not true.

The propositions in theory formulation languages must admit of amplified usage. For the same propositions of a theory formulation language may be used to describe the theory, physical systems, or phenomena. For example, 'the force of the entity is equal to the product of its mass and acceleration', a proposition in a formulation of classical particle mechanics, can be used to describe a characteristic feature of the theory, a particular physical system in the class of theory-induced physical systems, and a particular phenomenon within the theory's intended scope. The amplified usage of such propositions must be such that the proposition can be used *simultaneously* with reference to the theory, one or more physical systems, and phenomena within the theory's intended scope.

For example, suppose we are predicting the velocity of a block on an inclined plane at t' whose friction is negligible, and we experimentally determine that at t ($< t'$) the velocity is 32 ft./s. This yields the proposition 'the body has velocity 32 ft./s at t'', which describes the phenomenon. But in carrying out the procedure for prediction given at the end of section I, this proposition also must be taken as describing the initial state of the physical system corresponding to the phenomenon and as describing a state in the theory structure itself; otherwise the process of prediction given in section I would involve an equivocation fallacy. Hence amplified usage is necessary for predicting phenomena with theories.[20] Also, propositions in the theory formulation language may be used strictly with reference to the theory, physical systems, or phenomena. Key features of propositions in a theory formulation language can be discerned by investigating the strict usages of such propositions with reference to theories, physical systems, and phenomena—for the amplified usage of a proposition must be consistent with its various strict usages.

We begin by looking at the strict usage of propositions in a theory formulation language with reference to theories. Regardless whether a theory formulation is full or partial, it has the following basic features. First, there is a set of elementary propositions in the theory formulation language to the effect that a certain physical parameter p has a physical quantity q at time t. An elementary proposition ϕ will be true of a state s in the theory's domain if s has q as the value of parameter p at time t. For each elementary proposition ϕ there will be a maximal subset $h(\phi)$ of the theory's domain such that ϕ is true of all states in that subset; the function h from elementary propositions to subsets of the theory is known as the *satisfaction function* for the set of elementary propositions. Second, elementary propositions can be compounded together in accordance with some logic or other known as the *logic of the theory.* The logic must be such that every compound proposition is *empirically significant;* that is, it is true of at least one state which, according to the theory, is physically possible.[21] Thus the logic of the theory is determined by the theory, and different theories may impose different logics. For example, classical particle mechanics imposes a Boolean algebra mod-2 and quantum theory imposes a nondistributive lattice. The set of elementary propositions, the theory, the satisfaction function h, and the logic of the theory determine a *language of physical description,* which is a sublanguage of the theory formulation language. This language is capable of describing physically possible individual states, or collections of states, in the theory and thus is able to describe any physically possible state of affairs in a physical system. The logic of the theory enables one to deduce logical consequences of propositions in the language of physical description. In most circumstances, however, the logic of the theory is too impoverished to deduce changes in state over time (because these usually are nonlogical consequences of the propositions in the language of physical description). In order to do so, the language of physical description must be incorporated into a more comprehensive language, using an augmented logic able to express the laws of the theory and deduce various predictions, and so on, from these laws. For example, in classical particle mechanics the additional logic might include the vocabulary and mechanisms of differential equations, and in quantum theory the additional logic might be the algebraic theory of Hilbert spaces or matrix algebra. This expanded logic and language is the *theory formulation language.* The truth conditions for this expanded language (used strictly with reference to the theory) are specified in terms of the attributes of the theory and the truth conditions for the language of physical description. Third, a *formulation* (full or partial) *of a theory* is a set of propositions deductively closed under the logic of the theory formulation language such that every proposition in the set is true of the theory. For example, a formulation of the theory might be a set of *law statements* together with all their deductive consequences. The theory formulation language may be a natural or an artificial language, though typically it is a natural language such as scientific English.

We next turn to a consideration of the semantic relations holding between propositions in the theory formulation language and theory-induced physical systems. A physical system, it will be recalled, is a relational system consisting of a domain of states and a sequence defined over that domain; the sequence is the behavior of the physical system. If the physical system is in the class of theory-induced physical systems, then the domain of the physical system will be a subset of the domain of the theory, and the sequence will be one of the sequences determined by the theory's attributes. Thus a physical system in the class of theory-induced physical systems may be construed as the restriction of the theory to a single sequence. For example, the theory of ideal gases defines various equivalence classes of states in terms of the relation $PV = nRT$ and determines the set of all sequences of states such that the states in the sequence are equivalent. Each of these sequences is the behavior of some physical system in the class of theory-induced physical systems, and so the restrictions of the theory to individual sequences are physical systems.

Since physical systems are restrictions of the theory, propositions in the theory formulation language can be used to refer to and describe physical systems. The truth conditions for these propositions will be the restriction to the physical systems of the truth conditions for the propositions when they are used with reference to the theory. Only some of the propositions which are true of the theory will be true of a particular physical system in the class of theory-induced physical systems, but every proposition true of a physical system in that class will be true of the theory. Some propositions in the theory formulation language cannot meaningfully be used with reference to a particular physical system in the class of theory-induced physical systems (e.g., a true proposition about a four-body system in a formulation of classical particle mechanics cannot be meaningfully asserted about a two-body physical system). However, every proposition in the language of physical description will be true of some physical system in the class of theory-induced physical systems.

If the class of theory-induced physical systems is identical with the class of causally possible physical systems for the theory, then the semantic relations holding between propositions in the theory formulation language and the latter class of physical systems will be exactly the same as for theory-induced physical systems; and every proposition in the theory formulation language which is true of a causally possible physical system will be true of the theory. Suppose, however, that the theory is empirically false—hence that the class of causally possible physical systems is not identical with the class of theory-induced physical systems. In this case, the semantic relations holding between propositions in the theory formulation language and causally possible physical systems will be of exactly the same sort as those holding between such propositions and theory-induced physical systems. But in this case there is no guarantee that propositions true of a causally possible system will be true of the theory, or that any propositions true of the theory will be true of some causally possible physical system.

At this point, a further complication needs to be considered: The logic of the theory restricts the ways in which elementary propositions may be compounded together in the theory formulation language. If the theory is empirically false, then the logic of the theory may prevent certain compoundings of elementary propositions which are true of causally possible physical systems. As such, there may be propositions excluded from the theory formulation language which can be truly asserted of casually possible physical systems. Moreover, one way of falsifying a theory is to show that there are causally possible physical systems which do assume states the theory counts as physically impossible. If the theory is falsified in this way, propositions in the language of a physical description needed to describe the falsifying causally possible physical systems will not be propositions in the theory formulation language, and so the counterinstance to the theory is unstatable in that language.

This indicates that in order to describe and characterize all causally possible physical systems we must use an *expanded theory formulation language* which allows all truth-functional combinations of elementary propositions to be propositions. This language also must have a descriptive mechanism adequate to describe any logically possible behavior of any logically possible physical system (a feature not necessarily possessed by the theory formulation language). As a result, the deductive logical apparatus of this expanded theory formulation language may also have to be expanded. The theory formulation language is a sublanguage of the expanded theory formulation language, and the semantical relations holding between propositions in this expanded language and causally possible physical systems, or the theory, will be generalizations of those holding for propositions in the theory formulation language.

Finally, we turn to a consideration of the strict use of propositions in the expanded theory formulation language with reference to phenomena in the theory's intended scope; for such propositions employing theoretical vocabulary can be, and often are, used with reference to actual phenomena. When so used the propositions are making statements of putative fact and will be true if and only if the propositions are factually true. Thus to analyze the semantic relations holding between propositions in the expanded theory formulation language and phenomena, we will need to employ some account of factual truth.

For brevity I will assume an analysis of factual truth I have developed and used elsewhere.[22] In this analysis, empirical propositions are used to assert that certain states of affairs hold in the world, and if the asserted states obtain, the proposition is true. The world is assumed to consist of particulars having intrinsic properties together with various intrinsic relations particulars may enter into. Particulars may be simple or complex, with a complex particular being the continuous instantiation of an intrinsic relation by particulars.[23] For the purposes of the natural sciences, these particulars are assumed to be nonintentional, and the properties of particulars and the intrinsic attributes they enter into in principle can be specified in an extensional language containing the class abstraction operator, infinite disjunctions, and temporal names for particulars.[24]

When a proposition '*S* is *P*' is used to make a putative factual assertion, its sense determines ostensible referents for *S* and *P*; the ostensible referent for *S* will be a possible particular, and the ostensible referent for *P* will be a possible intrinsic property (an extensional property a particular could possess). Then

'*S* is *P*' is factually true if and only if
- (i) the ostensible referent of *S* is an actual particular in the world, and
- (ii) the ostensible referent of *S* has an intrinsic property which has the same extension as does the ostensible referent of the predicate *P*.

Factual truth conditions for simple relational propositions are defined analogously, and the analysis can be extended to compound propositions, universal and existential propositions, and so on, in standard ways.

Since, on this account of factual truth, the world consists of particulars which possess intrinsic properties and enter into various intrinsic relations, the phenomena within a theory's intended scope will be systems of particulars in the world which possess certain kinds of properties and enter into certain types of relations; these particulars, their properties, and the relations they enter into need not be observable, though they may be. Such systems will be known as *phenomenal systems*, and each causally possible phenomenon is a *causally possible phenomenal system*. From our previous discussion, there will be at least one causally possible physical system corresponding to each causally possible phenomenal system for a given theory.

Can elementary propositions be used to describe phenomenal systems? Elementary propositions, it will be recalled, make assertions to the effect that a certain physical parameter p has a physical quantity q at time t. But these physical parameters are nothing other than *kinds of attributes* which physical particulars may possess, and physical quantities are attributes of the requisite kinds which the particulars may possess. For example, in classical particle mechanics the physical parameters are the position and momentum coordinates of bodies, and so the elementary propositions of that theory each specify a position or a momentum coordinate of a body; that is, an attribute of a particular (a body) in the phenomenal system at time t. As such, elementary propositions can be used to refer to particulars in phenomenal systems and predicate attributes of them. When so used, the propositions will be factually true of the phenomenal system if and only if the particulars referred to have the attributes indicated. In a manner analogous to that for physical systems, nonelementary propositions in the expanded theory formulation language may be used with reference to phenomenal systems. Here the truth conditions are defined analogously relative to the truth or falsity of elementary propositions. If the theory is empirically false, there may be propositions true of causally possible phenomenal systems that are false of all theory-induced physical systems and also false of the theory.

If a theory is empirically true, does it follow that all propositions true of causally possible phenomenal systems in the theory's intended scope are true of some theory-induced physical system, hence true of the theory? To answer this we will have to investigate the relations holding between phenomenal systems and their corresponding physical systems (for the answer, see the end of sec. V).

IV. STRUCTURAL RELATIONS

Physical systems and phenomenal systems are relational systems. A physical system is a relational system having a domain of states and attributes defined over these states. A phenomenal system is a relational system having a domain of particulars and attributes defined over these particulars; these attributes include properties the particulars possess at various times and relations they enter into at various times. Collectively they are the behavior of the phenomenal system. Let *S* be a physical system corresponding to a phenomenal system *P*. *S* is characterized in terms of various parameters abstracted from *P*; these parameters are *kinds of attributes* the particulars in *P* may possess. A state in *S* is a simultaneous set of values for these parameters; but the value of parameters are physical quantities, which is to say that the values are attributes which particulars could possess.[25] Thus a state of *S* is a possible state of affairs *P* could assume (being a possible set of simultaneous attributes the particulars in *P* could possess). An important consequence follows from this discussion: Since states in *S* are possible simultaneous attributes of particulars in *P*, the states are not themselves particulars even though they are theoretical entities. Theories and physical systems have sets of theoretical entities (the states) as their domains, but do not contain any particulars in their domains. Phenomenal systems, on the other hand, do not have theoretical entities in their domains, having particulars instead.

The attributes in *S* determine a sequence of states over time and thus indicate a possible behavior of *P* (i.e., a sequence of changing attributes the particulars in *P* could have at various times). Accordingly, *S* is a kind of *replica* of *P*; however, it need not replicate *P* in any straightforward manner. For the state of *S* at *t* does not indicate what attributes the particulars in *P* possess at *t*; rather, it indicates what attributes they *would have* at *t were* the abstracted parameters the only ones influencing the behavior of *P* and were certain idealized conditions met. In order to see how *S* replicates *P* we need to investigate these abstractive and idealizing relations holding between them.

Begin by considering *abstraction*. Scientific theories do not describe the behavior of phenomenal systems in all their complexity, but rather attempt to characterize the phenomena in terms of a few selected parameters. For example, in classical particle mechanics, one treats mechanical phenomena involving interacting bodies *as if* the phenomena *only* involved certain specified bodies and *only* involved the positions and momenta of these bodies. As such, a mechanical phenomenon is treated as an *n*-body system whose behavior is wholly specified in terms of $6n$ position and momentum coordinates (parameters). In effect, one assumes the fiction that no other bodies and no other parameters exert an influence on these n bodies' behaviors.

Under certain circumstances this fiction can be approximately realized—if, for example, the experimental setup is such that other bodies and other parameters exert a negligible influence. In such special cases, we have abstracted certain aspects of the systems' behavior and duplicated them with a physical system. And any difference between the actual behavior of the phenomenal system and the corresponding states in the physical system will be negligible. In most circumstances, however, the fiction is not realized—other bodies or other parameters do exert a nonnegligible influence on the behavior of the physical system. For example, the density of the medium in which the bodies act is such that they retard motion, or the system is not frictionless, or other massive bodies are present altering the motion of the specified bodies. In such cases, the values of the parameters characteristic of the state s that the physical system S is in at t are *not* the actual parameter values characteristic of the phenomenal system P at t. Rather, they stand in the following *replicating relation* to the values in P:

> If P were an isolated phenomenal system in which all other parameters exerted a negligible influence, then the physical quantities characteristic of those parameters abstracted from P would be identical with those values characteristic of the state at t of the physical system S corresponding to P.

In this case, any proposition in the expanded theory formulation language true of S would be counterfactually true of P (being an assertion about what P's behavior would be were certain conditions met).

There are two ways this counterfactual could be true of P, and the difference between the two cases marks the difference between pure abstraction and idealization. First, it may be causally possible for P to realize the conditions such that P's behavior would be as S indicates. For example, an electrically charged sphere phenomenon might be such that, through experimental controls, the phenomenon could have occurred in circumstances where it in effect was isolated from other influences. Second, it may be such that it is causally impossible for P to realize the conditions such that P's behavior would be as S indicates. For example, in classical particle mechanics (as opposed to rigid body mechanics) spatial location (position) is a parameter, and the possible values for the parameter are the coordinates of points in space. This in effect imposes the requirement that at each time a body must have a unique point in space as its spatial location, which in turn requires that the body be extensionless or dimensionless. But this condition is causally impossible, since it would require that bodies have an infinite gravitational potential, which is impossible if classical particle mechanics is true.[26] As such, it is causally impossible, then, to put the physical system P in circumstances that would meet the conditions required to make statements about S (noncounterfactually) true of P. The reason why these conditions cannot be met is that they require that parameters characteristic of phenomenal systems have physical quantities as values which are causally impossible for any phenomenal system to possess.

The former case (causally possible) is an example of *pure abstraction*, and the latter case (causally impossible) is one of *idealization*. Roughly put, the difference between the two cases is this: In cases of pure abstraction, the theory simply ignores a number of parameters and other factors which in most cases exert a nonnegligible influence on the phenomena within the theory's intended scope, but it is causally possible that there could be phenomena within the theory's intended scope on which only the abstracted parameters exert nonnegligible influences. Cases of idealization also involve abstraction, but in a way that imposes conditions on the phenomena which are causally impossible for any phenomena to meet.

A physical system S corresponding to a phenomenal system P must be such that the above counterfactual condition is met. This condition can be met either by S being an abstract replica of P, or else by S being an idealized replica of P—depending on the circumstances in which it is causally possible that phenomena within the theory's intended scope can be realized.

V. THE EMPIRICAL TRUTH OF THEORIES

The next step in our investigation of the relations between phenomena, physical systems, theories, and linguistic formulations of theories is to consider what it is for a theory to be empirically true. As stated previously, there are two empirical truth conditions for theories:

(1) If the theory is empirically true, then the class of theory-induced physical systems is identical with the class of causally possible physical systems for the theory.
(2) If the class of theory-induced physical systems is not identical with the class of causally possible physical systems for the theory, then the theory is empirically false.

The first condition is a necessary condition for the empirical truth of a theory, but is it a sufficient condition? The second condition is a sufficient condition for a theory being empirically false, but is it a necessary condition? These questions must be answered if we are to know what it is for a theory to be empirically true or false.

First, let us see what is involved in asserting a theory. A theory has as its intended scope a natural kind class of phenomenal systems. In propounding a theory, one commits oneself to the existence of the phenomenal systems within the theory's intended scope.[27] Further, one presents a theory structure and asserts that the class of theory-induced physical systems is identical with the class of causally possible physical systems. This in turn commits one to the existence of a class of causally possible physical systems which *are* abstract and/or idealized replicas of the phenomena within the theory's intended scope.

A physical system is an abstract replica of a phenomenal system if for each t, its state stands in the counterfactual replicating relation to its corresponding phenomenal system. This replicating relation obtains only if the defining parameters of the physical system are parameters abstracted from phenomenal systems in the theory's intended scope or else are idealizations of them. Recall (from sec. III) that these defining parameters must be kinds of attributes characteristic of the particulars in phenomenal systems; hence the existence of the class of causally possible physical systems for a theory presupposes the existence of particulars having the kinds of attributes selected as defining parameters of physical systems for the theory.

Further, if the physical systems in the class of causally possible physical systems are pure abstractions of phenomenal systems, one also commits oneself to the notion that the physical quantities that are values of parameters in casually possible physical systems must also be attributes of causally possible phenomenal systems within the theory's scope. However, if the physical systems in the class of causally possible physical systems are idealized replicas of causally possible phenomenal systems, one does not commit oneself to the values of parameters in the causally possible physical systems being attributes of particulars in causally possible phenomenal systems. Rather, one commits oneself to their being idealizations of attributes which particulars in causally possible phenomenal systems do possess.

In order for a theory to be empirically true, all the commitments made in propounding the theory must be satisfied. Based on the previous discussion, these are the requirements for a theory to be empirically true:

> Let T be a theory with intended scope I whose defining parameters are $p_1, \ldots, p_n$. Then T is *empirically true* if and only if

(a) I is a natural kind set of causally possible phenomenal systems whose domains contain particulars of which $p_1, \ldots, p_n$ are characteristic kinds of attributes.
(b) The possible values of the parameters $p_1, \ldots, p_n$ allowed by T are attributes which particulars in the phenomenal systems in I do possess or else are idealizations of such attributes.[28]
(c) The set of theory-induced physical systems for I is identical with the class of causally possible physical systems for T.

If any of these three conditions fail to be met by T, then T is *empirically false.*

The requirement that I be a natural kind set in effect commits my version of the Semantic Conception to an antinominalism (to a view that some classes correspond to real divisions of nature and others do not) that van Fraassen would reject. I must postpone defense of that condition until chapter 7, where I develop and defend a specific version of the antinominalistic thesis.

Whenever a theory is propounded as *being empirically true*, a commitment to requirements (a) through (c) being satisfied is made. However, it is possible to propound a theory without propounding it as being empirically true; one may propound it instead as being some kind of *conceptual device.* For example, the theory may be known to be empirically false but approximately correct for certain types of phenomena within the theory's intended scope, and then used to make approximate predictions within that range of phenomena. For example, the ideal gas laws are known to be empirically false, yet they are used in this way. In such a case, the theory is propounded as *being approximately true for a restricted scope* $I' \subset I$, which carries the commitment that the restrictions of (a) through (c) to I' are satisfied.

A theory also can be propounded as *a simplification.* For example, for certain purposes it may prove convenient to employ an astronomical theory which considers planetary orbits to be circular rather than elliptical, even though the theory is known to be empirically false. Here there is no question whether the theory is empirically true—it is known to be empirically false since it fails to satisfy truth condition (c). And unlike the approximation case, the theory is not even true for a restricted scope I' contained in the intended scope I of the theory. Rather, it is used because it conveniently yields incorrect predictions which are close enough for the purposes at hand.[29]

Let T be a theory which is empirically true, and let ϕ be a proposition in the theory formulation language which is true of some physical system in the class of theory-induced physical systems for the theory. Since the theory is empirically true, ϕ will be true of some causally possible physical system for T; hence by the replicating relation, ϕ will be counterfactually true of some causally possible phenomenal system. And if the counterfactual conditions imposed on the physical system were true of the phenomenal system, then ϕ would be factually true of the phenomenal system. A proposition ψ factually true of a causally possible phenomenal system usually will not be noncounterfactually true of its corresponding physical system; whether it is counterfactually true of the physical system is an open question.[30]

VI. INSTRUMENTALIST VERSUS REALIST CONSTRUALS OF THEORIES

Our findings on the empirical truth conditions just discussed yield certain consequences for the issue whether theories are to be construed instrumentally or realistically which will be helpful in assessing the operational imperative. In its classic formulation, the instrumentalist-realist controversy has centered on the question whether statements of laws or theories incorporating nonobservable theoretical entities are factually true or false, or whether they are merely calculating devices for predicting observable phenomena. The realist answers the question as follows:

> Scientific theories are factually true or false; hence if true, the theoretical (i.e., nonobservable) entities postulated by the theory do exist.

The Instrumentalist answers the question this way:

> Insofar as scientific theories employ theoretical terms (i.e., nonlogical terms which do not designate observable entities or attributes), they are not factually true or false. Only if the theory makes no reference to theoretical entities is it factually true or false. Theories incorporating theoretical terms are mere calculating devices for predicting and controlling observable phenomena.

Typically both realists and instrumentalists identify theories with their linguistic formulations, though it is possible to divorce the main tenets of realism and instrumentalism from that identification.

The realist claims that theoretical assertions are factually true or false. Under the analysis of theories being considered here, the realist in effect is claiming that propositions in the theory formulation language which are true of the theory will be factually true of the phenomena within the theory's intended scope if the theory is empirically true. Our findings have shown that this generally will not be so; rather, it will be the case that such propositions will be counterfactually true of the phenomena. The realist also claims that the theoretical (i.e., nonobservable) entities postulated by the theory do exist if the theory is empirically true. Construed in terms of our analysis of theories, this becomes the claim that (a) that the particulars (both observable and nonobservable) in the phenomenal systems within the theory's intended scope do exist and have the kinds of properties (both observable and nonobservable) which are the theory's defining parameters, and (b) that the physical quantities which may be values of the theory's parameters are properties possessed by particulars in causally possible phenomenal systems. This in effect requires that the class of causally possible physical systems for a theory not be idealizations of phenomenal systems. We have seen, however, that there are theories (e.g., classical particle mechanics) which do incorporate idealizations, and so realism is descriptively false of the theories science actually uses.

Realism's mistake rests, I think, in the common but mistaken suppositions that all theoretical entities are particulars and that if a realistic theory is to be empirically true, the theoretical entities must exist. But, as Wilfrid Sellars has shown (1956, sec. 61), not all theoretical entities need be particulars; indeed, the states of theories and physical systems employed here are examples of theoretical entities that are not particulars. And as Dudley Shapere has shown (1969, Part II), theoretical entities can be invoked in a theory without committing oneself to their existence; idealizations typically are theoretical entities used nonexistentially. Thus idealizations can be employed in theories without committing oneself to their existence.

In particular, when the physical system S corresponding to a phenomenal system P is an idealized replica of P, one need not commit oneself to the idealized values of parameters being properties possessed by particulars in causally possible phenomenal systems. And using such idealized values does not preclude the theory from being empirically true or false, for by the empirical truth conditions of the previous section, theories are *counterfactually* true or false of the phenomena within their intended scopes, regardless whether causally possible physical systems are purely abstractive or idealized replicas of phenomena. The realists are correct, then, in supposing that theories are empirically true or false and may commit one to the existence of nonobservable particulars or attributes, but they are wrong in identifying empirical truth or falsity with factual truth or falsity and in supposing that to invoke theoretical entities in a theory *always* is to commit oneself to their existence as particulars. When these errors are eliminated from realism, a modified or *quasi-realism* emerges which is defensible and therefore is built into my empirical truth conditions for theories.

Viewed as a descriptive account of theories, instrumentalism fares worse: It becomes the claim that there are no scientific theories which deal with phenomena involving nonobservable particulars. For if theories did, then properties of these particulars would have to be parameters of the theory, which in turn means (by the empirical truth conditions) that the theory is empirically true or false—contrary to instrumentalism's claims. Since scientific theories dealing with phenomena of nonobservable particulars do exist, instrumentalism is descriptively false of actual science. If instrumentalism is to be viable, then, it must be viewed as a prescriptive thesis—saying that the only theories which should be used are those which presuppose only observable particulars and attributes. From the foregoing analysis of theories and from the empirical truth conditions for such theories it follows that such theories may employ no parameters of particulars *whose values* are not (in some cases at least) observable. If followed, prescriptive instrumentalism would rule out most present-day scientific theories.

To summarize, classical realism is untenable if my analysis of theories is correct; but modified or quasi-realism not only is tenable, but in effect is built into the analysis.[31] In addition, instrumentalism appears to be a conservative doctrine, demanding that science should restrict its attention just to observable phenomena; in effect, it is a modified realism restricted to a limited class of phenomena.

[Recall from chapter 1 that the classical realism-versus-instrumentalism debates have been replaced by what MacKinnon (1979) terms "the new debates" over scientific realism.[32] The Semantic Conception proves to be a particularly effective vehicle for displaying some fundamental oppositions within the new debates. For example, van Fraassen and I agree on nearly all aspects of the Semantic Conception except that he champions an antirealistic approach to the physical interpretation of theories, whereas I embrace a quasi-realistic one. The heart of my quasi-realism is the empirical truth analysis of theories expounded in section V and the replicating relation of section IV above. Van Fraassen rejects this quasi-realism, denying that theories are empirically true or false either in some literal sense or in accordance with my counterfactual ("nonliteral") empirical truth analysis. Instead he says we evaluate theories as to their "empirical adequacy":

> To present a theory is to specify a family of structures, its *models*; and secondly, to specify certain parts of those models (the *empirical substructures*) as candidates for the direct representation of ob-

> servable phenomena. The structures which can be described in experimental and measurement reports we can call *appearances*: The theory is empirically adequate if it has some model such that all appearances are isomorphic to empirical substructures of that model. [1980, 64; italics original]

Further, van Fraassen requires that "empirical adequacy concerns actual phenomena: what does happen, and not what would happen under different circumstances" (ibid., 60). By virtue of his antirealism's reliance on an observational/nonobservational distinction, it also is a modified realism restricted to a limited class (or aspect) of phenomena; this feature of his view is criticized in chapters 1 and 11.

The preceding discussion of the physical interpretation of theories provides another line of objection: My arguments in sections IV and V strongly support the contention that theories virtually always abstract their phenomena—which is to say, treat them as if they were isolated from outside influences. Moreover, this claim is separable from my quasi-realistic account of the empirical truth of theories. Such a process of abstraction carries no guarantee that any of the theory's models or substructures will be isomorphic to any actual phenomenal systems; hence there is no guarantee that there will be any models such that all appearances are isomorphic to empirical substructures of the model.

Empirical adequacy minimally would require that every phenomenal system within the theory's scope actually occur in isolated circumstances—which is a huge philosophical assumption (by no means obvious, but possibly defensible) that requires argument and defense. But even this is insufficient to establish that "*all* appearances are isomorphic to empirical substructures of that model," for in paradigmatic cases such as Newton's theory (an example van Fraassen uses [1980, 46]), appearances more often than not are at odds with the models of the theory, hence are not isomorphic. For example, real inclined planes usually are neither frictionless nor have constant coefficients of friction. In short, to the extent that empirical adequacy concerns "what does happen, and not what would happen under different circumstances" for *all* appearances, virtually no real-life scientific theory ever meets van Fraassen's empirical adequacy conditions. Differently put, van Fraassen's antirealism is a literal realism when it comes to appearances; my quasi-realism, with rare exceptions, is the most that is defensible—even for appearances.

The only way I can see to salvage van Fraassen's empirical adequacy conditions is to allow the idea that empirical adequacy concerns what would happen under isolated conditions, as well as what sometimes may happen under actual circumstances—which is to adopt something like the replicating relationship of section IV. Of course, this would be tantamount to abandoning his antirealism in favor of some sort of quasi-realism.]

VII. EXPERIMENTAL METHODOLOGY

Before turning to an assessment of the operational imperative, there is one last aspect of the relations holding between theory and phenomena which should be explored—how experimental methodology enables us to determine the states of a causally possible physical system by observing and measuring, and so on, the corresponding phenomenal system.

The most basic feature of the relation holding between a casually possible phenomenal system *P* and its corresponding physical system *S* is that *S* is a (possibly idealized) counterfactual replica of *P* satisfying the replicating relation discussed in section IV. This condition must be met regardless whether the particulars in *P*, their attributes, and the defining parameters of the theory are observable or not. For all empirically true theories stand in a counterfactual relationship to their phenomena regardless what the nature of the phenomenal system is (even if an occasional causally possible physical system is a noncounterfactual replica of its corresponding phenomenal system).[33]

In applying a theory to specific phenomena, it is the task of the experimental methodology to convert data about a phenomenal system *P* into data about the corresponding physical system *S* in such a way as to satisfy the replicating relation. Suppose, for example, that we wish to use the theory to predict the subsequent behavior of a phenomenal system *P*. In accordance with the discussion at the end of section I, we first determine the actual values of various parameters in *P* at *t*—we determine the physical quantities (attributes) of the various particulars in *P* at *t*. Then, in accordance with some experimental design, we convert this data about *P* into data about the corresponding physical system *S*—that is, we convert the data about *P* into data about what *P* would have been had *P* been isolated, dependent only on the defining parameters of *S* and the various idealized conditions met.

The *experimental design* governing this process is not part of the theory; rather, it is based on various extratheoretical regularities, other theories, laws, known regularities about the kind of phenomena involved, and so on. For example, for inclined planes we may employ various other theories which enable one to determine friction coefficients; for bodies falling in viscous media, we use various hydrodynamic laws to determine the retarding effects of the medium; for falling bodies we idealize the fiction that the position of the body is its center of gravity. These various laws, theories, and assumptions in effect are putative regularities which hold between the theory's parameters and other factors affecting them, and if the regularities do obtain for *P* (if the laws, theories, and assumptions are empirically true of *P*), then their use (together with experimental controls which succeed in keeping any other factors from influencing the phenomena) enables one to determine what states *S* is in at *t*. The truth of these then guarantees that the replicating relation is satisfied.[34] In different circumstances, and for different *P*s, different laws and theories will be used to determine that state of *S* at *t*. And there may be *P*s or circumstances in which no laws or theories exist which enable one to determine what state the *S* corresponding to *P* will be in at *t*.

Although the above discussion is exceedingly crude and does not even suggest the full complexity of the way experimental methodology enables us to determine the state of *S* at *t* from data about *P* at *t*,[35] it does enable us to draw several conclusions relevant to our assessment of the operational imperative.

(1) In order to apply a theory to a phenomenal system *P*, data about *P* must be converted into data about the corresponding physical system *S*; this must be done regardless whether the defining parameters in *S* and the particulars in *P* are observable or nonobservable.
(2) The conversion of data about *P* to data about *S* requires the employment of laws, theories, known regularities, and so on, which are not a part of the theory in question, as well as whatever idealizing conventions are associated with the theory. These are applied to data about *P*.
(3) Different phenomenal systems *P* within the theory's intended scope may require different laws, theories, and so on, to effect the conversion of data about *P* to data about *S*, and there may be *P*s such that we do not know how to effect the conversion.

These conclusions hold for all theories regardless whether the physical quantities which are values of the defining parameters of states in *S* are qualitative or measurable.

When the physical quantities involved in applying a theory to phenomena are measurable, certain complications result. The process of measurement consists of attaching numbers to physical quantities which describe the physical quantities. *Quantitative theories* are ones in which all of the defining parameters of the theory are measurable. These theories admit of *mathematic models,* and typically these models are employed in lieu of the theory.

I will briefly characterize them as follows: Let T be a theory with $p_1, \ldots, p_n$ having determinate attributes as values[36] and assume that $p_1, \ldots, p_n$ are measurable. Since states of physical systems are n-tuples of physical quantities $q_1, \ldots, q_n$, which are possible values of $p_1, \ldots, p_n$, we can represent or characterize states by numbers which are measures of $q_1, \ldots, q_n$. Such measures are determined relative to a *frame of reference,* and different frames of reference typically will assign different numbers to the same physical quantities. The various measures assigned to a given physical quantity under different frames of reference must be equivalent under some *system of transformations* (e.g., the Galilean system of transformations in classical particle mechanics or the Lorentz system of transformations in the special theory of relativity). For to say that a physical quantity is measurable is to say, inter alia, that it can be described by numbers up to some system of transformations.[37] These systems of transformations in effect are putative laws stating that certain physical quantities remain invariant under certain changes in frames of reference, measurement techniques, and so on. And the system of transformations is included among the laws, theories, and so forth, (mentioned in point [2] regarding methodology) used in converting data about P into data about S.

When a theory employs measurable parameters, formulations of the theory typically employ a mathematic vocabulary, and states of S are represented by n-tuples of numbers (or functions over numbers) which describe the physical quantities characteristic of the state in question. In such cases S often is represented by an *arithmetic model* describing S relative to some frame of reference, where the domain of the model consists of n-tuples of numbers which describe the states relative to the frame of reference, and where the behavior of S is represented by a sequence of these n-tuples of numbers. Similarly, the theory will be represented by a *phase space* model, which is an n-dimensional space whose points represent the states whose physical quantities are described by the coordinates of the point under the frame of reference in question. The laws of the theory are represented by trajectories directed through the space against time, subspaces, and the like.

Admissible frames of reference for a theory must be such that every possible physical system state can be represented by an n-tuple of numbers assigned to physical quantities relative to that frame of reference.[38] This enables us to construe each n-tuple of numbers representing states relative to a frame of reference as being coordinates of points in some n-dimensional vector space. And various configurations can be imposed on that space, such as trajectories, subspaces, probability measures, and the like. In particular, we can impose configurations on the points of such a space which correspond to laws of succession, laws of coexistence, and laws of interaction. When such a space, obtained relative to a frame of reference, has such configurations imposed on it corresponding to the attributes of the theory, we say the space is a *phase space model of the theory.*

Phase space models can be given linguistic formulations which are analogous to theory formulations as characterized in section I. Such formulations are called *model formulations* and are given in an (expanded) *model formulation language*, which is a mathematical *language* having the same structural properties as the (expanded) theory formulation language. In particular, it will have elementary propositions, a logic of the model, and a language of physical description. Deductions in this language are undertaken by solving equations.

An appropriately chosen phase space model will be isomorphic to the theory, and propositions will be true of the model if and only if their analogues in the theory formulation language are true of the theory. Likewise, the quasi-realistic interpretation of propositions in the model formulation language and the model will be analogous to that for propositions in the theory formulation language and the theory. As such, whenever a theory admits of phase space models, these models may be employed in lieu of the theory and model formulation languages may be employed in lieu of theory formulation languages. In actual scientific practice this is what invariably is done.[39]

VIII. TENABILITY OF THE OPERATIONAL IMPERATIVE

We now have discussed enough about the relations holding between theories, physical systems, phenomena, and theory formulations to assess the operational imperative. We concentrate our attention on the weak operational imperative, which says that every scientific theory must admit of a full linguistic formulation in which theoretical terms are operationally defined. Suppose we have a full formulation F of a theory T meeting the operational imperative. Then it must be possible to divide the nonlogical terms of the theory formulation language for F into *observation terms* which designate observable particulars or attributes and *theoretical terms* which designate nonobservable particulars or attributes.[40] And the formulation F must include propositions of the following form for each theoretical term Q:

> X is Q if and only if [or, only if] performing operations O in circumstances C would yield results R,

where O, C, and R do not involve theoretical terms. These are the *operational definitions*. For simplicity of argument we will assume that the theoretical terms Q designate defining parameters of the theory's physical system or physical quantities thereof, and we also will assume that physical systems are *not* idealized replicas of phenomenal systems.[41] The observational terms used to define O, C, and R may or may not designate defining parameters of the theory or attributes thereof, but they clearly must designate attributes or kinds of attributes possessed by particulars in phenomenal systems within the theory's intended scope.

Suppose T is empirically true of the phenomena within its intended scope. Then by the results of section V the sorts of particulars asserted to exist by the theory do exist in causally possible systems, and the defining parameters of the theory and physical quantities thereof are kinds of attributes and attributes possessed by these particulars. Hence the Q in the operational definitions designate attributes or kinds of attributes possessed by particulars in the phenomenal systems. Since T is empirically true, the propositions in F are counterfactually true of the phenomenal systems if they meaningfully can be asserted of these phenomenal systems (secs. III, V). But the Q, O, C, and R do designate attributes or kinds of attributes which can be possessed by particulars in these phenomenal systems, so the operational definitions are counterfactually true of the phenomenal systems. But then, the operational definitions specify that certain regularities hold between particulars being Q and particulars being R when operations O are performed on them in circumstances C. As such, operational definitions are true or false statements of empirical regularities holding between particulars in phenomenal systems. This precludes operational definitions from being definitions in the usual stipulative sense; for empirical regularities cannot be defined into existence, and definitions are analytically true whereas statements of empirical regularity cannot be analytically true.

The discovery that operational definitions are contingent statements of empirical regularity enables us to see the operational imperative in its true color. For we now see that the operational imperative in effect prohibits the employment of nonobservable parameters (or values thereof) in theories unless it is possible to state empirical regularities holding between these parameters (or values) and other observable parameters (or values). In a word, one cannot incorporate parameters (or values thereof) into testable theories unless one knows some of their observable manifestations. On the other hand, any observable parameters (or values thereof) are admissible. This makes the question whether operational definitions should be explicit definitions or reduction sentences rather silly; for to demand that operational definitions be explicit definitions is nothing other than a refusal to theorize with parameters which do not manifest themselves observationally the same way in all phenomenal systems within the theory's intended scope.

As we saw in section II, the operational imperative, inter alia, is supposed to explain how theories with nonobservable defining parameters can be tested and confirmed observationally. To see whether it does, first consider what happens if the strong operational imperative holds. To test a theory with nonobservable parameters it is not sufficient merely that the operational definitions in *F* be true *if* the theory is empirically true. For one tests a theory by comparing the theory's predictions (obtained in accordance with the procedures of sections I and VII) about *P* with *P*'s actual behavior. And this requires knowing the actual values in *P* of *T*'s defining parameters independently of the theory's empirical truth or falsity. When these parameters or values are nonobservable, the operational definitions are supposed to enable one to determine the actual values of the parameters in *P*, but they can do so with reliability only if the operational definitions are known to be true of the phenomenal systems prior to testing the theory *T*.[42] But the operational definitions—being empirical statements of regularities holding between observable and nonobservable parameters or values thereof (when all other parameters exert a negligible influence on the regularities)—are themselves formulations of (simpler) theories. And since the operational definitions must be known to be true, by the discussion of sections III through V, the theories they formulate must be known to be empirically true prior to the testing of *T*. How do we determine that this is so? By testing the theories, it seems. But under standard accounts this requires testing their predictions against the actual attributes of particulars in the phenomenal systems within their intended scope, and as before, to do this one must know in advance the actual attributes of those particulars. However, some of these attributes are nonobservable (namely, those designated by the term being operationally defined), and if the testing is to be noncircular, one must have some other theory (incorporating the laws, theories, and so on, of section VII) specifying empirical regularities holding between the nonobservable attributes and other observable attributes under various causally possible circumstances. Moreover, this theory would have to be known to be empirically true prior to the testing. The catch is that the same arguments apply to *this* theory, showing the need for still another prior theory—and so on into an infinite regress. Since this is a vicious regress, it follows that no theory embodying nonobservable parameters or values thereof can be tested. *Thus the strong operational imperative is untenable.* Rather than explaining how testable theories can employ nonobservable parameters or values thereof, the strong operational imperative in effect precludes the employment of such parameters or values, which in turn forces one to adopt the instrumentalist account of theories as characterized in section VI.

Next, consider whether the weak operational imperative does any better in explaining how theories employing nonobservable parameters or attributes thereof are observationally tested. Unlike the strong imperative, the weak operational imperative does not require that one only employ formulations of a theory in which theoretical terms are operationally defined. Indeed, it does not even require that users of a theory *possess* a formulation of the theory in which all theoretical terms are operationally defined; it only requires that the theory *admit* of such formulations. It follows from the previous arguments, then, that if an empirically true theory admits of such a formulation, the operational definitions will be counterfactually true statements of regularities holding between particulars—that is, that the nonobservable parameters or values thereof employed in the theory do admit of observable manifestations which can be operationally discerned. It is this fact that explains how it is that theories employing nonobservable parameters or values thereof can be observationally tested; it is surely true that one cannot observationally test a theory about phenomena which do not have observable manifestations. If this is all there is to the weak observational imperative, then it is defensible.

Few proponents of the operational imperative would be willing to say that this is all there is to the operational imperative, though, for the operational imperative is intimately connected with the following epistemological doctrine: Determining the truth or falsity of propositions about observables is straightforward and nonproblematic, but knowledge about nonobservables is problematic and ultimately must rest on our knowledge about observables. It is operational definitions which enable us to expand our knowledge of the observable to knowledge of the nonobservable: From our observable knowledge and operational definitions we are able to deductively gain knowledge of nonobservables.

There is some plausibility to this thesis *if* operational definitions are thought to be definitions, hence, analytically true. For then, from our sure knowledge of observables and our operational definitions which are nonproblematic as to truth, whatever we deduce about nonobservables must be true; hence operational definitions yield knowledge of the unobservable. However, we have seen that operational definitions are not really definitions and are not analytically true. Rather, they are counterfactual empirical truths about particulars and their attributes, some of which are nonobservable.[43] Hence if operational definitions are to enable us to deductively expand our observational knowledge to knowledge of the unobservable, the operational definitions must be known to be empirically true.

But this requirement gets the weak imperative into the same regress as the strong operational imperative does; hence, it does not explain how we come to have theoretical knowledge about the nonobservable. And if the only way we have of obtaining knowledge of the nonobservable is by using operational definitions to deductively expand our knowledge of the observable, then there can be no knowledge of the nonobservable. The epistemology of the operational imperative thus is a skeptical one which ultimately forces science into an instrumentalism.

The fact that science does yield knowledge of nonobservable particulars and attributes indicates that there is something drastically wrong with the operational imperative's associated epistemology. Part of the difficulty clearly rests in the operational imperative's embrace of the observational-theoretical term distinction which, as argued in chapter 2, is untenable in that it cannot be drawn on the basis of natural language usage. Indeed, without that distinction it is very hard to give a coherent formulation of the operational imperative. However, I suggest that the difficulty with the operational imperative rests with deeper problems with its associated epistemology—in its suppositions that one can *only* have observational knowledge of attributes of particulars which can be directly seen, and that knowledge of attributes and particulars which cannot be directly seen must be obtained *inferentially* from observational knowledge of what can be directly seen.

This thesis, I suggest, is fundamentally mistaken. While it is true that our sensations are essential to having observational knowledge, the sensations we have in various circumstances of observation are the product of empirical regularities involving both "observable" and "nonobservable" particulars (particulars we can see and ones we cannot see). What we observationally are able to determine to be the case depends on these regularities regardless whether the objects of our observational knowledge themselves can be seen or not. In fact, several recent philosophical analyses of perceptual and observational knowledge strongly suggest that these regularities enable us to have noninferential observational knowledge of particulars and attributes of particulars which we are unable to directly see; also, this knowledge can be had without knowing what those regularities are, provided that we know how to use language to describe them.[44] (For example, while looking at a bubble chamber I may be able to observe that or see that an α particle has been emitted even though I cannot see the α particle, and I thereby gain *observational*—not inferential—knowledge that an α particle has been emitted). Furthermore, it is possible to know how to use language to describe particulars which cannot be directly seen and their attributes without possessing operational definitions of them. (Consider, e.g., descriptions of men "too little to be seen" in children's stories.)

If these claims are philosophically defensible (and there is good reason to suppose they are), then the operational imperative is pointless, for then we can have observational knowledge of particulars and attributes which cannot be directly seen. Such knowledge enables us to test and confirm theories about such particulars and attributes; these theories can then be used to test and confirm other theories in accordance with the procedures of section VII. As such, the operational imperative seems plausible only if certain questionable epistemological suppositions are made about the nature of observation.

IX. CONCLUSION

We have seen that the operational imperative is a prescriptive thesis about formulations of theories, which imposes restrictions on the sorts of theories science may employ. We assessed the operational imperative by investigating a number of relationships holding between theory formulations, theories, physical systems, and phenomena and then applying our findings to the operational imperative. These applications showed that the operational definitions required by the operational imperative were not definitions at all, but rather statements of putative empirical regularities holding between particulars, which in effect are formulations of empirically true or false theories. From this fact it followed that the supposed epistemic advantages of following the operational imperative fail to obtain: Operational definitions do not enable one to go deductively from knowledge of observables to knowledge of nonobservables, and operational definitions do not provide a means for testing theories about unobservable phenomena. As such, the operational imperative should be rejected in both its weak and strong versions. However, we did discover a grain of truth in the operational imperative; namely, that theories with nonobservable parameters are testable only if these parameters have observable manifestations. But that grain of truth does not lead to the operational imperative as typically advanced, unless one embraces certain epistemological theories about observation which recent work on observation makes highly doubtful.

NOTES

1. The operational imperative apparently was first introduced by Bridgman in 1927 and elaborated in a number of his subsequent works. However, earlier anticipations of the operational imperative are found in C. S. Pierce's pragmatic theory of meaning (see Burks [1977] for discussion).

2. See Hempel (1954, reprint, 1965, 126). He imposes the observation language requirement on the specification of operations and results.

3. For arguments that this is so, cf. Hempel 1952, secs. 5–7; and Suppe 1974b, sec. II–B.

4. For an explicit statement of this argument, see Hempel 1952, 41. Hempel 1954 is largely devoted to developing the same line of argument. In it, Hempel also raises other objections against specific versions of the operational imperative which are not based explicitly on the Received View; these objections do not constitute an attack on the operational imperative in general.

5. Cf. Achinstein 1968, chs. 3–6; Putnam 1962; Suppe 1974b, sec. IV. A critical assessment of Achinstein's and Putnam's attacks is found in ch. 2 above.

6. That the alternative linguistic formulations need not be equivalent follows from the fact that one can have partial formulations of theories (e.g., the difference equation formulations of classical mechanics found in high school physics texts). See section III for a characterization of partial formulations.

7. For simplicity of exposition my examples are taken from the physical sciences, which are particularly susceptible of the analysis. For examples showing the applicability of the analysis to theories in other branches of science, see chs. 2, 7, and 8.

8. Physical quantities need not be measurable quantities. For example, if color were a parameter, then the physical quantities which were possible values of the parameters might be colors as differentiated in a Munsell color table. Physical quantities also might be probability distribution functions, as in quantum theory.

9. The sequence may have a wide variety of different ordinal properties depending on whether, for example, time is viewed as discrete or continuous by the theory. Even though each physical system has a unique sequence of states, the theory may be unable to predict that sequence to the exclusion of all others.

10. Other requirements imposed on the theory, related to this basic one, might ask that it yield certain kinds of predictions, be compatible with certain other theories, or yield answers to certain types of questions, and so on.

11. This is a generalization of the analysis of empirical truth for theories introduced in ch. 2. Note that it only states a necessary condition for the empirical truth of theories. In sec. V, a set of necessary and sufficient conditions is given.

12. See sec. VII for an elaboration.

13. For a detailed discussion of how this works and for arguments sup-

porting the contention that confirmation of theories is reducible to goodness-of-fit problems, see Suppe 1967, sec. 3.4 and ch. 4 below. The inductive logic approach to confirmation as a vehicle for obtaining general or theoretical knowledge in science is rejected in ch. 13.

14. Laws are being construed as extralinguistic entities here; for a defense of the legitimacy of doing so, cf. Achinstein 1971, ch. 1; and Bunge 1959, ch. 10. Here and elsewhere in this chapter, short passages from my 1974c work have been interpolated.

15. The precise properties of these different kinds of laws are explained in ch. 5.

16. If the states are numerically specifiable, or if unique sets of n-tuples of numbers are assigned to each state, then the theory can be construed as a space upon which certain configurations (e.g., trajectories, branching trees, subspaces) have been imposed by the laws of the theory. Beth (1949) and van Fraassen (1970) identify theories with such phase spaces, whereas I tend to view them as canonical iconic models of theories. For a discussion of the relative merits of the two approaches, see Suppe (1974b, 227, n. 565). Phase space models of theories are discussed in sec. VII below.

17. This account only applies to theories with deterministic and indeterministic laws of succession and coexistence. How theories with other sorts of laws can be used to make predictions is discussed in chs. 5 and 6.

18. I am using 'proposition' here in the medieval sense—not the modern sense—as a linguistic entity which can be propounded but need not be asserted. Propositions contain sentences as components, but the sentences do not completely determine the proposition, since the same sentence can be used to express more than one proposition. Roughly speaking, when a proposition is asserted with reference to some subject matter, it becomes a statement. As such, propositions are interpreted declarative sentences which may be propounded with reference to one or more situations, states of affairs, things, and so on. See Geach (1962, 25) for a fuller characterization of this sense of 'proposition'.

19. Beth introduced this distinction (1963, 479–80). My formulation of it is somewhat different.

20. If the friction of the inclined plane were significant, then the proposition would describe only the phenomenon, and a different proposition, embodying a corrected value specifying the corresponding frictionless velocity, would be used with reference to the physical system and the theory. The general conclusion about the necessity of amplified usage still follows.

21. This requirement is rather significant, since not all states in the domain of the theory need be physically possible states. For example, in quantum theory the only physically possible states are those which satisfy the Heisenberg uncertainty principle, and so not all truth-functional combinations of elementary propositions satisfy this empirical significance requirement; as such, quantum theory requires a special quantum logic. My formulation of the empirical significance requirement is one of several possible ones.

22. See my 1973 article where the basic analysis is developed. The same analysis is partially developed and employed in ch. 7 below.

23. This characterization of complex particulars is not the only possible one. This formulation, though, yields individuation characteristics for complex particulars which allow Leibniz' law and the principle of absoluteness of identity to hold, and it also enables one to solve various problems concerning identity and spatio-temporal continuity. See my 1973 article for a discussion of these issues.

24. Cf. ch. 7 and especially Suppe 1973 for details and a discussion of the necessity of so construing the scientific world.

25. Note that physical quantities *never* are numbers. Rather, as J. von Neumann and O. Morgenstein so aptly put it, "[Measurable] physical quantities ... are described by numbers up to [some] ... system of transformations" (1953, 22). As we are using 'physical quantity', the term includes differentiable qualities such as colors, preferences, and so on. (See n. 8 above.)

26. For a discussion of this case, see Shapere 1969, 147–49.

27. More precisely, one commits oneself to the existence of some of the phenomenal systems in the intended scope (past, present, or future), and to the assumption that it is causally possible for the remaining phenomenal systems to exist.

28. [A case can be made for the contention that theories with idealizations really cannot be empirically true, but rather are another sort of conceptual device. My own inclination is that whether one wishes to make this move or employ the definition given in the text probably will hinge on the physical interpretation one makes of the causal modalities implicitly contained in the notion of "causally possible physical system." Presently I believe that the more restrictive definition of truth (where "or else idealizations thereof" is deleted from clause [ii]) probably will prove to be preferable on such metaphysical and related epistemological grounds. If this is done, minor but obvious adjustments will have to be made in the discussion that follows. If the more restrictive empirical truth definition is adopted, one can avoid the counterintuitive conclusion that it is impossible that classical particle mechanics *(CP)* could be empirically true by distinguishing it from *CP**, which is *CP* augmented by the principle that the position of a particular is the point at which its center of gravity is located. Thus, although *CP* could not possibly be empirically true, *CP** possibly could be. In actual practice, physicists employ CP* in most circumstances.] (Recall that brackets indicate new material.)

29. There are other ways in which approximations, simplifications, and idealizations occur in working with theories. Often in manipulating a theory, in using some theory formulation, one makes idealized assumptions, employs approximations, or simplifies the statements of problems to facilitate calculations; these approximations, simplifications, and idealizations may be inconsistent with the theory in question, yet are employed for convenience. Such cases need to be distinguished from the ones we have been considering where the theories themselves are propounded *as* idealizations, approximations, or simplifications.

A pioneering discussion of idealization, approximation, and the like, with respect to theories is to be found in part II of Shapere 1969. In writing this chapter, I have relied heavily on Shapere's work, which contains a number of useful detailed case studies. Shapere's 1969 discussion is marred, however, by his failure to clearly distinguish theories as idealizations, approximations, or simplifications from idealized, approximate, or simplified manipulations of theories; his discussion is concerned largely, but not exclusively, with the latter case. [Cartwright (1983) shares much of my concern with ceteris paribus conditions, idealizations, simplifications, etc., but takes her analysis in a more instrumentalistic direction. See also Hacking (1983).]

30. The answer to the question turns on the following question: Suppose that

If P which is A were C, then P would be B

is true. Under what circumstances does it follow that

If P which is B were C, then P would be A

is true?

31. The fact that a quasi-realism is built into my analysis of theories may lead philosophers, such as Feyerabend (1962a), who are of the opinion that no realistic interpretation has been given and/or can be given for quantum theory to suppose quantum theory is a counterexample to the analysis. Although I do not have enough space to give detailed consideration to the matter here, quantum theory is not a counterexample. However, the following comments may help to dispel the illusion that it is.

Under quasi-realism, a realistic interpretation of quantum theory requires a commitment (1) that the subatomic particles to which the theory is applied exist and (2) that they have the kinds of attributes (with the sorts of physical quantities as values) the theory postulates. Only commitment (2) is problematic, as all proponents of quantum theory accept (1). Physical system states for quantum theory have position and momentum coordinates as parameters, and unlike classical mechanics, the possible values of these are probability distribution functions (characteristic of the Borel subspaces which are projections of wave packets in the von Neumann model), specifying the probabilities with which various measured values can be obtained. To mention just two possibilities, these probability distribution functions can be interpreted as specifying statistical dispositional attributes of the subatomic particles or (following von Neumann) attributes of interacting systems of subatomic particles and measurement apparatus. In either case, one is committed to the functions designating attributes which can be possessed by systems of interacting particulars. Thus under quasi-realism, quantum theory receives a realistic interpretation.

Note also that quasi-realism does not require that, e.g., the ψ function designate any particular or attributes. Its function, like most of the apparatus of quantum theory, is to determine sequences of states—which requires that the ψ function be a theoretical term, but does not require that it be used existentially or designate a particular or an attribute of particulars.

32. The bracketed material beginning here constitutes a substantive addition to the original text.

33. This reflects the fact that it is the task of a theory to characterize all phenomena of a given sort which are causally possible (not just to characterize those phenomena of the sort which happen to occur) and the fact that all theories necessarily must abstract from their phenomena.

34. If the replica is an idealization, then, of course, the satisfaction of the idealization conventions is not met just by following laws or theories. For example, there is no empirically true law or theory which says that if a massive body were extensionless, then its location would be that of its center of gravity. Rather, this is a convention built into the idealization itself. This is not to say, however, that the convention is arbitrary. In the case of classical particle mechanics, it is a convention which must be imposed if rigid body mechanics is to be reducible to classical particle mechanics.

35. For further discussion of what is involved, see Suppes (1962, 1967), Suppe (1974b, secs. IV–E, and V–C) and ch. 4 of the present book.

36. By doing so we are excluding theories, such as quantum theory, which have statistical attributes from our discussion of models. Such theories also admit of mathematical models, though of a quite complicated sort (usually being configurated, infinite-dimensional spaces). Nothing in this chapter involves the details of these complexities, so they will be ignored here. For discussions of models for such theories, cf. Suppe 1967, sec. 2.3, and van Fraassen 1974a.

37. See note 25 above. [These systems of transformations determine symmetry relationships (of which invariances are a species). For an illuminating discussion of symmetry relationships, see van Fraassen 1989, part II.]

38. If, as in classical particle mechanics, varying numbers of defining parameters are required to completely specify the states of physical systems (e.g., in classical particle mechanics an n-body system requires $6n$ defining parameters), then this requirement and the discussion which follows is to be restricted to the class of physical system states having the same number of defining parameters as the dimensionality of the frame of reference. (See ch. IV, sec. 2, for a precise characterization of frames of reference.)

39. For a detailed discussion of how these models relate to the theory and how employment of theories with measurable parameters can be accomplished by working just with these models, see Suppe 1967, secs. 2.3–2.4. A fuller discussion of how these models are employed in experimental measurement is found in ch. 3 of the same work.

When these models are employed in lieu of the theory, one often formulates the theory by specifying the model and system of transformations rather than giving a linguistic formulation of the theory itself (see Suppes 1967 for a discussion of this practice).

Since he has been concerned exclusively with quantitative theories, van Fraassen (1970) identifies the theory with (one of) its phase space models.

While legitimate for his purposes there, the Semantic Conception of Theories must distinguish the theory from its phase space models if its account of theories is to apply to qualitative theories as well.

40. The possibility of such an observational/theoretical distinction being drawn has been seriously challenged, and most philosophers of science deny that it can be drawn. (See ch. 2 for an assessment of these challenges.) Despite the evidence against it, I will grant the distinction to the operational imperative for the moment.

41. There is no guarantee that these assumptions will be met by all theories or their formulations under the foregoing analysis. However, theories meeting this assumption provide the most favorable cases for the operational imperative, and so it will suffice to consider just them.

42. For simplicity of exposition I am presenting the argument in terms of knowledge that they are true; in point of fact, however, on most epistemologies, it would suffice for them to be inductively highly confirmed. Either way, the same sort of argument leads to the same conclusions.

43. In ch. 7, I examine in detail species definitions which also are thought to be operational definitions and argue for a similar conclusion: In natural taxonomies, species definitions are not definitional; rather they are counterfactually true or false statements of putative laws and theories.

44. Cf., for example, Dretske 1969 and Sellars 1956. Chapters 4 and 11 through 13 below further develop these ideas and suggest the outlines of such an approach to epistemology in science.