The Journal of Mind and Behavior
Winter, 1983, Volume 4, Number 1
ISSN 0271-0137

Hypothetical Constructs, Circular Reasoning, and Criteria

Austen Clark

University of Tulsa

Two accounts of construct validation are critically analyzed. One is the claim that such validation inherently involves circular reasoning; the second is that construct terms are meaningless unless they are provided with an observational criterion, so that construct validation is nothing more than validation against a criterion. Both views are shown to rest on the assumption that each claim concerning a construct must receive empirical support which is independent of the rest of the theory, employing no other theoretical proposition as a premise. Once this unrealistic assumption is abandoned, a clearer analysis of construct validation emerges. Some of the implications of this account concerning the definition of constructs and the use of convergent indicators are sketched.

A "construct" or "theoretical term" is a term in a theory which ascribes a property or state to subjects but which is not an observation term (see the APA "Standards," 1967, p. 176). An "observation" term can be taken as a term for which there is high intersubjective agreement on its application, given the same concurrent stimulation across observers (Nagel, 1961, pp. 79-90; Quine, 1969, pp. 85-89). Such terms can be affirmed or denied without reliance on a particular theory. Theories clearly contain terms which ascribe states, properties, or relations but which do not meet this criterion. Such terms are "theoretical terms" or "constructs." They are not defined by explicit test procedures, but can be related to observations only through inferences linking them to other theoretical terms and to observations. They describe "hypothetical" states or processes which cannot be directly observed, but are postulated in order to account for observations (see Hempel, 1965, pp. 173-226; Nagel, 1961, pp. 91-105).

There are a number of confusions as to how theoretical terms are used in psychology, and some authorities even recommend that all such terms be avoided, so that only observation terms are employed (Skinner, 1974). The same issues recur in discussions of construct validation of tests, where the intent is to investigate which constructs are measured by a given test. In this paper I will examine some of these positions, and point out their weaknesses. I will then propose an analysis of theoretical constructs which is free of these problems, and in the last section show how that account can clarify some of the methodological issues associated with the use of hypothetical constructs.

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

The Problem of Ascribing Constructs

Suppose we have some test or procedure T which we think measures the presence or magnitude in subjects (human or otherwise) of some theoretical construct X. X may be a construct such as "fear," "expectancy," or "conditioned frustration." How do we establish whether or not T measures X?

If T measures X, then knowing the value of T for a subject should enable us to specify the magnitude of X in the subject. That is, T must correlate with X. But since X is not directly observable—as otherwise we would not need to establish the construct validity of T—we cannot establish such a correlation directly, for we have no way of directly assessing variation in X. We need some second measure T' which is a valid measure of X, so that the construct validity of T could be established by correlating it with T'.

How then do we establish that T' measures X? This problem is exactly the one we started with. In order to establish the construct validity of one measure, it seems that we must already have a measure with known validity. Inferences from observations to theoretical states can seemingly never be made. At best T correlates with some other observable test T', but neither T nor T' allow one to infer anything about X states.

For example, suppose that X is the construct of "fear" in a two-process learning theory, hereafter abbreviated as "F" states (see Gray, 1971, 1975; Miller, 1957; Mowrer, 1960). One measurement for F states might be an increase in rates of Sidman avoidance when a conditioned stimulus for punishment is presented. How do we know that increased Sidman avoidance indicates fear? One answer might be that an increased rate of Sidman avoidance happens to correlate with a potentiated startle reflex, and we accept a potentiated startle reflex as a valid measure of fear. But the question obviously arises: how do we establish that the potentiated startle reflex measures the construct F? Empirically, all that we have done is correlate two experimental measures, and this covariation does not allow us to infer the presence of an unobservable construct.

The fear construct is linked via conditional (if . . . then . . .) statements to stimulus conditions and to responses in particular situations. For example, we might have:

1. If S is presented with a negative reinforcer, then S will be in an F state.
2. If S is presented with a conditioned stimulus for a negative reinforcer, then S will be in an F state.
3. If S is in an F state, then the startle reflex of S will be potentiated.
4. If S is engaged in a Sidman avoidance task, elicitation of an F state will increase the rate of avoidance.

To establish the truth of any such conditional, it seems one must have some independent means of establishing the presence or magnitude of the fear

construct. For example, to show that punishing stimuli induce F states, one needs some means of establishing the presence of F states other than the presentation of punishing stimuli. But the presence of F states can only be established by some new conditional *itself* open to this challenge. One can always ask: "How do you know that T indicates F?" Producing some further conditional linking F to an experimental procedure or observation does not answer this question.

The First Mistake: Construct Validation as Circular

One response to this problem (Campbell, 1960; Campbell and Fiske, 1959; Cronbach, 1970; Cronbach and Meehl, 1955; Nunnally, 1967; Nunnally and Durham, 1975) is to admit that no single correlation will suffice to establish construct validity, but to insist that confirmation of a predicted *pattern* of results among tests and procedures will do so. For example, a theoretical claim concerning an antecedent or consequence of "fear" does not by itself entail observations, but will only yield experimental predictions when combined with other claims concerning unconditioned responses to punishment, expectancies, incentives, arousal, conditioning of responses, and so on. The construct does not confront the evidence directly, but only through a network of inferences provided by the theory.

A theory will predict a certain pattern of results—a pattern of relations among various constructs—and on this view, to show that T measures X, we need to show that the predicted pattern obtains. For example, a two-process theory will predict certain relations between constructs of expectancy, fear, safety signals, and so on (call these constructs X, Y, Z). We have purported measures A, B, and C for those constructs. The construct validity of A measuring X is established by showing that the pattern of relations expected between X, Y, and Z obtains for A, B, and C. On this view, a validity claim for an instrument is just one theoretical proposition among others, and its confirmation rests on the success of the theory as a whole.

This analysis of construct validation has been challenged on the grounds that it is circular (Bechtoldt, 1959, pp. 624-625; Nunnally, 1967, pp. 92-94; Nunnally and Durham, 1975, pp. 301-305). Nunnally argues that confirmation of the expected relations between A, B, and C demonstrates that they are valid measures of X, Y, and Z only if one assumes that the three constructs in fact stand in the relations postulated by the theory. The data are equally consistent with the hypothesis that X, Y, and Z are not related as the theory claims, so that even though the expected relations are found between A, B, and C, they do not measure the three constructs.

For example, why should one think that conditioned suppression is an indicator of fear? One reason to accept this hypothesis is that it leads one to

expect a number of correlations between behavioral phenomena such as increased Sidman avoidance, potentiated startle response, and others. Furthermore, those correlations are confirmed. Nunnally protests that such a procedure is circular. For unless one assumes that the construct 'fear' in fact is related as the theory claims to those other constructs and behavior, a pattern of correlations among behavioral phenomena gives one no reason to suppose that a particular index is linked to a particular construct. One might be wrong about the experimental indices for a given construct, or one might be wrong about the relations between the constructs. It seems that claims concerning experimental indices for "fear" must be assessed along with claims about the relations of "fear" to other constructs; validation of the instruments cannot be achieved without confirmation of the theory about the relations of states 'measured' by the instruments. Given the paradox of such a procedure, Nunnally concludes that "it logically is not possible to prove that any set of observables measure a construct . . . . In a strict sense, 'construct validity' is logically impossible . . ." (Nunnally, 1967, p. 98; Nunnally and Durham, 1975, p. 308). He urges that construct discourse or hypothetical terms do not refer to anything other than clusters of observables, and that to think they have real referents is incorrect. Nunnally proposes that construct validation is nothing more than investigation of the internal structure of sets of observables and of relationships across different sets of observables. Other authorities recommend that construct terms be dropped from the psychological vocabulary and that only observation terms be employed (Bechtoldt, 1959, p. 628).

The Second Mistake: Constructs Require Criteria

One response to the seeming circularity of construct validation is to claim that insofar as construct validation is possible, there must be some means of definitively establishing the presence or magnitude of the construct. One denies the claim that all of the conditionals relating a given construct to observables are empirical propositions for which evidence can be requested. Instead, some of those conditionals are thought to provide a *criterion* for the construct: they are *true by definition*, and the observation is linked to the construct in virtue of the meaning of the construct term (Bechtoldt, 1959, pp. 621-622; Clark, 1980, pp. 67-71). Because there can be high intersubjective agreement on the ascription of the given observation term, it thereby provides a criterion for definitively establishing the presence or magnitude of the construct.

For example, if "fear" cannot be linked to any specific observations in specific situations, one can argue that it is a meaningless term. Adding it to the theory does not improve one's ability to predict or explain phenomena, and

the principle of parsimony would rule out such an addition. Therefore, the argument goes, there must be some indices for "fear" provided by the meaning of the term. One might argue that part of the meaning of the construct "fear," as used in a particular theory, is that it is a consequence of a punishment stimulus or a negative reinforcer. Such a claim is not empirically testable, but is rather a definition. Punishment is in that case a *criterion* for the construct, as it provides an observational test for ascription of the term.

Once one has such a criterion, construct validation can proceed. To show that T measures X, we correlate T with the criterion of X. The criterion affords us certainty that X obtains, and allows us to determine whether T measures the construct. It is for this reason that some writers claim that the only legitimate validation of a construct is criterion validation, established by a high correlation with the variable of interest (Anastasi, 1950, p. 67; Bechtoldt, 1959, pp. 621-622; Brodbeck, 1963, pp. 59-72; Horst, 1966, pp. 345-347), or that a construct term can be defined by a set of observational terms (Anastasi, 1968, p. 122; Lord, 1955, p. 509; Nunnally, 1967, pp. 97-98; Nunnally and Durham, 1975, pp. 307-309). On this view, the only usefulness of construct talk is that it can serve to summarize complicated sets of observations and potentially intricate links to other observations. The definitional chains may be complicated, but nevertheless contain only observables.

One problem with the criterion view is that if a construct can be provided with a criterion, then (paradoxically), there seems to be no reason to use the construct, as its observational criterion will do as well. Suppose, for example, that by using punishment as a criterion for F states, we find that F states lead to increased avoidance tendencies in approach-avoidance conflicts. In effect, to link F states both to punishment and to increased avoidance tendencies is to claim that those two observations correlate. But in that case, why mention F states at all? Why not directly relate punishment to increased avoidance tendencies? If punishment provides a criterion for F states, then talk about F states seems to be just a useless form of words, replaceable by talk about punishment.

There are other problems with the criterion view, which can be mentioned but not elaborated here (see Hempel, 1965, pp. 109-118; Putnam, 1962, 1966, 1977; Quine, 1951, 1969). One is that examination of a theory provides no clear means of distinguishing those propositions within it which are true by definition from those asserting empirical claims. For example, in all the claims concerning antecedents or consequences of F states, there is no clear candidate for a criterion for the term. None of the conditionals in a theory are treated as being necessarily true or as immune from revision. Each of them is subject to test, and any one of them can be modified or abandoned. None are advanced as being true by definition.

An Alternative View of Theory Construction

Nunnally's circularity argument and the criterion view of constructs both make the same mistake about the epistemology of theory construction. Both assume that each proposition in a theory must receive empirical support which does not employ other parts of the theory as a premise. Both assume that a theory is built up by accumulating propositions each with empirical support independent of the rest of the theory.

The 'circular' procedure attacked by Nunnally is nothing more than use of one part of a theory as a premise in tests of another part. It seems illegitimate to do this because one could never obtain sufficient support for a given premise in a theory if such support required one to assume the truth of other premises in the theory.

The claim that one needs observational criteria for a construct term also makes this assumption. Such stipulative definitions were used to justify ascription of a construct in cases where the only other justification possible would employ some other conditional in the theory. An observational criterion allows one to justify ascription of the construct without appealing to the rest of the theory.

Both of these positions on construct validity seem reasonable if one accepts the assumption that each conditional in a theory must receive some empirical support which does not employ other theoretical conditionals as essential premises. For if such independent support is required, there seems to be no way to empirically establish that a given unobservable state or process has occurred. Every conditional linking a construct to observations can be challenged for its support. So one is led either to deny that construct talk means anything other than talk about observables, or to claim an a priori justification for some such conditionals, based on the meaning of the hypothetical terms.

There is, however, another way in which propositions involving hypothetical terms can receive empirical support. Instead of building a theory through accretion of independently supported propositions, one can think of it as an hypothesis concerning a network of inter-related hypothetical processes. Treat the entire theory as a premise, and see what sort of results it predicts. The justification for accepting any particular conditional within it is that the theory as a whole is well confirmed (Popper, 1962, 1968). A number of conditionals are tested when one tests a prediction, since no single theoretical conditional suffices to entail an experimental prediction. If the prediction is not confirmed, there are several different premises in its derivation which one could change to reconcile the theory to experimental findings; just as there are several conditionals at risk in a given experiment, confirmation of predictions tends to confirm a collection of conditionals rather than one alone (Cronbach, 1970; Cronbach and Meehl, 1955; Loevinger, 1957).

Not only is there nothing circular about using one part of the theory as a premise in experimental tests of another part, but such a procedure is unavoidable. At one point we may use a given conditional—such as "If S is presented with a negative reinforcer then S is F"—as a premise in deriving an experimental test of some other part of the theory—such as the link between F states and conditioned suppression. We would use the above premise as a justification for the claim that S is in an F state in the given experiment. We would justify use of that premise by appeal to the success of the theory to which it belongs. We do not need to justify it by claiming it is an a priori definition of "fear." Later we can proceed to test the conditional "If S is presented with a negative reinforcer then S is F" by a different experiment using other conditionals involving F states as premises. We may even revise the conditional linking punishment and fear, stating some necessary condition or parameter for that relationship.

This procedure may seem to employ circular reasoning, where one argues for P on the grounds of Q, and then argues for Q on the grounds of P. Here one seems to argue for construct validity by assuming a theory of their interrelations is true, then argues for the theory by assuming one has valid measures of its constructs. However, what here 'breaks the circle' and justifies the procedure is that by taking part of the theory as a premise, one can confront it with new empirical tests. Different predictions are derived from different propositions of the theory, and repeated experimental confirmations thereby give confidence in the entire structure. One may take P for granted while testing Q, but the justification is that P and Q together imply experimental predictions which are repeatedly confirmed.

On this account, ascription of a particular theoretical construct is justified to the same extent that the theory employing it is justified. To claim that F states exist is just to say that the theory about F states is true. We would know that F states occur only if we *knew* the theory were true. Since certainty can never wholly be obtained for a theory, ascription of theoretical states and processes can never be absolutely proven. For the same reason, one can never achieve certainty on the properties or relations of constructs. We know the antecedents and effects of a given construct only to the extent that we know a theory to be true.

Some Implications for the Definition of Theoretical Terms

The only 'definition' one can give for a theoretical term is one which implicitly involves all of the conditionals in which it occurs. A theory proposes a structure of relations between hypothetical states, stimuli, and responses, and a given theoretical state is characterized by the relations it bears to other hypothetical states and to observables. For example, "fear" is charac-

terized by its relations to punishment, conditioned stimuli for punishment, expectancies, incentives, arousal, and to particular tasks such as Sidman avoidance or approach-avoidance conflicts. The only way to define "fear" is to specify all these relations; that is, an F state is just that sort of state which bears such-and-such relations to punishment, conditioned stimuli for punishment, expectancies, incentives, and so on. More formally, a definition would take into account the conjectural status of propositions about F states, and read: if there exists a state that bears such-and-such relations to other hypothetical states, stimuli, and responses, then that state is an F state (see Carnap, 1966, pp. 265-274).

Such a 'definition' is peculiar because it does not involve any claim to a priori truth. It defines a hypothetical term solely in terms of its relations to other hypothetical terms and to stimuli and responses. To fully specify what sort of state is an F state, one must include every conditional relating F states to other states in the system, and so in a sense every conditional employing the fear construct plays a role in the definition of the term. Since those conditionals must mention other theoretical terms in the theory, this carries the further implication that the theoretical terms are defined collectively, in terms of their relations to one another and to stimuli and responses.

Some of the methodological implications of this account are as follows.

1. It is apparent that there is no need to provide a special 'definitions' section in one's theory, separate from the remaining empirical premises, which would specify the meaning of the theoretical terms in propositions true by definition. Instead, the meaning of those terms is conveyed by the conjunction of empirical premises making up the theory; any one of the premises can be tested, and there is no need to claim that some particular members of the class of conditionals employing a given theoretical term are immune from revision. The meaning of a theoretical term is conveyed by the place of the theoretical term in the structure of relations postulated between observables and hypothetical states. The theory itself suffices to spell out those relations. Additional a priori definitions are not required.

2. Construct validity can be established without the use of a definitional criterion for a term, and there is no need to search for such a criterion before research can proceed. Once one realizes that one can be certain that a given theoretical state X obtains only if one is certain the theory is true, the demand for such definitional criteria is seen to rest on a mistake. The grounds for believing that X obtains can only be as strong as the grounds for believing the theory itself. One can employ other premises of that same theory as assumptions in showing that X obtains—premises which in the given context are taken for granted—because such premises are part of a theory which successfully explains observations. But such premises no more provide certainty than any theory ever provides certainty. Every conditional adding to the pattern of

relations defining a theoretical term can be tested; none of them need to be set as true by definition or as immune from revision.

3. It is heuristically useful to think of hypothetical states as existing independently from the observable 'symptoms' by which they become manifest (Hempel, 1965, pp. 205-206; Loevinger, 1957, p. 642; Plutchik, 1968, pp. 37-38). Hypothetical states are sometimes conceived to exist as states distinct from observable stimuli or responses, having effects within the system which are as yet to be discovered. Later discoveries are sometimes described as if they add a new characterization to a current construct.

Obviously this sort of talk is difficult to reconcile with the claim that theoretical discourse is really just discourse about clusters of observations, and the idea of independent existence of constructs is not easy to accept along with the criterion view. But the account above can make sense of these ideas.

There is a difference between picking out or identifying a state in terms of a particular characteristic—that is, as standing in such and such relations to other states—and defining the state in those terms. Because the theory can always be revised, the relations between theoretical terms are always open to revision, and one can never claim that the states are exhaustively characterized by any version of the theory. It is true that the hypothetical terms are identified in terms of their relations to one another and to observables, but the proposed 'definition' of the terms allows for changes. That is, one defines such terms in the following way:

> If there exist states $Q_1 \ldots Q_n$ which bear such and such relations to one another and to stimuli and responses, then those states are $X_1 \ldots X_n$ states.

A revision of the theory will alter one's conception of the relations required between the constructs. When we arrive at a different idea of those relations, we may or may not decide to call the states by the same names. If we do, we show that a given hypothetical state can be characterized by other relations which are as yet undiscovered. The current understanding of relations between a given X, other theoretical terms, stimuli, and behavior does not exhaustively characterize X, nor are any particular claims in it necessarily true of X. This is why such 'definitions' allow for further elaboration or revision.

4. *Convergent indicators* are observations which mutually confirm an inference concerning the presence or magnitude of some construct (Campbell, 1960; Campbell and Fiske, 1959; Cronbach, 1970; Sullivan and Feldman, 1970; Webb, Campbell, Schwartz, and Sechrest, 1966). One might wonder how two or more observations can be tied to the same construct in this way. Why should we ever say that two observations are both indices for the same theoretical term? This problem is particularly acute when the responses providing indices for the given construct occur in altogether different stimu-

lus situations—for example, startle response potentiation and increased Sidman avoidance rates as indices for fear. Why should we ever relate different responses in different stimulus situations to the same construct?

An initial answer to this question is that the same hypothetical state is thought to be causally relevant to the different responses in different situations. The response serves as an index or indicator for the fear construct, because according to one's theory, that response would not occur in that situation if F states did not occur. The hypothetical state plays an essential role in both situations in the elicitation of the given behavior, and for that reason both behaviors are indicators of F states.

This analysis can be put in formal terms. Two observations are considered to be convergent indicators for a given construct if and only if a conditional using that construct is an essential premise in deriving (explaining) those observations. That is, within the theory as given, for response A to occur or for response B to occur, a necessary condition is that X be of a certain value. Occurrence of response A or B then allows one to infer something about X; hence A and B are convergent indicators for X.

The convergence of indicators cannot be demonstrated independently of a theory. One can always explain the occurrence of particular stimulus-response relations in many different ways, and alternative hypotheses are easy to propose. Indicators are convergent only if they are both explained by citing the same construct. If a superior alternative explanation for one of the behaviors was found, the indicators would no longer be convergent, for they would no longer both be explained by the same construct. For example, it is only because one accepts a particular learning theory that one thinks that increased Sidman avoidance and potentiated startle response are both indices of fear. They are classed as indices of 'fear' because within that theory, in order to derive those observations, an assumption about the fear construct is necessary. A superior theory may someday assimilate potentiated startle response to different internal processes, so that it will no longer be an index of fear. But this accords with the intuition that nothing in the observations obliges one to assimilate them to a hypothesized internal process, nor do they provide criteria or definitions for that process. The link between index and construct is only made by the success of a theory.

5. A final point this analysis helps to clarify concerns the claim that to establish the existence of a given hypothetical state, one must find at least *two* convergent indicators for that state (Campbell, 1960, p. 549; Campbell and Fiske, 1959, pp. 102-104; Cook and Campbell, 1979, p. 65; Cronbach, 1970, pp. 467-468; Plutchik, 1968, p. 37). This claim is often defended on the grounds that one must have evidence for the construct which is independent of its definition, so that if one response in a particular situation is used as a criterion for the hypothetical state, a second index is required so as to provide

empirical evidence of its existence.

While we do require at least two indices (or observational predictions) for a given theoretical term, the above explanation provides a poor rationale for this requirement. As argued above, if the second index provides independent evidence for the construct purportedly defined by the first, one would wonder why one needs to mention the hypothetical state at all. Why not simply note the correlation between observations, and drop the construct term? If one index is thought to provide a criterion for the hypothetical state, it is merely an empty form of words to explain the convergence between indicators by citing the hypothetical state.

A clearer explanation of the need for at least two convergent indicators appeals to the principle of parsimony. To claim that a hypothetical state X exists is just to claim that the theory employing X states is true. The truth of the theory would commit one to the existence of states and processes standing in the stipulated relations to one another and to observables. Now if one had but one empirical index for a given construct, dropping it from the theory would in no way decrease one's ability to explain the patterns of dependence between stimuli and responses. The construct would not fulfill the requisite role of assimilating different observations to a common explanatory variable. The principle of parsimony would therefore lead one to drop that construct from the theory. In that sense, the truth of the theory would not commit one to the existence of constructs linked to just one observation, as one could drop such a construct and still have both the same grounds for accepting the theory as true, and the same ability to explain observations. The principle that two convergent indicators are needed to assure the reality of constructs is not derived from the need for a definition of theoretical terms, but simply from the requirement for parsimonious theories.

References

American Psychological Association. Standards for educational and psychological tests and manuals. Reprinted in D.N. Jackson and S. Messick (Eds.), *Problems in human assessment*. New York: McGraw-Hill Book Company, 1967.

Anastasi, A. The concept of validity in the interpretation of test scores. *Educational and Psychological Measurement*, 1950, *10*, 67-78.

Anastasi, A. *Psychological testing* (3rd ed.). New York: Macmillan, 1968.

Bechtoldt, H. Construct validity: A critique. *American Psychologist*, 1959, *14*, 619-629.

Brodbeck, M. Logic and scientific method in research on teaching. In N.L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand Mcnally, 1963.

Campbell, D.T. Recommendations for APA test standards regarding construct, trait, or discriminant validity. *American Psychologist*, 1960, *15*, 546-553.

Campbell, D.T., and Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, *56*(2), 81-105.

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: Clarendon Press, 1980.
Cook, T.D., and Campbell, D.T. *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally College Publishing Company, 1979.
Cronbach, L.J. Test validation. In R.L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington D.C.: American Council on Education, 1970.
Cronbach, L.J., and Meehl, P.E. Construct validity in psychological tests. *Psychological Bulletin,* 1955, *22,* 281-302.
Gray, J.A. *The psychology of fear and stress.* London: Weidenfeld and Nicolson, 1971.
Gray, J.A. *Elements of a two-process theory of learning.* London: Academic Press, 1975.
Hempel, C.G. *Aspects of scientific explanation and other essays in the philosophy of science.* New York: The Free Press, 1965.
Horst, P. *Psychological measurement and prediction.* Belmont, California: Wadsworth Publishing Company, Inc., 1966.
Loevinger, J. Objective tests as instruments of psychological theory. *Psychological Reports,* 1957, *3,* 635-694. (Monograph supplement 9)
Lord, F.M. Some perspectives on "The attenuation paradox in test theory." *Psychological Bulletin,* 1955, *52,* 505-510.
Miller, N.E. Objective techniques for studying motivational effects of drugs on animals. In S. Garettini and V. Ghetti (Eds.), *Psychotropic drugs.* Amsterdam: Elsevier, 1957.
Mowrer, O.H. *Learning theory and behavior.* New York: John Wiley and Sons, Inc., 1960.
Nagel, E. *The structure of science.* London: Routledge and Kegan Paul, 1961.
Nunnally, J.C. *Psychometric theory.* New York: McGraw-Hill Book Company, 1967.
Nunnally, J.C., and Durham, R.L. Validity, reliability, and special problems of measurement in evaluation research. In E.L. Struening and M. Guttentag (Eds.), *Handbook of evaluation research.* Beverly Hills, California: Sage Publications, 1975.
Plutchik, R. *Foundations of experimental research.* New York: Harper and Row, 1968.
Popper, K.R. *Conjectures and refutations: The growth of scientific knowledge.* New York: Basic Books, 1962.
Popper, K.R. *The logic of scientific discovery* (2nd ed.). New York: Harper and Row, 1968.
Putnam, H. It ain't necessarily so. *Journal of Philosophy,* 1962, *49,* 658-671.
Putnam, H. The analytic and the synthetic. In H. Feigl and G. Maxwell (Eds.), *Minnesota studies in the philosophy of science* (Vol. 3). Minneapolis: University of Minnesota Press, 1966.
Putnam, H. Is semantics possible? In S.P. Schwartz (Ed.), *Naming, necessity, and natural kinds.* Ithaca, New York: Cornell University Press, 1977.
Quine, W.V. Two dogmas of empiricism. *Philosophical Review,* 1951, *60,* 20-43.
Quine, W.V. *Ontological relativity and other essays.* New York: Columbia University Press, 1969.
Skinner, B.F. *About behaviorism.* New York: Vintage Books, 1974.
Sullivan, J.L., and Feldman, S. *Multiple indicators: An introduction.* Beverly Hills, California: Sage Publications, 1979.
Webb, E.J., Campbell, D.T., Schwartz, R.D., and Sechrest, L. *Unobtrusive measures.* Skokie, Illinois: Rand McNally, 1966.