The Journal of Mind and Behavior
Summer, 1983, Volume 4, Number 3
Pages 339-352
ISSN 0271-0137

Functionalism and the Definition of Theoretical Terms

Austen Clark

University of Tulsa

Advocates of the philosophical thesis known as "functionalism" have recently proposed a solution to the problem of defining theoretical terms in behavioral science. They claim that such terms are *functionally defined*, or defined in terms of the *functional role* that the term specifies within the theory. This paper examines two versions of functional definition for theoretical terms. One version is shown to imply that some of the propositions in a theory are true a priori, solely in virtue of the meaning of the terms within them. Since no part of a theory can claim such independence from empirical test, that implication is shown to be unacceptable. An alternative account of functional definition is proposed. It allows one to define theoretical terms without stipulating that any of the propositions within the theory are true independently of evidence, or true in virtue of the meanings of their terms. This second account is shown to accord better with features of scientific practice.

The problem of defining theoretical terms in behavioral science has long proved elusive, yet recently proponents of the philosophical thesis known as "functionalism" have promised to provide a solution. Putnam, Fodor, and Harman claim that such terms are *functionally defined*, or that their meaning can be specified by describing a *functional role* within the theory (Fodor, 1968; Harman, 1973; Putnam, 1967, 1975a). Interestingly, all three of these proponents also claim that, at least within scientific theories, there is no clear distinction between analytic statements and others (Fodor, 1979; Harman, 1973; Putnam, 1962). An "analytic" proposition is one which is purportedly true solely in virtue of the meanings of its terms. An example is "All bachelors are unmarried." Such propositions are purportedly true a priori, independent of any possible empirical test. Typically, analytic truth is closely tied to definition, for a definition is simply a proposition setting out the meaning of a term, and is therefore a likely candidate for truth in virtue of meaning. For this reason it seems inconsistent to claim both that theoretical terms can be given a functional definition (or any definition) and that none of the propositions in a theory are true in virtue of their meaning.

I will show how certain versions of functionalism imply that some of the theoretical postulates employing a given term are analytically true. I will then

Request for reprints should be sent to Austen Clark, Ph.D., Department of Philosophy, University of Tulsa, 600 South College Avenue, Tulsa, Oklahoma 74104.

address the question of whether a functionalist position commits one to analytic truth. A method for defining theoretical terms is described in which one can claim both that theoretical terms are functionally defined and that none of the conditionals in a theory are true in virtue of their meaning.

Analytic Functionalism Applied to Theories

In order to understand what it means to give a term a "functional definition" or to specifiy a "functional role" we must first understand something about model building in psychology. Observations in that discipline describe either aspects of stimuli or characteristics of responses. The task of model building is to explain the relationships between stimuli and responses by proposing hypothetical intervening states and processes which are responsible for those regularities. The subject is considered a "black box," with stimulus "inputs" and response "outputs," and one attempts to elucidate mechanisms and processes within the black box which relate the two (see Clark, 1980, pp. 39-84).

More formally, a relationship between stimulus variables S and response variables R can be explained by proposing a set of internal hypothetical states (X, Y, Z, . . .) which are functionally related to one another and to stimuli and responses. For example, X and Y may be functions of different stimulus characteristics and in turn determine the values of variable Z, which leads to the given response. Each hypothetical state is linked to certain antecedent variables which determine its value, and to certain consequent variables whose value it influences. These functional relations can be described by theoretical conditionals with the following forms:

(1) If . . . S . . . then . . . X . . .
(2) If . . . X . . . then . . . Y . . .
(3) If . . . Y . . . then . . . R . . .

where "S," "X," "Y," and so on are stand-ins for predicates in the theory, and the elliptical ". . . X . . ." is to be filled in with a sentence ascribing the X predicate to subjects. Such a formatization allows one to derive a description of the response from a description of the stimulus conditions via premises describing hypothetical intervening states.

For example, suppose we wish to explain why some students consistently fail math tests which are well within their capabilities. We observe that presentation of mathematical expressions with multiply nested parentheses to students who have been repeatedly punished for passing math tests leads to failing the test. We attempt to explain the relations between these observations with the following theory, which I shall call the "mini-theory T" or just "T." Three theoretical terms are used: "math anxiety," "fear of success," and "low incentive motivation." T consists of four propositions:

(1) If a student is presented with nested parentheses, then the student feels math anxiety.
(2) If a student is repeatedly punished for passing, then the student has fear of success.
(3) If a student feels math anxiety and has fear of success, then the student has low incentive motivation.
(4) If a student has low incentive motivation, then the student fails tests.

This example is not intended to be realistic, and is deliberately oversimplified in order to make the logical form of its propositions manifest. Claims (1) and (2) link stimuli to theoretical terms, with the form "If . . . S . . . then . . . X . . . ." Claim (3) relates different theoretical terms to each other, and has the form "If . . . X . . . and . . . Y . . . then . . . Z . . . ." Claim (4) links a theoretical term to a response, and has the form "If . . . Z . . . then . . . R . . . ."

To functionally define a theoretical term is to define it in terms of its relations to other theoretic terms, stimulus variables, and response variables. Each such term marks a place within a pattern of relationships between hypothetical states, stimuli, and responses. The *function* of an internal variable is its job or role; that is, its typical effects on other states of the system. The function of a given state is described by specifying its relations to antecedents and consequences, so that its role in the workings of the whole is clear.

For example, in the mini-theory T, the theoretical term "low incentive motivation" is related to two theoretical terms "math anxiety" and "fear of success," and to the response characterization "failing the test." The functionalist idea is that the term "low incentive" is *defined* by those relations. The term purportedly picks out a state which plays a certain role within the system: the role of following from math anxiety and fear of success, and leading to test failures. By describing that functional role, one gives the meaning of the theoretical term "low incentive motivation."

These considerations lead to the following notion of functional definition: To give a theoretical term in psychology a functional definition is to claim that certain relations *must* obtain between the given term and terms describing other states of the system. Some of the conditionals relating "incentive motivation" to other theoretical terms and to observations are *true by definition*—for they describe relations which incentive motivation must bear to other states in the system in virtue of its being the functional state it is. They are part of the meaning of ascribing a functional state or of giving a functional definition of a state (see Block, 1980, pp. 173-181; Shoemaker, 1982). I shall call this view "analytic functionalism," as it implies that some subset of the propositions in a theory employing a given theoretical term are true in virtue of their meaning, and provide the definition for the term.

Arguments for Analytic Functionalism

Why should one think that a theoretical term is defined by a subset of the propositions within a theory? The strongest argument for this claim is that without such a definition, it seems impossible to test a theory or to compare different models. Because theoretical terms do not have their applicability determined by experimental tests and procedures, their use can be challenged by various skeptical arguments. The skeptic questions how any claim regarding a hypothetical state can be empirically substantiated (see Clark, 1983). Functional definitions of the above form could provide one answer to those questions.

One such line of questioning goes as follows. In order to set up a control group for an experiment, one needs to show that the manipulation of experimental and control groups gives different values for some independent variable X yet equally affects all other variables which might account for the behavioral phenomenon in question. To do this one must establish three claims. First, one must show that the stimulus conditions of the two groups differentially affect the independent variable X. For example, to experimentally manipulate math anxiety, one must be assured that altering the number of nested parentheses is sufficient to produce different levels of anxiety. Second, to show that it is only the difference in that internal variable which accounts for the behavior differences found, one must use the theory to show that the manipulations equally affect other internal variables. In a math anxiety experiment, one might randomize, so that history of punished successes is balanced in the two groups. Third, one must show that those controlled variables are the only other ones which could account for the behavior. To rule out such alternative explanations of R requires conditionals stating *necessary* conditions for the given behavior in terms of hypothetical internal states. For example, one would show that if the groups differ in failure rates, then either they differ in punished successes or they differ in math anxiety. Differential control hence requires conditionals of the forms:

(1) If . . . S . . . then . . . X . . .
(2) If . . . X . . . then . . . R . . .
(3) If . . . S . . . then if . . . R . . . then . . . X . . .

where "S" is a stimulus variable, "X" is the theoretical term, and "R" is a response variable.

Suppose we wish to test a proposition which provides a sufficient condition for the occurrence of a given hypothetical state:

(A) If a student is presented with nested parentheses, then the student feels math anxiety.

How could we establish that this conditional is true—that nested parentheses suffice to produce math anxiety? One must have sufficient grounds for justifiably ascribing math anxiety to the subject, and for that one needs evidence which does not rely on (A) as an essential premise and which establishes that the student feels math anxiety. For example, one might employ a proposition linking math anxiety to a different response, such as:

(B) If a student has been repeatedly punished for passing, then if the student does not fail the test, then the student does not feel math anxiety.

Claims (A) and (B) together allow a test of the theory, for students satisfying (A) should not satisfy (B), and conversely. However, the same problem of justifiable ascription of a theoretical term reappears for the second conditional, namely: how does one know that (B) allows one to rule out math anxiety in some subjects? Perhaps math anxiety is present in the repeatedly punished passers satisfying (B), but its relation to low incentive motivation is different. It is apparent that such a question can be raised of any proposition linking a theoretical term to an observation.

In order to stop this regress and establish sufficient grounds for ascribing a theoretic term, it seems that at some point one must refuse the request for additional data and say instead:

> We know that such-and-such antecedent conditions provide sufficient grounds for ascribing "math anxiety" because that is what the term means. It is part of the definition of "math anxiety" that math anxiety comes about when such-and-such conditions obtain.

Some of the conditionals relating "math anxiety" to other theoretical terms and to observables provide the definition for the term. Because such definitions would be exempt from empirical test, they would establish unimpeachable grounds for ascribing "math anxiety" to subjects—grounds which could not be challenged by the skeptic.

It seems impossible to establish any empirical claims concerning the relations between a theoretical state and its antecedents and consequences without beforehand defining the theoretical term. Since such a definition can only be a conditional relating the term to antecedents or consequences, it follows that not all of the conditionals relating a given theoretical term to other theoretical terms and to observables are empirical propositions. Some of them must have the special role of providing the theoretical term with a meaning. For example, some subset of the conditionals using the term "math anxiety" must be true a priori, and provide the term with its definition. Otherwise the other conditionals employing the term have no empirical content, and their truth cannot be established. Hence if any observable consequences follow from postulating a hypothetical state, it seems that some theoretical conditionals using that hypothetical term are not open to doubt. They define the term, so that its empirical consequences can be tested.

These considerations lead one to the view that some subset of the propositions in a theory are true a priori, in virtue of the meaning of the theoretic terms they contain; and that such a subset provides the definition for the theoretic term. A "functional definition" of a term, on this view, picks out those relations which obtain in virtue of the meaning of the term between a given hypothetical term and its antecedents and consequences.

Problems with Analyticity in Theories

There are many well-known objections to the idea that some propositions in ordinary language are true in virtue of their meaning (Putnam, 1962; Quine, 1951). These problems are even more marked when one attempts to find a subset of analytic truths within a theory. The fatal problem with this idea is that it is difficult if not impossible to find any such distinguished subset in any scientific theory. Instead, every proposition in a theory is open to test. None of them are immune from revision, as an analytic definition would be. None have the requisite independence from data and none can be assured of truth no matter what the empirical results. Hence none of the propositions in a theory could be analytic definitions.

The difficulties with analytic definition can be vividly illustrated using an analogy of Hempel:

> A scientific theory might therefore be likened to a complex spatial network: Its terms are represented by the knots, while the threads connecting the latter correspond, in part, to the definitions and, in part, to the fundamental and derivative hypotheses included in the theory. The whole system floats, as it were, above the plane of observation and is anchored to it by rules of interpretation. These might be viewed as strings which are not part of the network but link certain parts of the latter with specific places in the plane of observation. (Hempel, 1952, p. 36)

Suppose that each theoretical term is linked to some of its antecedents and to some of its consequences by propositions true in virtue of meaning. Since such links are true independently of empirical test, they should be distinguished from the rest, perhaps by marking them with threads of a different color. The colored threads are absolutely rigid, as they provide the definitions needed to test the rest of the linkages.

This view has an unacceptable implication. Every theoretic term is thought to have analytic relations with some of its neighbors. Every node (term) in the network is then connected to some neighbors by colored threads, and therefore from any point in the network one can reach the observation plane on a pathway consisting entirely of colored threads. But since each of the analytic links is true a priori, it follows that the relation between the endpoints of the pathway is also true a priori. Therefore, the relationship found between distinct observations is true a priori, so long as they can be connected by a chain of definitions. We get the absurd consequence that some relationships between observations do not need to be empirically established, but are necessary truths.

A classical requirement for definitions in a formal system is that they are not *creative*. Definitions must not allow one to derive relationships between existing terms which one could not derive without the definition (Suppes, 1957, p. 154). The definitions of the analytic functionalist are creative, however. In effect those definitions constitute a connected sub-network within the network of the theory. Every theoretical term is connected to observations via a path of analytic relations among terms. Not only would such definitions allow one to derive new relationships between observations, but those relationships would be established as true independently of any empirical test. Because such a claim is an implication of analytic functionalism, that account of theoretic definition must be rejected.

Why Define?

The question remains: In what sense can theoretical terms in psychology be given a functional definition if one denies that any of the conditionals in the theory are true by definition? In particular, how can a functionalist account of theoretical terms be reconciled with a repudiation of the idea of analytic truth within a theory?

To approach this problem we should first examine the role which analytic definitions were thought to have. It will be recalled that experimental differentiation of models requires conditionals of the forms:

(1) If . . . S . . . then . . . X . . .
(2) If . . . X . . . then . . . R . . .
(3) If . . . S . . . then if . . . R . . . then . . . X . . .

In order to justifiably ascribe a theoretical term X, one must have grounds for the ascription of X provided by justified belief in the truth of conditionals of the above form. It seems that one could never justify belief in any conditionals of the above form unless at least some such conditionals were analytically true. Otherwise, to substantiate the ascription of X, one must rely on yet other conditionals of the above forms, thereby raising again the problem of justifiable ascription of X. Only if some conditionals of the above forms were true in virtue of their meanings, and independent of empirical challenges, can this skeptical regress be stopped.

One step in this argument is incorrect, however. In order to justifiably ascribe X, one need not have conditionals of the above form which are analytically true; their truth will suffice. For example, instead of requiring that "If . . . S . . . then . . . X . . ." be analytically true, it would suffice to establish that as a matter of fact, if S occurs, X does as well.

The appeal to analytic truth occurred precisely because it was unclear how such a claim could be empirically established. Each empirical test of a theoretical proposition employs some part of the theory as a premise, and the project of empirically testing *every* characterization of a hypothetical state seems infested with a fatal circularity. Analytic truth provided a starting point for investigation, so that none of the empirical premises of the theory were needed as assumptions in testing theoretical propositions. One could be sure of at least some of the conditionals, because they were true by definition.

A different sort of justification for such conditionals is possible, however (see Clark, 1983). Instead of attempting to ground the ascription of a theoretical term on premises which are independent of other empirical premises of the theory, one can justify acceptance of a given conditional on the grounds that it is part of a theory which is well confirmed. A conditional of the form "If . . . S . . . then . . . X . . ." does not require justification which is independent of the other empirical premises of the theory. Instead we may be justified in believing that X occurs, because the predictions made by supposing that it does occur are verified. The success of the theory as a whole provides the justification for employing its conditionals as premises in derivations. Those conditionals do not require—and cannot be given—a justification which does not use other parts of the theory as premises.

Functional Definition without Analyticity

Theoretical terms can be given a definition—although it does not consist of analytically true statements. A theory describes an abstract pattern of relationships between terms which is purportedly satisfied by unobserved states and processes. Obviously, if the theory were true, then there would exist a collection of states which stand in those relations to one another. I suggest that to claim hypothetical terms are functionally defined is to claim that the converse holds. That is, one accepts a functional definition of one's theoretical terms if and only if one accepts the claim that *if* there exist states which stand in the specified relations to one another, then those states *are* the hypothetical states in question. To say that one's theoretical terms are functionally defined is just to say that satisfaction of the proposed relations between the terms is *sufficient* to identify a set of states as being the proposed hypothetical states.

For example, low incentive motivation is characterized by the mini-theory T as bearing certain relations to states of math anxiety, fear of success, and failing tests. Obviously, a state which does not satisfy the appropriate pattern of relations to those states is not a state of low incentive motivation. Those relations therefore provide a necessary condition for identity of low incentive motivation. I suggest that one accepts a *functional definition* for the term if one accepts that those relational characterizations also provide a *sufficient* condition for the identity of low incentive motivation. That is, the theoretical term is defined by a conditional of the form:

If any state bears such-and-such relations to math anxiety, fear of success, and test failures, then that state is low incentive motivation.

A functionally defined term has identity criteria provided by its relations to other terms in the theory.

This claim can be clarified with the notion of a Ramsey sentence (see Carnap, 1966, pp. 246-274; Lewis, 1970, 1972). A Ramsey sentence of a theory is a sentence which implies all of the same observational predictions as the theory itself, but which employs no theoretical terms. It is formed as follows. First, conjoin all of the propositions in the theory. For the mini-theory T, the result of this can be abbreviated as follows:

If . . . parentheses . . . then . . . math anxiety . . ., *and* if . . . punishment . . . then . . . fear of success . . ., *and* if . . . math anxiety . . . and . . . fear of success . . . then . . . low incentive . . ., *and* if . . . low incentive . . . then . . . failure . . . .

Second, replace each theoretical term in the theory with a variable. Replacing "math anxiety," "fear of success," and "low incentive motivation" with U, V, and W respectively, we have:

If . . . parentheses . . . then . . . U . . ., *and* if . . . punishment . . . then . . . V . . ., *and* if . . . U . . . and . . . V . . . then . . . W . . ., *and* if . . . W . . . then . . . failure . . . .

The intent of this replacement is to avoid all connotations associated with the theoretical terms by treating them as variables. They are predicate variables, ranging over states defined by their place in the schema. To derive any observational predictions from the above schema, one must convert it to a sentence by prefixing it with existential quantifiers for the predicate variables:

There exist states U, V, W *such that* if . . . parentheses . . . then . . . U . . ., *and* if . . . punishment . . . then . . . V . . ., *and* if . . . U . . . and . . . V . . . then . . . W . . ., *and* if . . . W . . . then . . . failure . . . .

This last construction is the Ramsey sentence for T. I will call it "R(T)." Given any theory T, one can form the Ramsey sentence R(T) by following the above steps. Any prediction concerning observations which can be derived from T can be derived as well from R(T).

A theory entails its Ramsey sentence, for "If T then R(T)" is necessarily true. The converse ("If R(T) then T") does not necessarily hold. To give a functional definition of one's theoretical terms is, I suggest, just to assert that the converse does hold.

For example, the Ramsey sentence above is a consequence of the theory T, for if there are states of math anxiety, fear of success, and low incentive motivation satisfying the relations specified in the theory, then those states also satisfy the predicate variables U, V, and W; and therefore R(T) is satisfied as well. The converse does not logically follow, for it is possible that there exist states U, V, and W which satisfy R(T), but which are not math anxiety, fear of success, and low incentive motivation. Unless the relations in R(T) between U, V, and W are *sufficient* to identify them as math anxiety, fear of success, and low incentive motivation, R(T) may be true while T is false. There may in some theories be significant characterizations of the hypothetical states other than the relations proposed between them. But a functional definition of the theoretical terms claims there are no such other characteristics, and that the relations between terms and observations are sufficient to identify the hypothetical states. The definition for the theoretical terms "math anxiety," "fear of success," and "low incentive motivation" is then given in one complicated sentence:

> IF *there exist states* U, V, W *such that* if . . . parentheses . . . then . . . U . . ., *and* if . . . punishment . . . then . . . V . . ., *and* if . . . U . . . and . . . V . . . then . . . W . . ., *and* if . . . W . . . then . . . failure . . . .
>
> THEN if . . . parentheses . . . then . . . math anxiety . . ., *and* if . . . punishment . . . then . . . fear of success . . ., *and* if . . . math anxiety . . . and . . . fear of success . . . then . . . low incentive . . ., *and* if . . . low incentive . . . then . . . failure . . . .

Schematically:

If (R)T then T.

This asserts that if there are states U, V, and W satisfying the required pattern of relations, then those states are math anxiety, fear of success, and low incentive motivation.

How does such a claim *define* the theoretical terms? It defines them by specifying the sufficient conditions for identifying a state as a given theoretical state. To say that "math anxiety" is functionally defined is just to say that any state satisfying such-and-such relational characterizations *is* math anxiety. We do not need to claim that the definition ("If R(T) then T") is analytically true for it to give a definition to the theoretical terms. (It is difficult to see how this claim could be true in virtue of meaning, since R(T) contains colorless variables in place of terms.) Instead the schema "If R(T) then T" is a description of how the terms in some theories are used. A functionalist account is true if the conditional is true; the conditional need not be analytically true.

Note that the Ramsey sentence conjoins all of the propositions of the theory, and therefore in effect defines all of the theoretical terms collectively. One cannot partition the theory into a set of sentences which are relevant to the definition of "math anxiety" and a set which are not. All of the conditionals of the theory must be mentioned. To give a functional definition of "math anxiety," for example, is to assert something of the form:

> If there exists a state U that bears such-and-such relations to fear of success, low incentive motivation, presentation of nested parentheses, punishment for passing, and failing tests, then that state is math anxiety.

The problem with this example is that "fear of success" and "low incentive motivation" must themselves be defined in order for the above to successfully define "math anxiety." To define either of those theoretical terms, however, one must mention the rest of the theory. This is why a functional definition of theoretical terms is accomplished with a single conditional which embraces the entire theory at once.

Implications of Definitions Without Analyticity

Some implications of this analysis of functional definition should be sketched. First, definition via a Ramsey sentence accords with the intuition that in order to know the meaning of a theoretical term, one must understand the entire theoretical structure in which it plays a part. The meaning of any one theoretical term is bound up with that of all the other theoretical terms with which it is associated. A network of propositions contributes to the meaning of each theoretical term, yet none of the propositions need to be accepted as true-by-definition or as analytically true. Any one of them can be tested. In effect therefore a collection of conditionals is required in order to specify the meaning of each theoretical term X, yet none of its members are analytically true. The definition schema "If R(T) then T" accords with these facts.

The above analysis of functional terms does not require any partitioning of theoretical conditionals into those true in virtue of meaning and those facing empirical tests. Instead theoretical terms are defined by a collection of conditionals whose observational consequences are confirmed. The discourse combines empirical and stipulative uses (Hempel, 1965, p. 207). Such an analysis can easily deal with seeming paradoxes such as the possibility of performing empirical tests on all of the conditionals making up a definition of a term, or the possibility of falsifying a definition.

For example, change in a theoretical definition can be analyzed as follows. Theories constantly evolve during inquiry, and the content of the conditionals employing a given theoretical term is constantly changing. This change does not implicate the theoretician in contradictory or meaningless statements (Hempel, 1965, p. 205; Putnam, 1962). New indices may be introduced for a given term, and new connections discovered between a given hypothetical state and other states of the system. The theoretician may abandon those conditionals which originally were used to introduce the term, and introduce new ones, while still claiming to refer to the same hypothetical state. It is very difficult to make sense of these facts if one holds to the view that some fixed set of conditionals employing a theoretical term are analytically true, and constitute its definition.

However, the analysis above can account for changes in the definition of theoretical terms. Amendments to a theory can be minor or major. With minor amendments, one often continues to think that the new version concerns the same theoretical entities as past versions—only we now know more about them, and have corrected past errors. After major shifts in the theory content, however, we may say that we have a new theory, and that the theoretical conditionals no longer describe the same hypothetical states, or have the same meanings. In effect the distinction between modified knowledge about theoretical terms with the 'same' definitions and new definitions for the terms is just the distinction between minor and major amendments to theories.

On a strict Ramsey sentence approach, adding or deleting conditionals for a theory T to give a modified T′ implies that the definitional schema "If R(T) then T" does not define terms in T′, since T and T′ postulate different relations among the terms. Strictly speaking, changing a single proposition of the theory alters the meaning of all the theoretical terms within it. However, the ordinary use of "definition" and "change in meaning" is not so strict. Instead I suggest that theoretical terms are provided "new definitions" or given "new meanings" if and only if the theory employing them is "new," and not merely a slightly modified version of the same theory. Criteria for changing the meaning of theoretical terms are no clearer than criteria for changing the theory. When do revisions become extensive enough that we say we have a 'new' theory?

I suggest that certain conditionals involving a theoretical term are relatively central to the conception of the term, and others are relatively peripheral. Trans-theory identity of hypothetical entities is lost when enough of the central characteristics are abandoned or changed. Of course, the content of the class of core characteristics can change over time; and furthermore, there is no firm rule for deciding when enough of the conditionals have changed for a new entity to be declared. What is often meant by a 'definition' for a theoretical term is a set of propositions which suffice to establish an initial understanding of the term. Knowledge of the key conditionals employing a term is necessary if one is to know how to use it. When a theory is first proposed or first learned such core conditionals may be advanced as a sort of definition (Putnam, 1970, 1975b). Here "definition" really means nothing more than instruction in how to use the term. A theoretical term has changed its definition when enough of those core conditionals have changed. When that happens, new instructions on how to use the term are required. The definition of the term changes through modification of the class of conditionals employing it; no analytic propositions are involved.

The above analysis does not entail that all theoretical terms can be functionally defined, although it is interesting to note that many of the arguments for functional definitions in psychology are quite general and would seemingly apply to theoretical terms in other disciplines as well. However, a functional definition of terms can only be provided for theories in which the relational characterizations of theoretical states or processes suffice to identify their referents. If this is not the case—if for example some theoretical state is described with a one-place predicate, or has intrinsic features not characterized by its relations to other states—then there is more to the theory than a structure of relations between hypothetical states and observables, and a functional definition of terms will not succeed. Functionalists have not discussed whether theoretical terms in all disciplines are to be functionally defined, and if not, what distinguishes theoretical terms in psychology from other theoretical terms. But that is a problem which must elaborated elsewhere.

References

Block, N. Introduction: What is functionalism? In N. Block (Ed.), *Readings in the philosophy of psychology* (Vol. 1). Cambridge, Massachusetts: Harvard University Press, 1980.

Carnap, R. *An introduction to the philosophy of science.* (M. Gardner, Ed.). New York: Basic Books, 1966.

Clark, A. *Psychological models and neural mechanisms.* Oxford: Oxford University Press, 1980.

Clark, A. Hypothetical constructs, circular reasoning, and criteria. *The Journal of Mind and Behavior*, 1983, 4(1), 1-12.

Fodor, J.A. *Psychological explanation.* New York: Random House, 1968.

Fodor, J.A. *The language of thought.* Cambridge, Massachusetts: Harvard University Press, 1979.

Harman, G. *Thought.* Princeton: Princeton University Press, 1973.

Hempel, C.G. Fundamentals of concept formation in empirical science. In O. Neurath, R. Carnap, and C. Morris (Eds.), *International Encyclopedia of Unified Science* (Vol. 2, No. 7). Chicago: University of Chicago Press, 1952.

Hempel, C.G. The theoretician's dilemma: A study in the logic of theory construction. In C.G. Hempel, *Aspects of scientific explanation.* New York: The Free Press, 1965

Lewis, D. How to define theoretical terms. *Journal of Philosophy*, 1970, 67(13), 427-444.

Lewis, D. Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, 1972, 50(3), 249-258.

Putnam, H. It ain't necessarily so. *Journal of Philosophy*, 1962, 59(22), 658-671.

Putnam, H. Psychological predicates. In W.H. Capitan and D.D. Merrill (Eds.), *Art, mind, and religion.* Pittsburgh: University of Pittsburgh Press, 1967.

Putnam, H. Is semantics possible? In H.E. Kiefer and M.K. Munitz (Eds.), *Language, belief, and metaphysics.* Albany: State University of New York Press, 1970.

Putnam, H. Philosophy and our mental life. In H. Putnam, *Mind, language, and reality: Philosophical papers*, (Vol. 2). Cambridge, England: Cambridge University Press, 1975. (a)

Putnam, H. The meaning of 'meaning.' In K. Gunderson (Ed.), *Language, mind, and knowledge. Minnesota studies in the philosophy of science* (Vol. 7). Minneapolis: University of Minnesota Press, 1975. (b)

Quine, W.V. Two dogmas of empiricism. *Philosophical Review*, 1951, 60, 20-43.

Shoemaker, S. Some varieties of functionalism. In J.I. Biro and R.W. Shahan (Eds.), *Mind, brain, and function.* Norman, Oklahoma: University of Oklahoma Press, 1982.

Suppes, P. *Introduction to logic.* New York: Van Nostrand Reinhold Company, 1957.