

# Organizational Theories: Some Criteria for Evaluation

**SAMUEL B. BACHARACH**

Cornell University

*A set of ground rules and vocabulary to facilitate focused discussion about the structure of organization and management theories are proposed. The many previous efforts at defining and evaluating theory help establish criteria for theory construction and evaluation. In the establishment of these criteria, description is distinguished from theory, and a matrix of criteria for evaluating the variables, constructs, and relationships that together compose a theory is developed. The proposed matrix may be useful both for defining the necessary components of good theory and for evaluating and/or comparing the quality of alternative theories. Finally, a discussion of the way theories fit together to give a somewhat broader picture of empirical reality reveals the lines of tension between the two main criteria for evaluating theory.*

In order to talk about the nature of the universe and to discuss questions of whether it has a beginning or an end, you have to be clear about what a scientific theory is. (Hawking, 1988, p. 9)

A theory is a statement of relations among concepts within a set of boundary assumptions and constraints. It is no more than a linguistic device used to organize a complex empirical world. As Hall and Lindzey (1957, p. 9) pointed out, the function of a theory "is that of preventing the observer from being dazzled by the full-blown complexity of natural or concrete events." Therefore, the purpose of theoretical statements is twofold: to organize (parsimoniously) and to communicate (clearly).

Many current theories in organizational behavior fail to accomplish this purpose, primarily because they ignore certain generally accepted rules about theoretical statements. Just as a collection of words does not make a sentence, a collection of constructs and variables does not necessarily make a theory.

Students of theory construction have tried to develop a set of rules for the examination of the constructs and variables which are the units of theoretical statements (cf. Dubin, 1969; Chronbach & Meehl, 1955; Blalock, 1968; Schwab, 1980). They also have attempted to develop a set of rules for the examination of the relationships among these units (cf. Blalock, 1969; Cohen, 1980; Nagel, 1961; Hempel, 1965; Stinchcombe, 1968; Popper, 1959; Dubin, 1976; Gibbs, 1972). Nevertheless, the diversity of these perspectives suggests the need for a more specific examination of their rules as applied to organizational studies.

## **What Theory Is Not: Data, Typologies, and Metaphors**

*Description*, the "features or qualities of individual things, acts, or events" (Werkmeister, 1959, p. 484) must be distinguished from *theory*. As Hempel (1965) pointed out, the vocabulary of science has two basic functions: (a) to ade-

quately describe the objects and events being investigated and (b) to establish theories by which events and objects can be explained and predicted. While descriptions may be the source material of theories, they are not themselves theoretical statements. In the organization and management literature, the two are often confused. Specifically, three modes of description must be distinguished from theory: categorization of raw data, typologies, and metaphors.

While some forms of descriptive analysis are often confused with theory, all researchers agree that categorization of data—whether qualitative or quantitative—is not theory. In this context, much of the work in organizational and management science should not be thought of as theory. Categorization characterizes much of the work in these fields, particularly in the realms of business policy/strategy and human resource strategy. One theme in the former case, for example, has been the search for empirical categorizations, or *gestalts*, of organizational environments and characteristics (e.g., Miller, 1986; Miller & Friesen, 1977). One characteristic in the latter case has been the search for a goodness of fit between empirically derived categorizations of business strategy and human resource strategy (e.g., Schuler & Jackson, 1987; Wils & Dyer, 1984). Some of these studies, both quantitative and qualitative, are often particularly rich and thus useful as grounds (Glaser & Strauss, 1967) for theory building (e.g., Dyer & Holder, 1989; Miller, 1987; Miller & Friesen, 1984). In and of themselves, however, they clearly fall in Hempel's (1965) realm of description, not theory.

Other descriptions—specifically, those based upon typologies—have been more abstract in organizing observations (e.g., Blau & Scott, 1962; Etzioni, 1975; Gouldner, 1954). Typologists have implicitly emulated Weber's ideal construct, in that most typologies meet his classic definition of an ideal type . . . "a mental construct formed by the synthesis of many diffuse . . . individual phenomena which are arranged, according to certain one-sidedly accen-

tuated points of view, into a unified analytical construct" (Shils & Finch, 1949, p. 90).

Yet even these abstractions should not be viewed as theory. The one-sided accentuation of which Weber spoke is found in most typologies. For example, in creating a typology of organizations, Blau and Scott (1962) emphasized the beneficiaries; Gouldner (1954), leadership style; and Etzioni (1975), compliance structure. While this one-sided accentuation achieves one of the goals of theory (i.e., eliminating some of the complexity of the real world), and while typologies are more abstract than a categorical description of raw data, such typologies are limited to addressing the primary question asked by descriptive researchers—the question of *what*, rather than the more theoretical *how*, *why*, and *when*.

In recent years, metaphors have become popular in organizational studies. Broadly speaking, a metaphor is a statement that maintains that two phenomena are isomorphic (i.e., they have certain properties in common) (Brodbeck, 1959). Unlike the case of categorical analysis of raw data (What are the phenomena?), or the case of typology (What is the most important aspect of the phenomenon?), the metaphor is used to ask how the phenomenon is similar to another (often unrelated) phenomenon. Some of the most well-known metaphors include the notions of organizations as "loosely coupled systems" (Weick, 1976) and as "garbage cans" (Cohen, March, & Olsen, 1972).

Metaphors are powerful literary tools. Robert Burns' comparison of his love to a red rose evokes strong emotional imagery. It does not need to evoke a series of analytical questions about love; the description itself suffices. To be of use in the development of theory in organizational behavior, a metaphor must go beyond description and be a useful heuristic device. That is, the imagery contained in the metaphor must assist the theorist in deriving specific propositions and/or hypotheses about the phenomenon being studied. In this context, metaphors are not theories but may well serve as precursors to theories, and should

be judged on that basis. For example, if one chooses to view organizations as "nonconflictual zeppelins," it's his or her prerogative to do so. What must be evaluated is not whether organizations are in fact "nonconflictual zeppelins," but rather the propositions and hypotheses derived from the imagery. If one's image of organizations as "nonconflictual zeppelins" is to thrive, then it is because the quality of propositions and hypotheses generated by this image is better than the quality of those generated by other alternative images such as "garbage can models" or "loosely coupled systems."

### **What Theory Is**

Building on the works of previous students of theory construction (e.g., Dubin, 1969; Nagel, 1961; Cohen, 1980), researchers can define a theory as a statement of relationships between units observed or approximated in the empirical world. *Approximated* units mean *constructs*, which by their very nature cannot be observed directly (e.g., centralization, satisfaction, or culture). *Observed* units mean *variables*, which are operationalized empirically by measurement. The primary goal of a theory is to answer the questions of *how*, *when*, and *why*, unlike the goal of description, which is to answer the question of *what*.

In more detailed terms, a theory may be viewed as a system of constructs and variables in which the constructs are related to each other by propositions and the variables are related to each other by hypotheses. The whole system is bounded by the theorist's assumptions, as indicated by Figure 1.

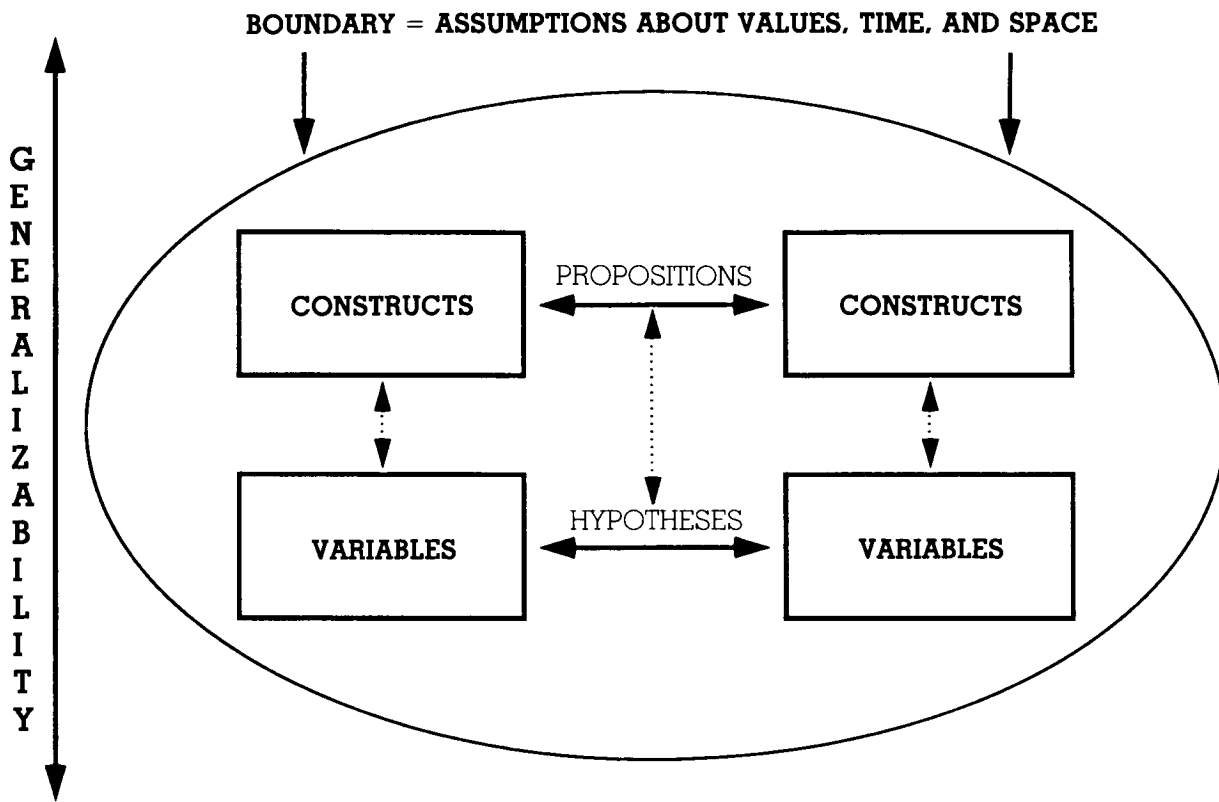
### **Boundaries of Theories**

The notion of boundaries based on assumptions is critical because it sets the limitations in applying the theory. As Dubin (1969) maintained, all theories are constrained by their specific critical bounding assumptions. These assumptions include the implicit values of the theorist and the often explicit restrictions regarding space and time.

Values are the implicit assumptions by which a theory is bounded. Theories cannot be compared on the basis of their underlying values, because these tend to be the idiosyncratic product of the theorist's creative imagination and ideological orientation or life experience. This may explain why perpetual debates such as those between Marxists and Structural Functionalists have made so little progress over the years. As Weber pointed out, the value-laden nature of assumptions can never be eliminated. Yet if a theory is to be properly used or tested, the theorist's implicit assumptions which form the boundaries of the theory must be understood. Unfortunately, theorists rarely state their assumptions. Thus, while Mintzman's (1970) extensive and fascinating discussion of the influence of Weber's personal reality on his theoretical product does not serve to expand or change Weber's theory, it does assist in explicating the implicit values which bound his theory.

An example of how a theorist's values may be manifest in a theoretical debate may be found in a classic debate over the concept of power. Parsons maintained that power is essentially the mobilization of resources and thus is not a zero-sum game. On the other hand, C. Wright Mills (1956) maintained that power is the control of resources, and thus is a zero-sum game. An argument may be made that this differential orientation toward power is based on these theorists' different values. Parsons, viewing society and organizations as functional and consensual systems, ignored the possibility of finite resources as a potential source of conflict. Rather, he saw resources as being capable of perpetual expansion. Mills, on the other hand, seeing society and organizations as stratified and conflictual entities, ignored the expansive nature of resources and instead emphasized finite resources and a zero-sum notion of power. The interaction between these two theorists is thus implicitly a collision of values.

The current debate over the application of the concept of culture to the organizational context also may be viewed as an example of the im-



**Figure 1. Components of a theory.**

explicit debate over values if not ideology. During the early resurrection (by organizational theorists) of the concept of culture (Schein, 1985; Deal & Kennedy, 1982), most theorists chose to view it as an integrative normative device. In doing so they implicitly drew on the values that underlie functionalist theorizing by scholars such as Radcliffe-Brown (1949), Malinowski (1962), and Durkheim (1933). Therefore, their implicit functional orientation (placing emphasis on sustaining the organization as a whole) may make them vulnerable to criticism that they serve the interests of management. In more recent work from the conflict theory perspective, culture is viewed as an organizational mechanism for the normative coercion of the individual worker. For example, Kunda (in press), in studying a com-

pany that is often held up as a shining model of the positive normative impact of culture, unveiled a story of oppression where others told a tale of productivity. While these two orientations toward culture are not inherently inconsistent, they do show the effects of different values on the construction of theories about organizations.

While values often can only be revealed by psychoanalytic, historical, and ideological studies of the theorist (e.g., Gay's work on Freud, 1988; Mintzmann's work on Weber, 1970), spatial and temporal assumptions are often relatively apparent. Spatial boundaries are conditions restricting the use of the theory to specific units of analysis (e.g., specific types of organizations). Temporal contingencies specify the historical

applicability of a theoretical system. Taken together, spatial and temporal boundaries restrict the empirical generalizability of the theory.

While most theories are limited by spatial and temporal restrictions, some are more bounded by one than the other. For example, some theories may be unbounded in time, but bounded in space. That is, these theories are only applicable to specific types of organizations, but can be applied over different historical periods. Other theories are unbounded in space (that is, they may be applicable to many types of organizations), but very much bounded in a specific temporal context. Finally, theories may be relatively unbounded in both space and time. Such theories have a higher level of generalizability than those bounded in either or both space and time, *ceteris paribus*. Of course, this generalizability requires a higher level of abstraction, which means that the theory sacrifices the level of detail needed to fit a specific situation. This leads to the paradox that some of the most detailed theories and elaborate studies about organizations are not generalizable enough to build a cumulative body of research on (e.g., Goffman's theory of total institutions). On the other hand, some of the most abstract and broad perspectives on organizations, while not necessarily rich in detail, have provided a critical basis for cumulative research (e.g., Hannan and Freeman's, 1977, population ecology, and Kimberly and Miles', 1980, life-cycle theory).

Implied in the notion of generalizability are different levels on which one can theorize. This implicit continuum stretches from empirical generalizations (rich in detail but strictly bounded in space and/or time) to grand theoretical statements (abstract, lacking in observational detail, but relatively unbounded in space and/or time).

### **Variables, Constructs, and Relationships**

Within these boundaries lies the stuff of theory. On a more abstract level, propositions state the relations among constructs, and on the more concrete level, hypotheses (derived from the

propositions) specify the relations among variables. In this context, theorists must be specific in how they use the notions of constructs and variables. Theorists should not use these terms synonymously. Constructs may be defined as "terms which, though not observational either directly or indirectly, may be applied or even defined on the basis of the observables" (Kaplan, 1964, p. 55). A variable may be defined as an observable entity which is capable of assuming two or more values (Schwab, 1980). Thus, a construct may be viewed as a broad mental configuration of a given phenomenon, while a variable may be viewed as an operational configuration derived from a construct. Schwab listed a number of examples of such constructs and their related variables (e.g., performance: sales or return on investment; cohesion: rate of interpersonal interaction or member voting patterns; leader consideration: member perceptions of specific supervisory behavior).

Created within the context of specified boundaries and built from abstract constructs or their more concrete manifestations (variables), theoretical systems take the form of propositions and proposition-derived hypotheses. While both propositions and hypotheses are merely statements of relationships, propositions are the more abstract and all-encompassing of the two, and therefore relate the more abstract constructs to each other. Hypotheses are the more concrete and operational statements of these broad relationships and are therefore built from specific variables.

### **Generation of Criteria for the Evaluation of Theories**

No evaluation of a theory is possible unless researchers first establish those broad criteria by which it is to be evaluated. Based on previous work (e.g., Popper, 1959; Nagel, 1961; Hempel, 1965), the two primary criteria upon which any theory may be evaluated are (a) falsifiability and (b) utility.

## **Falsifiability**

Falsifiability determines whether a theory is constructed such that empirical refutation is possible. While the idealistic goal of science is the pursuit of universal truth, most philosophers of science would agree that theories can never be proven, only disproven (cf. Nagel, 1961; Popper, 1959). As Popper (1959, p. 41) maintains, "It must be possible for an empirical scientific system to be refuted by experience."

Theories are thus like the accused in an American courtroom—innocent until proven guilty. The problem with organizational studies is that theories are often stated in such a vague way that the theorists can rebut any discrediting evidence. Just as no person can be above the law, no theory ought to be constructed in such a way that it is forever exempt from empirical refutation. If researchers are to avoid wading through ever deeper piles of irrefutable statements disguised as theories, they must be able to discard such false theories. To be able to do this, they must try to construct theories that are coherent enough to be refuted.

## **Utility**

Utility refers to the usefulness of theoretical systems. As Bierstedt (1959) pointed out, utility may be viewed as "the bridge that connects theory and research" (p. 125). At the core of this connection are explanation and prediction. That is, a theory is useful if it can both explain and predict. An explanation establishes the substantive meaning of constructs, variables, and their linkages, while a prediction tests that substantive meaning by comparing it to empirical evidence.

One problem of incomplete theoretical systems is that they are often used to make predictions, yet they do not provide explanations. In this context, Kaplan (1964) spoke of the ancient astronomers, who were able to make superb predictions but were incapable of providing adequate explanations of observed phenomena. Thus, in organizational behavior, when re-

searchers accept predictive statements as theory (e.g., the proposition: the greater the organizational size, the greater the horizontal differentiation; or its derived hypothesis: the greater the number of employees, the greater the number of departments), they sound a bit like the ancient astronomers. Only when theory shows how and why larger organizations have more departments will it be able to explain as well as predict.

## **The Falsifiability of Variables, Constructs, and Relationships**

With an understanding of the components of theory at different levels of abstraction (variables, constructs, and the relationships that connect them) and the two main types of criteria (falsifiability and utility), researchers can begin to understand the way these criteria can be applied to theory. Because constructs and variables are the building blocks of hypotheses and propositions, theorists must first evaluate them before analyzing the relational properties of theories. If they are working with inappropriate constructs and variables, how these constructs and variables are assembled into hypotheses and propositions is irrelevant. All parts of a bridge may fit together perfectly, but if this bridge is constructed of "silly-putty," it is not a good idea to drive over it.

By beginning the analysis with variables and constructs, researchers are not excluding the possibility that theory building or evaluation is a process which begins with the examination of the relationships in hypotheses and propositions, or what Kaplan refers to as the paradox of conceptualization. As Kaplan (1964) noted, "The proper concepts are needed to formulate a good theory, but we need a good theory to arrive at the proper concepts" (p. 53).

## **The Falsifiability of Variables: Measurement Issues**

Although constructs contained in propositions may be defined in terms of other constructs, con-

|               | Falsifiability     | Utility               |
|---------------|--------------------|-----------------------|
| Variables     | Measurement Issues | Variable Scope        |
| Constructs    | Construct Validity | Construct Scope       |
| Relationships | Logical Adequacy   | Explanatory Potential |
|               | Empirical Adequacy | Predictive Adequacy   |

**Figure 2. A Framework for evaluating theories.**

struct-derived variables in hypotheses must be defined in an operationally specific manner. By definition, the *raison d'être* of a variable is to provide an operational referent for a phenomenon described on a more abstract level (e.g., a construct). As such, in order for a variable to be operationally specific, that variable must be defined in terms of its measurement. Kerlinger referred to the first type of definition (i.e., clarity and parsimony, which are sufficient for constructs) as a "constitutive" definition, while the second type (i.e., necessary for variables) is referred to as an "operational" definition. In an employment setting, the former is exemplified by the construct *power* being defined in terms of *dependence* (Emerson, 1962), while the latter is exemplified by defining the variable *deduced* from the construct *power* (i.e., *alternatives*) in terms of "the number of other job opportunities an individual has available at a given time" (Bacharach & Lawler, 1980).

Furthermore, for a theory to be falsifiable, these operationalized variables must be coherent. That is, they must pass the tests of being a good measurement model: validity, noncontinuousness, and reliability. Since discussions of variable validity (content and face validity) are beyond the current scope of this article, the discussion will be confined to the importance of using variables with adequate variance for logical analysis, and adequate reliability (stability).

Unless the theorist's hypotheses incorporate

variables which can be meaningfully and correctly measured, any variance that may exist in the object of analysis is essentially unobservable, making the theory not subject to disconfirmation. This situation describes the test of noncontinuousness. Continuous antecedents and consequences make it impossible to specify relevant time and space parameters. Thus, hypotheses incorporating such variables are never subject to empirical disconfirmation.

If the antecedent is a necessary condition for the consequent condition, a continuous antecedent would render the proposition untestable. However, if the antecedent is a sufficient condition for the consequent, a continuous consequent would make testing impossible. For example, Young's (1988) critique of population ecology theory is that it has "difficulty with the definition of organizational death" (p. 7). Specifically, she argued that across studies, the definition of the constructs included in the primary propositions (i.e., organization/species, birth, and death) varies and is ambiguous. Thus, if the population theorist assumes that increased structural inertia is a sufficient condition for survivability (selection), the derived hypothesis would be untestable given a derived operationalization of death that is so ambiguous as to allow all organizations to have an equally high selection rate. Likewise, if the theorist assumes that increased structural inertia is a necessary condition for an increase in survivability, derived hypotheses cannot be tested if inertia, which is itself a problematic construct according to Young (1988), is at the same level in all organizations and species.

Thus, this condition demands that the theorist specify the time and space parameters of the variables embedded in the hypothesis so that constructs may be meaningfully measured. That is, the failure to specify the time and space parameters embedded in construct measurement makes it impossible to falsify constructs, and hence theory. Young (1988), in claiming that population theorists fail to specify the time and space parameters of death, implicitly argued

that these theoretical statements may not be subject to falsification.

Yet having noncontinuous variables is not enough. Many scholars in organizational and management theory have taken the criterion of reliability far too lightly. For example, in a recent evaluation of the technological innovation literature, Shenhav, Haberfeld, and Cohen (in press) showed that across numerous studies, measures of innovation and scientific productivity achieve an acceptable level of reliability (i.e.,  $\alpha = .70$  or greater) only in specific contexts. These researchers argued that such relatively unstable measures have often been used inappropriately to test theories. Their instability is attributed to the fact that most innovation and productivity measures are based on socially constructed notions which take on different meanings in different contexts.

### **The Falsifiability of Constructs: Construct Validity**

It may be useful to define constructs in terms of other established and well-understood constructs. If the purpose of a proposition is to communicate the relationship between two or more constructs, then (unlike for variables) the only operational criteria which these constructs must meet is that they have good clarity and parsimony. However, constructs which are explicitly operationalized must have their variables undergo the previously mentioned tests for variable falsifiability as a first step in the test for construct validity.

When combined, the indicators of construct and variable falsifiability are no less than the criteria for construct validity itself. In this context, in order to provide evidence of falsifiability, individual constructs contained in propositions must meet minimum standards of construct validity, while individual variables, derived from constructs and contained in proposition-derived hypotheses, must meet the measurement model criteria.

To achieve construct validity, at the very least the responses from alternative measurements of

the same construct must share variance (i.e., convergent validity) (Schwab, 1980), while the identified objects of analysis must not share attributes and must be empirically distinguishable from one another (discriminant validity).

In determining convergent validity the theorist must confirm that "evidence from different sources gathered in different ways all indicate the same or similar meaning of the construct" (Kerlinger, 1973, p. 463). In determining discriminant validity, the theorist must confirm that "one can empirically differentiate the construct from other constructs that may be similar, and that one can point out what is unrelated to the construct" (Kerlinger, 1973, p. 463). If two independent variables have high collinearity, it is impossible to talk of their independent effects.

The formulation of propositions incorporating valid constructs is a responsibility not always taken seriously by scholars of organizational and management theory. Examples regarding problems with both convergent and discriminant validity may be found. Pelz and Andrews (1966) examined the convergence of their self-reported, nonjudgmental measures of the construct *scientific productivity*. They found a relatively low magnitude of correlation among the various indicators (e.g., application outcomes, lab reports). In spite of the low magnitudes of convergence, the validity, and hence, falsifiability, of the construct was confirmed by these researchers on the grounds that low magnitude correlations could be expected given that each measure was designed to tap a different aspect of scientific productivity. Rejecting this argument, Shenhav and Haberfeld (in press) argued that Pelz and Andrews' findings should be taken as an indication of low construct validity, since different measures of the same construct should be highly correlated. That is, Shenhav and Haberfeld suggested that Pelz and Andrews did not take the issue of construct validity seriously enough.

Regarding discriminant validity, Young (1988) was critical of Freeman and Hannan (1983) who, in examining the relative impact of variability in



sales and seasonality on the viability of resort towns, measured seasonality on a quarterly basis. For some resort towns (e.g., South Lake Tahoe), sales and seasonality measures (i.e., variables) are nearly perfectly confounded, thus raising the possibility of limited discriminant validity.

Similarly, Locke, Saari, Shaw, and Latham (1981, p. 145) are critical of researchers examining the impact of goal difficulty on the goal-task performance relationship. There is often no empirical difference between an easy-goal condition and a moderate-goal condition, for "a common problem with easy-goal subjects is that their goals are so easy that once they are revealed, they set new higher goals to have something to do, which means that they are no longer genuine easy-goal subjects" (Locke et al., 1981, p. 142). Furthermore, in attempting to explain the inconsistent findings regarding the impact of individual differences on the goal setting-task performance relationship, Locke et al. found inadequate reliability and validity of personality measures to be a primary source of these inconsistent findings. This situation reminds researchers that embedded in the construct validity criteria is the assumption that the variables were correctly measured.

The key to meeting the falsifiability criterion for constructs thus lies in showing that variables which should be derived from the constructs are indeed correlated with the construct, regardless of the procedure used to test correlation, and those variables which should be unrelated to the construct are indeed uncorrelated. In adopting such a definition of convergent validity, Schwab (1980) maintained that Campbell and Fiske's (1959) multitrait-multimethod matrix methodology, while the most popular method for assessing construct validity, may provide only limited evidence, because correlation coefficients are likely to be strongly influenced by sample size, and convergence is likely to reflect hard-to-avoid common-method variance.

For this reason the explication of construct validity as an evaluative indicator began with the

phrase "at the very least. . . ." In determining the falsifiability of constructs, convergent and discriminant validity tests should not be used in isolation. Thus, a well-grounded assessment of the falsifiability of a construct should most likely go beyond assessments of convergent and discriminant validity. Other supplementary tests include factor analysis and tests of concurrent and predictive validity. Van de Ven and Chu (1987), for example, subjected their innovation effectiveness construct to rigorous analysis, including multimethod convergent and discriminant validity tests at two points in time, factor analysis, and tests of concurrent and predictive validity. While construct validity (like a theory) can only be rejected, never confirmed, Van de Ven and Chu's tests provide stronger evidence that their construct is indeed falsifiable.

The importance of this falsifiability criterion cannot be overstated. Unless evidence is presented of the falsifiability of constructs and variables, construct validity—critical to the building of theory—can never be achieved. Schwab (1980) maintained that because organizational and management theorists and researchers place such emphasis on the examination of relations between independent and dependent phenomena, without examining the characteristics of these phenomena, knowledge of the relationships among phenomena "is not as great as is believed, and (more speculatively) not as great as would be true if the idea of construct validity received greater attention" (p. 4).

With this in mind, Anderson and Gerbing (1988) proposed that in undertaking LISREL-based theory testing, researchers should first assess validity for the building blocks of the theory (i.e., confirm the falsifiability of the constructs and variables embedded in propositions and hypotheses) by separate estimation and, where necessary, respecification of the measurement model, and only afterwards simultaneously estimate the measurement and structural submodels. When a theory is evaluated, the boundary between theory construction and theory testing often becomes blurred. As such, theorists have

the responsibility to ensure that their hypotheses and propositions contain constructs and variables which can be researched (i.e., are capable of disconfirmation).

### **The Falsifiability of Relationships: Logical and Empirical Adequacy**

A number of criteria with which to evaluate constructs and variables have been emphasized. Now the adequacy of the relational elements of a theoretical system (i.e., the linkages established among the component constructs and variables) must be evaluated. Going back to the analogy of the bridge, now that indicators can determine that the bridge is not made out of "silly-putty," it should be established that the linkages among the component elements are sound.

When evaluating the falsifiability of the relational properties of theoretical systems, theorists must examine both the logical adequacy of the propositions and hypotheses (and their interrelationships) as well as their empirical adequacy (the capacity of the relationships implied in propositions and hypotheses to be operationalized).

#### **Logical Adequacy**

*Logical adequacy* may be defined as the implicit or explicit logic embedded in the hypotheses and propositions which ensures that the hypotheses and propositions are capable of being disconfirmed. In this context, individual propositions and hypotheses must satisfy the following two criteria: (a) They must be nontautological, and (b) The nature of the relationship between antecedent and consequent must be specified.

*Criterion A.* For a proposition or hypothesis to be falsifiable, the antecedent and the consequent may not be epiphenomenal. That is, the sheer existence of the antecedent may not automatically imply the existence of the consequent. A proposition such as "the greater the work load of an individual, the greater the level of felt role conflict" would clearly be tautological if the ex-

istence of role conflict automatically implied the existence of high work load or role overload, or if role conflict was measured by an index including work load or role overload (Newton & Keenan, 1987).

However, note that such contamination at the variable level (in this example, if role conflict was measured by an index including work load or role overload) does not mean that the hypothesis itself is tautological. It may simply be necessary to operationalize it with differently defined (i.e., noncontaminated) variables. Similarly, the presence of a tautological hypothesis does not necessarily imply a tautological proposition. Clearly, this issue is related to the issue of discriminant validity and the question of how constructs are defined.

In a recent critique of the population ecology model, Young (1988) maintained that she found a tautology in the logic underlying Hannan and Freeman's (1984) theory of structural inertia (with inertia, in this case, defined as a low rate of change). She claimed the tautology emerged when these theorists defined organizational reproducibility in terms of "having nearly the same structure today as it had yesterday" (p. 154), and then posited that organizational reproducibility generates "strong inertial pressures."

There are similar examples in micro-organizational behavior theory. Vecchio (1987), in examining Hersey and Blanchard's Situational Leadership Theory (SLT), pointed out another example of tautology in theory. SLT proposes that, in order to account for leader effectiveness, theorists must consider the appropriateness of leader style, operationalized in terms of task orientation and relationship orientation, in a given situation, operationalized in terms of individual or group level maturity. Tautological reasoning is used by these theorists when they define "effectiveness" (the dependent variable) in terms of "appropriateness of leader style" (a key component of their independent variable).

In these two examples, a tautological proposition or hypothesis is self-verifying and, therefore, not subject to disconfirmation.

*Criterion B.* The theorist must incorporate in propositions and hypotheses an explicit statement of whether the antecedent is a necessary, sufficient, or necessary and sufficient condition for the consequent. This specification determines the nature of the data required to adequately test the theory. For example, if theorists claim that job dissatisfaction is a necessary condition for physical stress symptomology, they must search for physical stress symptoms not preceded by job dissatisfaction in order to reject the hypothesis. On the other hand, if job dissatisfaction is considered sufficient for physical stress symptoms (Kemery, Bedeian, Mossholder, & Touliatos, 1985; Kemery, Mossholder, & Bedeian, 1987), they must search for a case in which individuals experience job dissatisfaction without reporting physical stress symptoms in order to reject the hypothesis.

By failing to explicitly specify the nature of these logical links, organizational theorists make it impossible for their theories to ever be disproved. Each critique of a theory inevitably leads to a response that the criticism was based on a misunderstanding of the author's original logic. The responsibility for the specification of these linkages must be taken seriously if theorists are ever to move beyond the deluge of criticism and counter-criticism. If "charismatic leadership" is sufficient to bring about organizational change, then that should be clearly stated. On the other hand, if "charismatic leadership" is only necessary to bring about organizational change, then that should be explicitly stated. If theorists are so self-confident as to posit that a particular antecedent is both necessary and sufficient for a particular consequence, then by all means, let them run it up the flagpole.

### **Empirical Adequacy**

*Empirical adequacy* is the second criterion for evaluating the falsifiability of the relationships embedded in a theory. An empirically adequate theory is one in which the propositions and hypotheses may be operationalized in such a man-

ner as to render the theory subject to disconfirmation. If a theory is operationalized in such a way as to preclude disconfirmation, then it is clearly not falsifiable. Specifically, propositions and hypotheses should satisfy the following criterion:

*There either must be more than one object of analysis or that object of analysis must exist at more than one point in time.*

In order for a theory to be subject to disconfirmation, some variation in the object of analysis must be observable. For example, a theory of interorganizational relations in a specific industry necessarily deals with just a single object of analysis—a multinational interorganizational network (Aldrich, 1980). Because the theory pertains to only one network (i.e., it is spatially bound), the theory must be proposed in such a way as to permit longitudinal analysis. On the other hand, a theory examining a number of individual organizations within that network may be tested at a single point in time, since it is not spatially bound.

In addition, empirical adequacy at the relational level cannot be achieved if the variable does not meet standards of a good measurement model. If the variables included in a hypothesis under evaluation are inherently unmeasurable or unstable, the satisfaction of the present criterion becomes impossible.

### **Utility of Constructs, Variables, and Relationships**

Not only do theorists have the responsibility for evaluating the falsifiability of these variables, constructs, and relationships, but they also have the responsibility for evaluating the utility of these variables, constructs, and relationships.

#### **The Utility of Variables and Constructs: Scope**

For adequate scope, the variables included in the theoretical system must sufficiently, although parsimoniously, tap the domain of the constructs in question, while the constructs

must, in turn, sufficiently, although parsimoniously, tap the domain of the phenomenon in question. This is because constructs and variables with broader scope allow hypotheses and propositions to have greater overall explanatory power.

A construct or variable which is content deficient (Schwab, 1980) limits the generality of the theory encompassing the construct or variable. For example, as Bamberger (in press) maintains, in the case of technological innovation, a useful measure should, at the very least, tap that construct's quantitative (i.e., the number of innovations initiated or adopted) and qualitative elements (i.e., the innovation's utility, depth, originality/radicalness). Furthermore, to recognize that not all technological innovations emerge out of an R&D subunit, measures of organizational innovativeness which are broad in scope, and thus higher in overall utility, should be able to tap all innovations, regardless of where they originate within the organization. However, as often is the case with organization and management theory, the scholars concerned with innovation theory have sacrificed broad scope for the sharper focus that they believe enhances accuracy and parsimony.

Unfortunately, this process has led to theories which are no more than compilations of isolated variables and constructs, often trapping theorists on the level of empirical generalization and variable-driven analysis, and making impossible a truly parsimonious theoretical system. The goal must therefore be the achievement of a balance between scope and parsimony.

### **Utility of Relationships: Explanatory Potential and Predictive Adequacy**

When evaluating the utility of the relationships embedded in theoretical systems, theorists must examine both the substantive as well as probabilistic elements of propositions and hypotheses (and their interrelationships). In Figure 2, these two indicators of utility are referred to as

(a) explanatory potential and (b) predictive adequacy.

The explanatory potential of theories can be compared on the basis of (a) the specificity of their assumptions regarding objects of analysis, (b) the specificity of their assumptions regarding determinative relations between antecedent and consequent, and (c) the scope and parsimony of their propositions.

(a) The Specificity of the Assumptions About the Objects of Analysis. As noted in the previous discussion of boundaries, theories cannot be compared on the basis of the content of their assumptions. However, a theory in which these assumptions are explicit is clearly preferable to one in which they are not. For example, a continuing underlying tension in macro-organizational theory is the usually implicit assumption of independence between units of an organization versus the assumption of an interdependence between those units. Theorists supporting the environmental-determinism perspective, particularly those adopting *selection* as a key concept (i.e., population ecology theory), have consistently assumed that the units under study (i.e., members of a species) are free actors, independent of, and not part of any other higher level organization (Young, 1988).

On the other hand, theorists supporting the strategic-choice perspective have implicitly stressed the interdependence of units within and across organizational species and, in particular, the political implications of such interdependence (Thompson, 1967; Lawrence & Lorsch, 1967; Child, 1972; Van de Ven, 1979; Bacharach & Lawler, 1980; Perrow, 1986). It should be obvious that the elaboration or synthesis of organizational theories cannot proceed without specifying this sort of assumption. Some attempts at such elaboration and synthesis have been provided by Tolbert (1985) and Hrebiniak and Joyce (1985).

(b) The Specificity of the Substantive Nature of the Relationship Between the Antecedent and Consequent. Although the specification of necessary or sufficient conditions is essential to the

testability of the theory, the application of a theory in explanation requires additional assumptions about the substance of the linkages (i.e., the sense in which the antecedent is necessary or sufficient for the consequent). For example, one of the primary critiques made by supporters of the environmental determinism perspective (e.g., Hannan & Freeman, 1977; Aldrich, 1979) against the strategic choice perspective is that supporters of the latter fail to recognize the environmental constraints acting upon the proposed linkage between managerial decision making and organizational adaptation and change. In his critique of Child's (1972) defense of the strategic choice perspective, Aldrich (1979, p. 160) made that argument most succinctly: "Environmental selection processes set the limits within which rational selection among alternatives takes place. Prior limits and constraints on available options leave little room for maneuvering by most organizations. . . ." Aldrich did not argue that strategic choice is irrelevant as a tool by which organizational change may be explained. Rather, he argued that the strategic choice theory is lacking in explanatory potential because (a) the potential for organizational decision makers to cause change within the organization is severely limited by environmental pressures (Van de Ven, 1979) and (b) the theory does not clearly specify those conditions necessary for the hypothesized determinative relations existing between the variables to be valid.

Another way in which the substantive nature of the relationship between antecedent and consequence can be blurred is if the nature of causal linkages is dictated by methodological convenience. Both Bailey (1970), and Hage and Meeker (1987) have pointed out that theorists have been prone to concern themselves almost exclusively with recursive causal linkages to the point of nearly ignoring other possible types, such as the teleological (i.e., the notion that a cause is the end state toward which an event is leading), dialectical, or reciprocal (e.g., feedback mechanisms). The clearest example of this

is the proliferating misuse of path analysis (and most recently, LISREL analysis) to determine the causal ordering of a set of variables.

Although path and structural equation (e.g., LISREL) models provide a systematic format for expressing the assumed relationships among variables and estimating the strength of these relationships, the actual ordering of the variables and the nature of their relationship (e.g., causal, simultaneous) can only take place on the assumptive level. For example, by empirically confirming the existence of a reciprocal (rather than recursive) linkage between job satisfaction and physical stress symptomology, Kemery, Mossholder, & Bedeian (1987) provided a clear illustration of how the recursive causal logic implicit in path analysis can have a damaging effect on stress-related theory development.

Furthermore, while the explanatory power of a theory is influenced by the specification of the substantive nature of the relationship among the variables (e.g., is it causal, teleological), explanatory power is also contingent upon the extent to which the actual empirical form of the relationship (e.g., linear, J-curve) is stated. Even though this is an empirical question, it must be answered if the hypotheses are to be useful in explaining the phenomena in question. For example, while Hage and Aiken (1969) hypothesized and confirmed the existence of an inverse linear relationship between routinization and job satisfaction, there may be instead a curvilinear relationship between these two variables. However, in neither case was a curvilinear relationship hypothesized. For the most part, relationships between antecedents and consequences in organizational behavior are often assumed to be linear. Such assumptions of linearity may be viewed as expeditious in terms of limited theory-testing techniques, but naive in the context of theorists' qualitative understanding of people in organizations.

As Hage and Meeker (1987) pointed out, qualitative field research, exemplified by such classic studies as Lipset, Trow, and Coleman's (1956)

*Union Democracy*, and Whyte's (1955) *Street Corner Society*, may go much further in enhancing the explanatory adequacy of a theory than quantitative research. Such techniques are ideal when the theory constructionist is seeking to find and explain causal relations, while quantitative methods are better when the researcher wishes to test these relations on the basis of confirmed or disconfirmed predictions.

(c) Scope and Parsimony. Scope is defined as the range of phenomena encompassed by the theory, and *parsimony* as the ratio of hypotheses to propositions. A theory which can accommodate a wide range of objects of analysis (from small groups of individuals to entire organizational types or more general processes) is clearly preferable to one only applicable to a single object. For example, all-encompassing theories of motivation (e.g., drive, instinct, and conditioning) are preferable to more limited theories which do not presume to explain all motivational phenomena (e.g., goal setting) (Locke et al., 1981).

Given the range of phenomena encompassed by the theory, one with a higher ratio of hypotheses to propositions is preferable to one with a lower ratio. That is, a theory where each proposition covers five hypotheses is preferable to one where each proposition only covers two hypotheses. Although some theorists may view this as merely an aesthetic criterion, the role of theory in science is the integration and simplification of experience. A theory which can best approximate this ideal (i.e., a parsimonious theory) is preferable to one that does less to reduce the complexity of the empirical world. For example, in her critique of Hannan and Freeman's (1984) theory of structural inertia and organizational change, Young (1988) argued that these theorists used more theoretical statements than were absolutely necessary and that this lack of parsimony, as a result, leaves the theory "unnecessarily vulnerable" (presumably to critique and empirical disconfirmation).

Of course, the primary risk in theoretical parsimony is the underspecification of the model.

An example of this may be found in Kemery et al.'s (1987) LISREL-based comparative analysis of stress theory. In this study, the most parsimonious theories, such as those of Beehr and Newman (1978) and Locke (1976), were found to be less plausible than the least parsimonious model, that of Schuler (1982), across a number of model "fit" indices.

The second basis for evaluating the utility of the relationships embedded in a theory is the predictive adequacy of that theory. Does one have to have a predictive theory in order for a theory to be acceptable? According to Kaplan (1964), and Hempel and Oppenheim (1948), the answer is yes. A theoretical system can meet all the criteria discussed previously but will be unacceptable if it is incapable of predicting or is inferior in its predictive adequacy. From hypotheses and propositions some order in the otherwise complex empirical world must be deduced. The degree to which hypotheses and propositions approximate this reality is predictive adequacy—the final determinant of a good theory.

However, in discussing predictive adequacy, it is important to differentiate between two types of prediction: probabilistic and theory-based. Probabilistic predictions are based on universal laws of probability. If one tosses a coin in the air often enough, he can predict that the coin will land heads up fifty percent of the time. Theory-based predictions are grounded in propositions and deduced hypotheses. Unlike probabilistic predictions, they must be made within a delineated period of time or number of cases. On the other hand, probabilistic predictions are not based on the explanatory power of propositions and derived hypotheses, but rather on the assumption that given enough time or enough cases, all outcomes may be probabilistically predicted. Theory-based predictions are contingent upon hypotheses and propositions (Kaplan, 1964). That is, a theory-based prediction is based on the specification of a relationship between particular antecedents and consequences.

In social science, the distinction between

probabilistic and theory-based prediction emerges as a sampling issue. Given a large enough sample, and/or a long enough period of observation, theorists can predict on the basis of some of the worst explanations or no explanations at all. In other words, given a large enough sample and/or a long enough period of observation, one is able to predict for all the wrong reasons. For example, one might predict that managerial practices based upon the idea of indentured servitude will result in higher levels of job satisfaction than managerial practices based upon the idea of participative decision making.

Thus, the predictive adequacy of a theoretical system must be judged in terms of its ability to make predictions within delineated spaces and time. This goes back to the earlier discussion of theory, specifically to the assumptions which a theorist must make about space and time. From these assumptions the final and most widely accepted evaluative indicator may be stated: The theory should provide a mechanism for predicting beyond chance.

However, a note of caution must be raised. As Hawking maintained, although a theory is always provisional (i.e., it can never be proven), the predictive adequacy of two alternative theories may be comparatively assessed on the basis of the degree of confidence researchers have in the theory (i.e., statistical significance). That is, assuming that Hawking was correct in stating that "no matter how many times the results of an experiment agree with some theory, you can never be sure that the next time the result will not contradict the theory," the key to predictive adequacy is minimizing the probability of disconfirmation (Hawking, 1988, p. 10).

As shown in Figure 3, all these criteria present a multidimensional approach to the critical analysis and evaluation of theory. A theory cannot be deemed acceptable by meeting only one or two ideal criteria. Rather, in the context of such a framework, a theory may be found acceptable in one respect yet unacceptable in another. Only that rare theory which meets all of the evaluative indicators may be considered (at a given point in time) to be acceptable and gen-

|               | Falsifiability  | Utility  |
|---------------|---|--|
| Variables     | Operationally Defined?<br>Measurement Issues<br>face & content validity<br>noncontinuousness<br>reliability                                   | Variable Scope   |
| Constructs    | Clarity & Parsimony<br>Construct Validity<br>convergent<br>discriminant   | Construct Scope  |
| Relationships | Logical Adequacy<br>nontautological<br>specified nature of<br>relationship<br><br>Empirical Adequacy<br>more than one object<br>or time frame | Explanatory Potential<br>specificity of assumptions<br>regarding objects<br>specificity of assumptions<br>regarding relations<br>scope and parsimony of<br>propositions<br>Predictive Adequacy<br>probabilistic versus<br>theory-based |

**Figure 3. Criteria for evaluating theories.**

erally superior to the alternative theories against which it has been compared.

### **Evaluating the Conceptual Coherence of a Theory: The Fit**

Previously, the need for organizational theorists to be in a position to discard some previous theories if the field is to avoid sinking under its own weight was mentioned. The other side of that coin is the need for a cumulative body of more-or-less universally accepted theories of organizations as a basis for further theory construction. If theorists are to begin that monumental task, how will they decide which theories to include?

To answer this question, theorists need a clear understanding of how a given theory fits in with the other preexisting and apparently related theories. Two qualitative dimensions to this fit can be described: connective and transformational. *Connectivity* refers to the ability of a new theory to bridge the gap between two or more different theories, thus explaining something between the domains of previous theories. In this way, new knowledge is created, and a more nearly continuous mapping of the empirical universe is achieved. As an example, institutional theories of bargaining (Chamberlain & Kuhn, 1965) were connected to economic theories of bargaining (Pen, 1952) by the more recent social-psychological work by Axelrod (1980) and Bacharach and Lawler (1980). The social-psychological approach does not subsume the two, but instead incorporates selected elements from each. Tactics and coalitions are adopted from the institutionalists, while the motivational elements (e.g., ends maximization) are adopted from the economists.

The theory is said to be *transformational* if it causes preexisting theories to be reevaluated in a new light. Some theories even have the potential to change the older, established theories that they were built upon, just as the placement of a bridge has a profound effect upon the

economic development of the communities along the two shores connected. As Kaplan (1964) said, "A new theory requires its own terms and generates its own laws: the old concepts are not merely reorganized, but reconstituted, the old laws not just connected, but given a new meaning" (p. 297).

How does this happen? Theories are collections of constructs which are related to each other by propositions. Thus, the boundary spanners between theories are the constructs embedded in and shared by them. Yet this notion of boundary spanners is still problematic, because constructs are still theoretical terms which are not directly observable. In the empirical world, these constructs are operationalized as variables and related to each other by hypotheses. Kaplan pointed out that a hypothesis may be as much confirmed by fitting it into other theories as by fitting it to the facts, because the theory then enjoys the conceptual coherence and support provided by evidence for all other related theories. Just as the superstructure of a bridge relies upon each supporting piece of metal and each individual rivet and weld connecting them, the constructs and propositions of a theory rely upon the many hypotheses and variables generated from them.

This systemic quality is what makes the construct validation process so problematic; establishing the content of a construct in a variable (inevitably reducing scope) ends up interfering with the necessity of that construct being used to connect the new theory with other theories (which requires broad scope). This is further revealed when theorists try to find a list of qualities which make a conceptually coherent theory. Systemic openness, dynamic openness, flexibility, and vagueness are cited as being good for creativity. Thus, they give theories utility in that they have the scope to be put to work in the broader intellectual world of the social sciences. On the other hand, qualities such as closure, precision, accuracy, and the exactness of meaning of terms are cited as being necessary if a theory is to be capable of disconfirmation. In



other words, the latter list of qualities is essential for a theory's falsifiability.

A quick look at Figure 2 reveals that this tension breaks along the line between the criteria of falsifiability and utility. Interestingly enough, it seems to break in all three boxes (variables, constructs, and relationships). No matter how detailed the analysis, the same issue comes back. However, an increasing sensitivity to the problem should cause theorists to be more rigorous in attempting to think through all the issues when finding the balance between the competing forces of focus.

## Conclusion

It would be foolish to assume that on the basis of any set of criteria, one could determine that the insights of Marx are more or less profound than those of Weber. However, as the reading of any organizational journal will testify, most of us do not theorize on the level of Marx or Weber. To a large degree today's students of organizational behavior are craftspersons working in the context of the middle range (Merton, 1957). As such, the goal is to ensure that theoretical systems and statements can be empirically tested, and provide some source of explanation and prediction.

The use of the criteria should improve theory building and evaluation by

1. Ensuring the delineation of theoretical boundaries, while at the same time ensuring the explication of assumptions (values, scope, and time) which bound the theory.
2. Ensuring a common language of constructs and variables across levels.
3. Specifying the distinctions between propositions and hypotheses, and the relationships implied in them.
4. Improving the parsimony of our theories.

Organization and management studies are not the only delinquents in regard to strict evaluation of theory. Indeed, this has been a com-

mon symptom throughout the social sciences. Maybe it is more apparent because this discipline is so diverse and the methodologies are so divergent. It also may be that the spate of publicity surrounding such popular books as *In Search of Excellence*, although on the one hand strengthening the field, has resulted in a rash of broad descriptions and sometimes irrefutable theories.

If theorists don't take the rules of theory seriously, individually they will continue to cling to theories in almost cultist fashion. Getting beyond this clinging behavior, which tends to drive theorists from fad to foible, demands a precise discourse, one which allows theorists to focus on the specific strengths and weaknesses of particular theories. If nothing else, the list of criteria presented here may enhance the accuracy of discourse.

Finally, the use of criteria for the evaluation of theory also may assist in demystifying certain false dichotomies. Although these criteria could be dismissed as just another rhetorical call for more organized empiricism, nothing about the appropriateness of various modes of data collection is implied. Too often, a false distinction is drawn between those who manipulate large data sets and those who rely on more interpretive methodologies such as ethnographic or case studies. To the degree that both are dealing with the empirical world, the principles herein are equally applicable to quantitative data and ethnographic or case studies. No matter how the data are collected, researchers have the obligation to present them in a way that allows other scholars a fair chance at using and or disproving the data.

The second false dichotomy is between theory construction and theory testing. The message to the theorist should be clear. If it is not testable, no matter how profound or aesthetically pleasing it may be, it is not a theory.

The third false dichotomy is between those who view themselves as theoreticians and researchers and those who view themselves as

consultants and practitioners. The role of consultants is to assist their clients in both diminishing the complexity of their empirical world and explaining and predicting events. The goal of theory is to diminish the complexity of the empirical world on the basis of explanations and predictions. Thus, both practitioners and theoreticians need a clearer understanding of the rules of theory construction.

Nevertheless, this article should end on a cautionary note. To dangle criteria above the head of a theorist like the Sword of Damocles may stifle creativity. In most of our work, flaws in theoretical logic can be found. However, during the early stages of theory building, there may be a fine line between satisfying the criteria of the internal logic of theory and achieving a creative contribution. A good theorist walks this line carefully.

## References

- Aldrich, H. (1979) *Organizations and environments*. Englewood Cliffs, NJ: Prentice Hall.
- Anderson, J. C. & Gerbing, D. W. (1988) Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411–423.
- Argyris, C., & Schon, D. A. (1977) *Theory in practice: Increasing professional effectiveness*. San Francisco: Jossey-Bass.
- Axelrod (1984) *The evolution of cooperation*. New York: Basic Books.
- Bacharach, S. B., & Lawler, E. (1979) *Bargaining*. San Francisco: Jossey-Bass.
- Bacharach, S. B., & Lawler, E. (1980) *Power and politics in organizations*. San Francisco: Jossey-Bass.
- Bailey, K. D. (1970) Evaluating axiomatic theories. In E. F. Borgatta & G. W. Bohrnstedt (Eds.), *Sociological methodology* (pp. 48–71). San Francisco: Jossey-Bass.
- Bamberger, P. (in press) Re-inventing innovation theory: Critical issues in the conceptualization, measurement and analysis of technological innovation. In S. Bacharach (Ed.), *Research in the sociology of organizations*, Vol. 8. Greenwich, CT: JAI Press.
- Beehr, T., & Newman, J. (1978) Job stress, employee health, and organizational effectiveness: A facet analysis model and literature review. *Personnel Psychology*, 31, 665–699.
- Bierstedt, R. (1959) Nominal and real definitions in sociological theory. In L. Gross (Ed.), *Symposium on sociological theory* (pp. 121–144). New York: Harper & Row.
- Blalock, H. M. (1969) *Theory construction: From verbal to mathematical formulations*. New York: Prentice-Hall.
- Blalock, H. M. (1968) The measurement problem: A gap between the languages of theory and research. In H. M. Blalock, Jr., & A. B. Blalock (Eds.), *Methodology in social research* (pp. 5–27). New York: McGraw-Hill.
- Blau, P., & Scott, R. (1962) *Formal organizations*. San Francisco: Chandler.
- Brodbeck, M. (1959) Models, meaning, and theories. In L. Gross (Ed.), *Symposium on sociological theory* (pp. 373–403). New York: Harper & Row.
- Campbell, D., & Fiske, D. (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Chamberlain, N. W., & Kuhn, J. W. (1965) *Collective bargaining* (2nd ed). New York: McGraw-Hill.
- Child, J. (1972) Organizational structure, environment, and performance: The role of strategic choice. *Sociology*, 6(1), 1–22.
- Chronbach, L. J., & Meehl, P. E. (1955) Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Cohen, B. (1980) *Developing sociological knowledge: Theory and method*. Englewood Cliffs, NJ: Prentice Hall.
- Cohen, M., March, J., & Olsen, J. (1972) A garbage can model of organizational choice. *Administrative Science Quarterly*, 17(1), 1–25.
- Deal, T. E., & Kennedy, A. A. (1982) *Corporate cultures*. Reading, MA: Addison-Wesley.
- Dubin, R. (1969) *Theory building*. New York: Free Press.
- Dubin, R. (1976) Theory building in applied areas. In Marvin D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 17–40). Chicago: Rand McNally.
- Durkheim, E. (1933) *The division of labor in society*. New York: MacMillan.
- Dyer, L., & Holder, G. W. (1989) Toward a strategic perspective of human resource management. In L. Dyer (Ed.), *Human resource management: Evolving roles and responsibilities* (pp. 1–46). Washington, DC: Bureau of National Affairs.
- Emerson, R. M. (1962) Power-dependence relations. *American Sociological Review*, 27, 31–40.
- Etzioni, A. (1975) *A comparative analysis of complex organizations*. New York: Free Press.

- Freeman, J., & Hannan, M. (1983) Niche width and the dynamics of organizational populations. *American Journal of Sociology*, 88, 1116–1145.
- Gay, P. (1988) *Freud: A life for our times*. New York: Norton.
- Gibbs, J. (1972) *Sociological theory construction*. Hinsdale, IL: Dryden.
- Glaser, B., & Strauss, A. (1967) *The discovery of grounded theory*. New York: Aldine.
- Gouldner, A. W. (1954) *Patterns of industrial bureaucracy*. New York: Free Press.
- Hage, J., & Aiken, M. (1969) Routine technology, social structure, and organizational goals. *Administrative Science Quarterly*, 14(3), 366–377.
- Hage, J., & Meeker, B. (1987) *Social causality*. Boston: Unwin Hyman.
- Hall, C. S., & Lindzey, G. (1957) *Theories of personality*. New York: Wiley.
- Hannan, M., & Freeman, J. (1977) The population ecology of organizations. *American Journal of Sociology*, 82, 929–964.
- Hannan, M., & Freeman, J. (1984) Structural inertia and organizational change. *American Sociological Review*, 49, 149–164.
- Hawking, S. (1988) *A brief history of time: From the big bang to black holes*. New York: Bantam.
- Heise, D. R., & Bohrnstedt, G. W. (1970) Validity, invalidity, and reliability. In E. F. Borgatta and G. W. Bohrnstedt (Eds.), *Sociological methodology* (pp. 104–129). San Francisco: Jossey-Bass.
- Hempel, C. (1965) *Aspects of scientific explanation*. New York: Free Press.
- Hempel, C. G., & Oppenheim, P. (1948) The logic of explanation. *Philosophy of Science*, 15, 135–175.
- Hrebiniak, L., & Joyce, W. (1985) Organizational adaptation: Strategic choice and environmental determinism. *Administrative Science Quarterly*, 30, 336–349.
- Kaplan, A. (1964) *The conduct of inquiry*. San Francisco: Chandler.
- Kemery, E., Bedeian, A., Mossholder, K., & Touliatos, J. (1985) Outcomes of role stress: A multi-sample constructive replication. *Academy of Management Journal*, 28(2), 363–375.
- Kemery, E., Mossholder, K., & Bedeian, A. (1987) Role stress, physical symptomology, and turnover intentions: A causal analysis of three alternative specifications. *Journal of Occupational Behaviour*, 8, 11–23.
- Kerlinger, F. N. (1973) *Foundations of behavioral research* (2nd ed.). New York: Holt, Rinehart & Winston.
- Kimberly, J. R., Miles, R. H., & Associates. (1980) *The organizational life cycle*. San Francisco: Jossey-Bass.
- Kunda, G. (in press) *Engineering culture*. Philadelphia: Temple University Press.
- Lawrence, P., & Lorsch, J. (1967) *Organization and environment*. Cambridge, MA: Harvard University Press.
- Lipset, S., Trow, M., & Coleman, J. (1956) *Union democracy*. New York: Free Press.
- Locke, E. (1976) The nature and causes of job satisfaction. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 1297–1350). Chicago: Rand McNally.
- Locke, E., Saari, L., Shaw, K., & Latham, G. (1981) Goal setting and task performance. *Psychological Bulletin*, 90(1), 125–152.
- Malinowski, B. (1962) *Sex, culture, and myth*. New York: Harcourt, Brace & World.
- Merton, R. K. (1957) *Social theory and social structure*. New York: Free Press.
- Miller, D. (1986) Configurations of strategy and structure. *Strategic Management Journal*, 7, 233–250.
- Miller, D. (1987) The genesis of configuration. *Academy of Management Review*, 12, 686–701.
- Miller, D., & Friesen, P. H. (1977) Strategy making in context: Ten empirical archetypes. *Journal of Management Studies*, 14, 259–280.
- Miller, D., & Friesen, P. H. (1984) *Organizations: A quantum view*. Englewood Cliffs, NJ: Prentice-Hall.
- Mills, C. Wright (1956) *The power elite*. New York: Oxford University Press.
- Mintzberg, A. (1970) *The iron cage: An historical interpretation of Max Weber*. New York: Knopf.
- Nagel, E. (1961) *The structure of science: Problems in the logic of scientific explanation*. New York: Harcourt, Brace and World.
- Newton, T. & Keenan, A. (1987) Role stress reexamined: An investigation of role stress predictors. *Organizational Behavior and Human Decision Processes*, 40, 346–368.
- Parsons, T. (1962) *The social system*. Glencoe, IL: Free Press.
- Pelz, D. C., & Andrews, F. M. (1966) *Scientists in organizations*. New York: John Wiley.
- Pen, J. (1952) A general theory of bargaining. *American Economic Review*, 42, 24–42.
- Perrow, C. (1986) *Complex organizations: A critical essay* (3rd ed.). New York: Random House.
- Popper, K. (1959) *The logic of scientific discovery*. New York: Harper & Row.

- Radcliffe-Brown, A. R. (1949) Functionalism: A protest. *The American Anthropologist*, 51, 320–323.
- Schein, E. H. (1985) *Organizational culture and leadership*. San Francisco: Jossey-Bass.
- Schuler, R. (1982) An integrative transactional process model of stress in organizations. *Journal of Occupational Behavior*, 3, 5–19.
- Schuler, R. S., & Jackson, S. E. (1987) Organizational strategy and organizational level as determinants of human resource management practices. *Human Resource Planning*, 10(3), 125–141.
- Schwab, D. P. (1980) Construct validity in organizational behavior. In B. M. Staw & L. L. Cummings (Eds.), *Research in Organizational Behavior* (Vol. 2, pp. 3–43). Greenwich, CT: JAI Press.
- Shenhav, Y., & Haberfeld, Y. (in press) The various faces of scientific productivity: A contingency analysis. *Quality and Quantity*, 22.
- Shenhav, Y., Haberfeld, Y., & Cohen, B. (in press) Contextual analysis of team productivity in the R & D industry. *Scientometrics*.
- Shils, E. A., & Finch, H. A. (1949) *Max Weber on the methodology of the social sciences*. Glencoe, IL: Free Press.
- Stinchcombe, A. L. (1968) *Constructing social theories*. New York: Harcourt, Brace and World.
- Thompson, J. (1967) *Organizations in action*. New York: McGraw-Hill.
- Tolbert, P. (1985) Institutional environments and resource dependence: Sources of administrative structure in institutions of higher education. *Administrative Science Quarterly*, 30, 1–13.
- Van de Ven, A. (1979) Review of Howard Aldrich, *Organizations and environments*. *Administrative Science Quarterly*, 24(2), 320–326.
- Van de Ven, A., & Chu, Y. (1987) A psychometric assessment of the Minnesota innovation survey. *Technical Research Report*. Minneapolis, MN: The Strategic Management Research Center, University of Minnesota.
- Vecchio, R. P. (1987) Situational leadership theory: An examination of a prescriptive theory. *Journal of Applied Psychology*, 72(3), 444–451.
- Weber, M. (1947) *The theory of social and economic organizations*. New York: Oxford University Press.
- Weick, K. (1976) Educational organizations as loosely coupled systems. *Administrative Science Quarterly* 21, 1–19.
- Werkmeister, W. H. (1959) Theory construction and the problem of objectivity. In L. Gross (Ed.), *Symposium on sociological theory* (pp. 483–508). New York: Harper & Row.
- Whyte, W. (1955) *Street corner society*. Chicago: University of Chicago Press.
- Wils, T., & Dyer, L. (August 1984) Relating business strategy to human resource strategy: Some preliminary evidence. Paper presented at the 44th annual meeting of the Academy of Management, Boston.
- Young, R. (1988) Is population ecology a useful paradigm for the study of organizations? *American Journal of Sociology*, 94(1), 1–24.

*I thank Peter Bamberger and Bryan Mundell, whose patience, research, and insights made this paper possible; Andy Van de Ven, who proved to be an exceptional colleague by making many insightful comments throughout this paper; Joe Francis, William Sonnenstuhl, and Lawrence Williams, whose suggestions have been incorporated throughout this paper; and Lee Meiser, who is truly a team player. Finally, I thank Anna Mundell.*