PROBLEMS IN SYSTEM THEORETIC EMPIRICAL RESEARCH

Mark H. Bickhard

James R. Murray

University of Chicago

Paper presented as part of the

SYMPOSIUM ON SYSTEMS ANALYSIS AT THE

AMERICAN SOCIOLOGICAL ASSOCIATION MEETINGS

Denver, 1971

(Sponsored Jointly with the Society for General Systems Research)

PROBLEMS IN SYSTEM THEORETIC EMPIRICAL RESEARCH*

Mark H. Bickhard
James R. Murray

University of Chicago

Although the use of a systems perspective in sociological thinking has been growing rapidly in recent years, its use in empirical research has remained relatively limited. This lag in empirical applications of systems theory is paralleled by, and perhaps explained by, a relative lack of attention in sociological literature to what systems theoretic research looks like and to its special problems of design and analysis. We would like to address ourselves to a few of the issues involved in systems theoretic empirical research.

In particular, among the many issues effected by a systems perspective, we will be concerned with a characterization of and methodological rationale for systems research as differentiated from classical modes of research, and with some of the general logic and statistical procedures which apply to the testing of systems models. We conclude with an example of a systems theoretic statistical model from ongoing research in organizational behavior.

Paper presented as part of the Symposium on Systems Analysis at the American Sociological Association Meetings, Denver, 1971 (Sponsored jointly with the Society for General Systems Research).

A RATIONALE FOR SYSTEMS THEORETIC EMPIRICAL RESEARCH

It might be contended, and appears to be frequently assumed on at least an implicit level, that, however useful systems theory may be in theory construction and model building, it has no real impact on empirical research-“that systems models are stated, data is collected, and analyses are performed in the same ways and for the same reasons as with any other kind of model. The main point of this section is that systems research is different from classical research--in particular, different from prediction analysis and experimental analysis--and that there is a powerful methodological rationale for engaging in such research. After some contextual preliminaries, the argument runs roughly as follows:

1) Causal models are legitimately the ultimate goals of the scientific process;
2) Experimental analysis is unfortunately taken as the ultimate methodological foundation for the imputation of causality; and
3) Structural or systems analysis forms an independent and equally sufficient basis for causal imputation.

The section ends with a few historical precedents for the above position.

In the literature of mathematical systems theory, the most sophisticated area of general systems theory, we find almost exclusive concern with two general problems --the analysis of the properties of a particular system design, and the design of a system with particular properties (Klir, 1969). These problems are encountered frequently in such cases as the design or analysis of electronic circuit diagrams, computer programs, or oil refineries. Sociological examples do perhaps exist--e.g., the analysis of the dynamic properties of a particular urban model (Forrester, 1969), or the design of a complex organization with certain functional properties--but essentially these problems are problems of application and require a greater knowledge of the relevant systems and their components than is generally to be found in the behavioral sciences.

More commonly in the behavioral sciences, we encounter a real system about which we know little or nothing concerning either its structure or its behavior, and we wish to understand more about both. The logically first task when confronted with an unknown system like this is one of system definition. Systems definable by their physical location or physical structure seldom present definitional problems, but when the essential system characteristics involve less obvious dimensions, such as in the case of families or cultures, rigorous definition can encounter difficulties ranging into the philosophical.

Although the problem of definition is sometimes 'solved[5] through the use of partial descriptions, the description of a designated system is not accurately to be considered a definition of that system, but rather the beginning of and foundation for the process of analyzing that system once defined. The process of description consists of the development of increasingly precise and replicable procedures for the assignment of descriptors, the organization of those descriptors into attribute classes--generally on the basis of procedural similarities--and the increase in knowledge of the logical structure of these attribute classes. A description procedure which meets certain logical requirements is called a measurement process; a symbolic representation of the outcome of a measurement process is called a variable; and a particular outcome of such a process is simply a measurement. We think of a measurement process as a description procedure which precisely and replicably generates an attribute class, and whose attribute class is provably exhaustive and possesses a generative logical definition--sometimes, however, the term "measurement" is reserved for processes whose attribute classes have in addition certain specified structures, such as an ordering or a field.

As various descriptive (or measurement) procedures become developed around or applied to designated systems, constraints will begin to be observed and sought after in the complex descriptor space generated by the attribute classes. Constraints may exist along a temporal dimension, a spatial dimension, or any dimension definable within the attribute space (i.e., with respect to any attributes of the space). Observed constraints may be expressed probabilistically, in terms of computing algorithms, by state transition laws, or in terms of any other appropriate mathematical language. Sequential constraints along a time dimension represent the classical case of laws of prediction.
The more general search for arbitrary constraints along arbitrary dimensions might also be considered to have prediction as its ultimate goal, but we choose to designate the process as signal analysis--considering the measured attribute values of the system as signals generated by the system.

Among the constraints in a system's attribute space may be some that we call causal. Causal constraints are characterized by a certain logical form and by various epistemological properties which we require of and ascribe to them (Bunge, 1959, Bunge, 1968). There is some controversy over causality as a metaphysical property, and an accompanying controversy over its usefulness in science. We feel that causal claims are of maximal usefulness in science, not because of any convictions regarding causality's metaphysical nature or existence, but rather because of some convictions regarding the informational properties of causal statements. In particular, the causal statement appears to us to be the maximally informational form of statement of empirical constraint among all possible such forms--maximally informational in Popper's sense of falsifiability (Popper, 1959). That is to say, statements of causality are more falsifiable than any other form of statement of empirical constraint. This is not to claim that all possible falsifications of a causal claim ever will or can be tested--in this sense a causal statement remains always a hypothesis, but so also does any other empirical induction--but rather that it is the strongest form of empirical induction possible. Such a view appears to us to be of importance in a systems theoretic context because, among the several logical and epistemological characteristics which we ascribe to a constraint in calling it causal, are two which we feel have special significance for system theoretic research--control and structure.*

Perhaps the characteristic most strongly associated with causality in science is that of control--if we know what causes something, then we can, at least in principle, control it. Around this characteristic has developed the vast body of epistemological methodology known as experimental design and analysis. This methodology is constructed precisely in order to isolate the experimentor's arbitrary (randomized) acts of control from as many other known or potential constraints on the outcome as possible, in order to be as certain as possible of the controlling effects of those acts. Mathematical statements of the constraints established by experimental analysis may have complex mathematical structure, but they will have only the simplest of topological structures--some set of variables which constraint (control) some other set of variables. Statements of such constraints are called causal laws.

Although control is certainly one of the characteristics imputed to causal constraints, it has in addition become a sort of methodological sine qua non for imputing causality to a potentially causal constraint--and this has created great difficulties in those sciences, particularly the behavioral sciences, where experimental manipulation is frequently not practical or ethical. There is, however, at least one additional epistemological foundation, with accompanying historical precedents, for the imputation of causality which we feel has been given insufficient attention, and which bears directly on the rationale and need for systems theoretic empirical research.

---

* Actually, the characteristics are 'necessary production' and 'instantaneousness' -- control and structure are resultant epistemological foundations for causality. This distinction will not be consistently maintained in the text.

This foundation--causal structure--derives from another of the basic characteristics attributed to causality--the characteristic of instantaneousness (continuity if temporal continuity is accepted). The causal constraint between two events is presumed to be mediated by a temporally continuous sequence of events. When the nature of the original constraint is investigated, this will be done by exploring the events between the original events in time, and attempting to explicate the causal relationships that are manifested in this mediating sequence. Although the mediating events are by definition sequential, the mediating relations may exhibit various topological structures: they may proceed in a single sequence from the first of the original events to the second, they may branch or conditionally branch into tree-like structures, they may branch and turn back on themselves, forming causal loops, and they may exhibit complex versions and combinations of these. Investigating the mediating causal constraints is thus investigating the mediating causal structure.

Note that in investigating the causal structure that mediates the constraint between two external variables of an unknown system, we are in fact investigating the internal causal structure of the system--we are engaged in what is called black box analysis.

The point to be made now is two-fold: 1) that such structural analysis of a system is both procedurally and informationally independent of a control or experimental analysis of that system, and 2) that a structural analysis of a system can serve as a legitimate foundation for the imputation of causality to the modeled constraints. The procedure independence is easily established: structure is a general characteristic among variables. It may exist in space, in time, or with respect to any other set of attributes. Clearly, then, structure is not necessarily even causal, and may be discerned with or without experimental manipulations.

A structure discerned without control manipulations will of course be weaker in its claims of control relationships (which it does as part of its claim of causality) than one using control manipulations, but the import of the point above is not that structural analysis is the equivalent of experimental analysis, but rather that structural analysis is independent of experimental analysis, and forms an additional basis for the imputation of causality that is at least as legitimate as the basis of control.

Accepting that structural analyses and control analyses are procedurally independent, we must now argue their informational independence--that is, we must argue that each kind of analysis contributes information not available in the other. That control analysis adds something to structural analysis is trivial: a direct test of the characteristic of control which is imputed by causality. In a sense it is also trivial that structural analysis adds something to control analysis--structure. But to what advantage? Just as we originally referred to a justification for concern with causal constraints on the basis of their greater information with respect to general signal analysis constraints, so here we claim that causal structural models offer greater information than causal laws--but, in this case, we would like to at least outline the argument.

A simple control causal constraint, or causal law, is established with a simple induction--many times under many conditions "B" has occurred when I did "A," therefore I will generalize to the relationship " 'A' is a cause (or the cause) of 'B'." Useful as such generalizations are (more so than " 'A' predicts 'B' "), they tell us nothing about what will happen if we do "C" instead of "A." A particular structural model may or may not tell us something about "C," but all non-vacuous structural models manifest more than one resultant causal constraint (one of which may involve "C"), and non-causal constraints too. It is precisely in the richness of unobserved constraints and in the parsimony of expression of observed constraints that a causal model contains more information than the (set of) simple causal laws for which it was developed--more information again in Popper's sense of falsifiability (Popper, 1959).

Thus we argue for causal laws over non-causal laws for their greater falsifiability, and for causal models over causal laws for the same reason.

To this point we have argued that structural analysis is both procedurally and informationally independent of control analysis--the clear conclusion is that even if we have experimental data, we should attempt structural analyses as well (and vice versa). But, accepting their full independence, what is the justification for imputing causality solely on the basis of structural analysis? This is, of course, logically a purely arbitrary question --we can (and do) impute causality to simple signal analysis constraints without having evidence for either structure or control. We can, then, impute causality to any constraint that meets the basic requirements of logical form--the question, then, focuses on the issue of justification for such imputation. Is evidence of control necessary to justify causal imputation, or are control evidence and structural evidence similarly sufficient for causal imputation? At this point, the arbitrariness of choice overtakes the argument, and we can only state that structural analysis appears to us to be in principal sufficient for causal imputation equally with control or experimental analysis, and cite some reasons and precedents for that choice.

Note that our argument thus far has implicitly assumed the sufficiency, at least, of experimental analysis for the imputation of causality--this assumption has been made precisely because of the prevalence of the experimental method as the methodological sine qua non of causality. If we look at what we actually do with experimental evidence, however, we find that we are never really confident that our experimental controls, randomizations, etc., have ruled out potential ’true[1] causes alternative to the one we’re testing until we have an accepted structural model which explains to us the control relationship we’ve been investigating. That is, we look for a causal model underlying (mediating) the apparent causal law(s). Ultimately, then, both control and structural analysis are asked of a causal claim, but it is clear that imputation of causality is frequently accepted on the basis of experimental analysis alone, and it is also true, as we shall show in a moment, that structural analysis alone is also accepted as a basis for causal imputation.

Perhaps the strongest evidence of the legitimacy of structural claims to causality is Newton’s law of gravitation: no experimental test of this law was technologically possible until long after it had been generally accepted throughout the scientific community--and this initial acceptance was based precisely on the law’s structural (or systemic, if you will) explanations of the orbit of the moon, Kepler’s law of planetary motion, the tides, etc. Einstein's theories show a similar history, though in this case the lag before experimental tests became possible was much shorter. Note that the basis of acceptance of these models was their general success in stating old and predicting new signal analysis constraints of all kinds--e. g., the progression of the perihelion of mercury, and the ’bending’ of light by the sun--and was not restricted to their success in stating and predicting control analysis constraints. Areas in which initially structural models abound are astronomy, meteorology, ecology, geology, economics, and so on. It seems clear that wherever control is difficult, including in particular most of sociology and much of psychology, at least the initial claims for causal models in those areas will in general have to be based on structural success, not control success--in many cases, in fact, no control or experimental manipulation will even be attempted until a minimal confidence is reached in the available structural models, e.g., meteorology.

The claim, then, is that causal models are the ultimate goals of science, that ultimately both experimental and structural verification will be required of a causal model, that either structural or experimental analysis serves as a sufficient methodological foundation for the initial imputation of causality, and that different of these methodological approaches will be most feasible and appropriate for different subject areas. The intended message is that structural analysis has a much larger role to play in sociological research than it has evidenced to date.

LOGIC AND PROCEDURES OF STRUCTURAL MODEL TESTING

But, if the usefulness of structural analysis is accepted, what then does it look like, what are its procedures and special problems? These questions cannot be answered here with any detail, but we would like to outline some of the answers and discuss examples of some others.

Our discussion begins with one of the considerations involved in the mathematical statement of a systems model--the mathematical language in which it is to be stated. A precise statement is a necessary preliminary to any empirical test, and many particulars of the test procedures will depend on the language and the form in which the model is stated. It seems that little attention has been given to the problem of selecting a mathematical language most appropriate to the model one is attempting to express.

Given a model statement, the next task is one of empirical comparison. Skipping in this paper the issues of data collection, we attempt to organize the numerous strategies of empirical model testing through a discussion of the general logic of model testing. Differences in testing strategies appear as adjustments to differences in available information at various points in the general logic of empirical test. The discussion proceeds from maximal information conditions to minimal information conditions.

The section ends with a discussion of one of the most common statistical criteria and its procedures--least squares analysis. We discuss assumptions of classical least squares procedures which are violated in general in structural analyses, and outline modifications and innovations in these procedures which have been made to accommodate such violations. There is also some discussion of how these assumptions related to prediction and control analyses.

The Mathematical Statement of a Model

Passing by the general problem of developing a systems model, we would like to first address some of the issues concerned in the mathematical statement of a model. Much of the discussion in this section applies similarly to both causal and logical structural models, e.g., the logical structure of an attribute class. However, although the mathematical forms for logical and causal models may frequently be similar, they retain many differences of significance for later sections--e. g. , considerations of measurement or classification error are usually different, dynamic models always raise the issue of longitudinal versus cross sectional data sampling while non-temporal models do not, etc.

The variables involved in a systems model may be conceived of as being continuous or discrete; the time transitions (if any) in a model may likewise be conceived of as continuous or discrete; and the constraints in a model may be conceived of as being deterministic or probabilistic. These dichotomies alone induce an eight-celled classification of models, each placing its own special demands on the mathematical language needed to state them. Mathematical languages which meet these various demands are not equally well-developed: as an example, deterministic constraints on continuous transitions among continuous variables are frequently modeled by systems of differential equations--a well-developed area of mathematics, while probabilistic constraints under the same circumstances might be modeled by stochastic differential equations--a very poorly developed area. Theories of stochastic processes, such as markov theory or martingale theory, can express many probabilistic models, particularly those involving discrete variables, but at the cost of restrictions on the model assumptions which may or may not be acceptable to the model builder.

The consequence, unfortunately, is that some modelling needs are met fairly well, while others are met very poorly or not at all. Also, some mathematical languages are applicable to more than one kind of model, while others apply only to very restricted cases within, say, the above classification of model types, while still others, e.g. , the theory of computation, or graph and network theory, are not obviously assignable to any of the above categories.

Although it was not intended to be exhaustive, the above model classification system was nevertheless generated by dichotomies which were intended to be meaningful from a modeling perspective. The difficulty of these model dichotomies to make clear and interpretable distinctions among the mathematical languages available for the statement of conceptual models illustrates part of the difficulty of moving from a conceptual model to its mathematical statement--the process is far from being a well-defined one, and, in fact, represents another area of relative neglect in the general methodological literature. What happens in practice, of course, is that we constrain and distort our models to fit the mathematics with which we are familiar. It would seem to be a major task of a useful general systems theory to construct classifications of model types and modelling languages that bore some meaningful relationship to each other and could provide some guidance for the process of stating a conceptual model mathematically.

The most general mathematical language is category theory (MacLane and Birkhoff, 1967, Mitchell, 1965). All other mathematical languages can be embedded within it, and it seems reasonable, therefore, that systems theoretic applications of category theory may provide the framework within which such tasks as the above (and many others) might best be pursued. Category theory, however, is relatively young, and its system applications even younger; nevertheless, even preliminary developments within category theoretic systems theory have clarified a number of conceptual issues and have begun to show instances of immediate application (Goguen, 1970). Category theory shows promise of being the language par excellence for general systems theory.

The Logic of Model Testing

Whatever language is used to state a model, it must eventually be tested, and to do this it must be compared to empirical data. Considerations of empirical data collection will effect a conceptual model in many ways. A conceptual model, for example, may contain continuous variables, while the empirical operationalizations may constrain these to be discrete variables; or we may have continuous transitions in the conceptual model, which become discrete transitions in the empirical model because we cannot collect data continuously. The transition from a conceptual model to a corresponding empirical model, however, is not generally very profound (although it may effect the choice of mathematical languages for stating the model). The significant effect of data consideration is the unavoidability of error in data, and the resulting necessity of moving from an empirical model to a corresponding statistical model for the purpose of empirical test. We will neglect in this paper most of the problems of the logic and process of data collection, and move directly to a discussion of the logic and process of data analysis.

A test of a model results ultimately in a decision to accept or reject the model as a basis for further action--conceptualization, research, or application. If the model is fully specified, the task of decision may be approached directly. More commonly, however, a model is specified only up to some set of arbitrary constants, or parameters, and some fully specified version of the model must be selected before the decision procedure can begin --i.e., specific values must be selected for the parameters.

Since the selected version of the general model serves to represent the general model for the purpose of decision, we wish to select the 'best' version of the model for that representation--that is, we wish to test a version of the model which represents its limit of usefulness, not merely a random example. The problem of specifying the model thus has two parts: 1) defining a rigorous measure of 'goodness,' or usefulness, of specific models with respect to which different model versions can be compared and 2) finding a procedure by which the 'best' version with respect to this measure can be found.

In general, measures of model 'goodness' (usefulness) may be defined for many purposes (e.g., cost of construction), but the object and purpose of empirical models is to express empirical constraints: we ask for agreement or fit between the constraints expressed in the model and constraints observed in reality. Considerations of measurement error, unconsidered variables, etc. , make it clear, however, that such agreement can never be complete--that we can ask for 'fit' between a model and relevant data, but only relatively, not absolutely. The measure of 'goodness' which is needed for empirical tests of models is thus a measure of empirical 'fit': we wish to select the 'best-fitting' version of a model, and our final decision will be whether or not the model 'fits' 'well- enough' according to some criterion.

The most common measure of empirical fit is the 'sum of squares' measure. This measure constructs an 'error vector' out of the discrepancies between the model constraints and the observed constraints, and then uses the length of that vector as a measure of fit between the model and the data--the selection problem thus becomes one of finding the parameterization of the model that produces the minimum length error vector. There are many ways to define the length of a vector, but the most common is the simple Euclidean length--the square root of the sum of the squares of its elements. Minimizing this square root is equivalent to minimizing the sum of squares--thus the common name 'least squares' given to the procedures motivated by this approach. Given general stochastic interpretations of the errors, least squares procedures can be proven to provide the best parameter selections (estimates) with respect to such selection criteria as bias and variance.

Frequently, the discrepancies between a model and empirical data are given particular stochastic interpretations--i.e., are assumed to have particular kinds of probability distributions. Under these conditions, it is possible to define a probabilistic 'likelihood' for a specified model given the actually observed data, and to use this likelihood as a measure of empirical fit. The selection criterion is then to find the model parameterization that has the 'maximum likelihood' of all possible parameterizations. Under common normality assumptions concerning the error probability distributions, the maximum likelihood selection criterion is equivalent to the least squares selection criterion; maximum likelihood procedures with non-standard distribution assumptions are frequently equivalent to minimum length error vector procedures with non-Euclidean vector lengths. If some probability distribution over the space of model parameterizations is accepted as known or assumable prior to the particular empirical test under consideration, then that prior distribution can be taken into account in the calculation of the model likelihoods using the procedures of Bayesian analysis.

Defining a measure of empirical fit in general allows comparison among any particular model parameterizations, but the task of finding the best parameterization with respect to that comparison is commonly the most difficult step. Ideally, there is a direct explicit expression for the best-fitting parameter vector given an observed data set and a particular measure of fit. In these circumstances, the only remaining consideration is that of the computational efficiency of the various computational versions of the expression (e.g., Bickhard, 1970). Unfortunately, such maximal information procedures are known only for the simplest cases (primarily for the least squares criterion).

Commonly, an explicit expression is known for the measure of fit, given a parameter vector and a data set, but none is known for the best parameter vector with respect to that measure--it must somehow be found in the general space of parameter vectors.* Of the general strategies for attempting to solve such problems (part of the subject matter of numerical analysis), the most common in statistical analysis is the stepwise search strategy. A parameter vector is selected, its measure of fit is calculated, a direction of search is determined, and a new parameter vector is selected some distance in the search direction from the prior vector--the process repeats until no direction offers improvement in measure of fit over the last selected parameter vector. (Clearly such a strategy offers only a locally best parameterization, but, under certain reasonable assumptions, this can frequently be shown to be the overall best as well-and, in any case, it may represent the most satisfactory procedure available.)

The most difficult part of such strategies is to find a direction of search for each new parameter vector--an explicit expression in terms of the parameter vector is again clearly desirable if available. If an expression for the first derivative of the measure of fit with respect to the parameter vector is known, it can be used as the direction of search in the method of steepest descent. Expressions for both first and second derivatives are used in the Newton-Raphson method, while several theorems, e.g., König's theorem (Householder, 1953), provide procedures for using higher order derivatives--if expressions are available and worth the computational labor.

---

* With different kinds of models and correspondingly different measures, this becomes familiar as the problem of heuristic problem-solving.

However, explicit expressions in terms of the parameter vectors may not be available for search directions, in which case the directions must somehow be determined from the prior set of parameter selections and their measures of fit. Again, there may or may not be explicit computational expressions, e.g., polynomial surface fitting followed by extrapolation--if not, perhaps the minimal information procedure is to graph the measures of fit against the parameter vectors and attempt to determine search directions by visual inspection. Certain procedures become very powerful by combining information from derivative expressions with information from prior selections--a primary example is the Fletcher-Powell method which uses an explicit expression for the first derivative and builds up information about the second derivative from prior selections (Fletcher and Powell, 1963, see also Shanno, 1968).

These procedures have all been discussed with the assumption of the availability of an explicit expression for a real number measure of fit. This assumption can fail in two ways--there is no explicit expression for the measure of fit, or the measure of fit is not a real number. The lack of explicit expression is frequently solvable by ‘model simulation’-- that is, by computationally simulating the model for each selected parameterization in order to determine the empirical constraints which that parameterization imposes, and defining the measure of fit directly in terms of these constraints (least squares is defined this way anyway). This might be combined with a search procedure based on prior selections, but clearly, since there is no measure of fit expression, there will be no explicit expressions for such directional information as derivatives.

Replacing explicit measure of fit expressions with model simulations in this way is more cumbersome, more computationally costly, and is compatable only with generally low efficiency search procedures, but, nevertheless, aside from these considerations of cost and efficiency, it introduces no new logical difficulties in the general selection process. Non-real number measures of fit, on the other hand, violate the assumptions of almost all known selection procedures. A measure of fit with a less rich mathematical structure than the real numbers, e.g., a partial ordering, may still be useful for comparing different model parameterizations, and in principle there may even still be a best-fitting parameterization with respect to that measure, but it is highly unlikely that mathematical procedures for finding it have been developed. Perhaps the most efficient procedure available would be graphical search, but commonly the goal of selecting the <u>best</u> parameterization is given up altogether, and a grid sampling of some portion of the parameter space is performed to see if <u>any</u> parameterizations can be found which meet the final decision criterion. It is possible, of course, for a well-defined measure of fit with an explicit computational expression to have less structure than the real numbers; most commonly, however, this condition is found when the fit criteria are so complicated as to defy any explicit measure, and the best that can be done is to make some subjective comparisons or judgments of 'reasonable' or 'unreasonable' concerning the manifested empirical constraints. Such conditions are currently to be found in such areas as personality modeling or urban modeling.

By whatever criterion and procedure, when a particular model parameterization has been selected, there remains the task of deciding whether or not to accept it--whether or not to act on the basis of it. (It should be noted that Simon's logic of satisficing combines the procedures of selection and decision--Simon, 1969.) Decisions can only be made among alternatives, and the alternatives in the case of model testing are alternative models. For some purposes, the alternative models may be interesting and viable models in themselves, perhaps involving different constraint structures, but usually they are various forms of nullification of the model being considered--null models of fewer or no constraints, rather than alternative constraints of similar degree. In some cases, e.g., prediction models, the constraints are relatively separable, and the overall model is generally compared for decision purposes with independent nullifications of each of these constraints in turn, e.g., with zero values for each parameter in turn--in other cases, especially with structural models, the constraints are not so easily separable, and the most relevant alternative model is the completely null model of no constraints. A model decision with a completely null model as alternative constitutes an overall goodness of fit test of the model--it is a decision as to whether the model fits the data well enough to be considered at all, aside from whether it is a significant improvement over various of its partial nullifications. Such an overall goodness of fit test is always desirable, but unfortunately it is frequently ignored even with structural models.

Decision rules are in principle definable with respect to any measure of fit (or of any other kind of model 'goodness'). They are commonly in the form of a simple threshold --the ubiquitous 5% significance level is an example of a familiar, if somewhat arbitrary, decision threshold. The issues and procedures involved in setting an acceptance threshold or region become interesting and complex, however, as consideration is given to the various types of errors possible in such a decision, the probabilities of those errors, and the costs of those errors. This is the subject matter of decision theory--though, as with numerical analysis, its applications are not limited to statistical models.

Least Squares Procedures for Structural Models

Explicit expressions and computational algorithms have been developed most fully in the case of the least squares selection criterion applied to models stated in algebraic form, in particular in linear algebraic form. In terms of the eight-celled classification mentioned earlier, algebraic statements are most appropriate to models with continuous variables, discrete transitions, and deterministic constraints--probably the most common conceptual model in the behavioral sciences. At least for the empirical versions of the models, however, the variables are frequently discrete, and, although the algebraic statements are usually stretched to handle these cases as well, such models with discrete variables are in fact state-transition models--a fact which generally remains implicit.

Fundamental to the logic of the least squares selection criterion (and other minimum length error vector criteria) is the construction of the error vector. The vector represents the overall discrepancy between the observed data and the modeled constraint, and each vector element is the discrepancy or error for a particular observed instance of the presumed constraint. These constraint errors are variously interpreted as representing unconsidered variables or error in the dependent variables (those presumed to be constrained)--but, however interpreted, it is important to note that there is only one such error for each instance of the constraint. In particular, the least squares statistical or error model does not allow for the consideration of separate errors of measurement for the particular variables involved. Maximum likelihood procedures do not suffer this general restriction and are more appropriate (when available) in cases where it seems desirable to introduce measurement error models for particular (independent) variables.

It should be noted, however, that this restriction on measurement error models in least squares analyses is not necessarily of great weight. In particular, in standard prediction analyses, the distinction between predictors and predicted is clear (constraining variables and constrained variables), and the legitimate concern is only with errors of prediction--errors with respect to the predictive constraint. A model for measurement error in a variable introduces a distinction between the ’true’ variable and the various sources of error in its measurement--in effect allowing models for the measurement processes separate from the model of the constraints among the true variables. With prediction as the goal, however, the concern is with the amount of predictive information in the measured variables, not with whether that information derives from ’true’ or ’error' sources. Prediction constraints are by definition among observed variables, while constraints among presumed ’true' variables represent conceptual inferences or constructions beyond the directly observable.

Causal constraints, even simple causal laws, are examples of presumed true variable constraints--and measurement error is correspondingly of potential salience in the analysis of even the simplest experiment.

An assumption in the derivation of ordinary least squares procedures, and thus another restriction on their application, is that the constraining variables (regressors) are not correlated with error. In models for which the distinction between constraining variables and constrained variables is clear and unambiguous, the constraining variables are considered to be non-stochastic and thus could not be correlated with error. A constrained variable, however, is stochastic by virtue of the contribution of the stochastic constraint error, and, if it should in turn serve to constrain some other variable, then it will be a stochastic constraining variable and potentially correlated with error--likely correlated, in fact, since the stochastic component is itself constraint error. Any model beyond the simple constraining-constrained dichotomy will thus have stochastic regressors by virtue of the reciprocal and sequential patterns of constraint in that structure, and will risk error correlated regressors.

A modification of the ordinary least squares procedure has been developed for the case of error correlated regressors--called two stage least squares. The first stage of the procedure is to regress all constrained variables on all unconstrained variables (thus ignoring any additional structure in the model), thereby obtaining ordinary least squares estimates of the constrained variables. The second stage consists of substituting these estimates for the corresponding stochastic regressors in the structural equations, and then proceeding with ordinary least squares on the structural equations. The ordinary least squares estimates are intended to approximate the stochastic regressors with as little correlation with error as possible. Details can be found in Goldberger, 1964, or Theil, 1971.

Again it should be noted that this consideration is of no weight for prediction analyses. Prediction models have a structure, but it is the dichotomous structure of 'before' and 'after' in time, and is clearly within the constraining-constrained dichotomy that does not introduce stochastic regressors. It might be thought that the goal of prediction might be better served with a more complex, perhaps more 'realistic,' structural model--one that would introduce stochastic regressors--, but in fact the information needed to estimate the extra parameters introduced by this extra structure will be subtracted from the fixed amount of information in the data, and will thus be unavailable for prediction. Simple causal laws share with prediction models the constraining-constrained dichotomy, and thus standard experimental analysis is also not effected by considerations of error correlated regressors.

Another critical assumption of ordinary least squares procedures is that the errors are uncorrelated among themselves and have equal variances. This assumption may be untenable with any kind of model, including a simple prediction model, but the likelihood of serious violation of the assumption is greatest with respect to dynamic structural models, especially with longitudinal data. The possibility of violating this assumption is frequently ignored in practice, but the consequences are potentially serious and the risk is usually taken out of ignorance rather than necessity.

A least squares procedure exists--called generalized least squares--which takes into account correlated errors and unequal error variances. It depends, however, on a knowledge of, or estimates of, the error variances and covariances, and these may or may not be available in the general case depending on the extent of prior information, the sample size, and the logical constraints one is willing to assume among the variances and co- variances. For models of sufficient complexity, however, a sample size sufficient to estimate the model parameters will be, together with a reasonable logical constraint, sufficient to estimate the error variances and covariances. A least squares procedure for utilizing this information in the structural equation case--called three stage least squares--consists of an initial two stage least squares calculation from which variance-covariance estimates are derived, followed by a generalized least squares solution to a modified form of the structural equations. Details are again to be found in Goldberger, 1964 and Theil, 1971.

Three stage least squares represents the most general least squares procedure available. Its generality is restricted by the lack of measurement error considerations, and by the necessity of a sufficient set of variables unconstrained in the model, but these restrictions seem fundamental to the definition of the least squares criterion, not just expedient assumptions for the derivation of efficient procedures. The initial values of, and the primary inputs to, an open system will commonly provide a sufficient set of model unconstrained variables, leaving the lack of measurement error considerations as the most salient least squares restriction for the testing of structural models. It should be pointed out that we have discussed only least squares procedures, and only a limited number of them. The general question of the best procedure for testing structural models has not been settled, and probably has no single answer (Goldberger, 1964), but there is nevertheless no question that some are better than others.

A MODEL OF ORGANIZATIONAL BEHAVIOR

In this section, we would like to very briefly develop a particular issue of systems theoretic model building, and illustrate it with a model we are currently investigating of social-psychological processes in organizations.

The Structure in Structural Models

A system is commonly thought of as an interconnected and interacting set of components. Similarly, a systems model is frequently thought of as a model of these component interconnections, or couplings. If we model a system solely by its real interconnections, however, e.g., this resistor is connected to that capacitor, or this person reports to that person, we may have the static structural knowledge of a circuit diagram or an organization chart, but with little consequent knowledge of the system's dynamics and behavior--interconnections determine which components will interact, but they do not per se determine what the interactions will be.

Knowledge of the dynamics of a system depends on a knowledge of the dynamic constraints among its variables. If we model a system solely by the constraints among its external variables, however, we may know something of the lawful behavior of the system, but will still remain ignorant of the internal system processes that manifest that behavior. Clearly, then, we wish to have dynamic models of a system's internal processes--dynamic because we wish to model the system in time, not just at an instant of time, and of internal processes because these will both explain and extend our knowledge of the externally manifested constraints. In other words, we seek models of the constraints among the <u>internal</u> variables of the system, as argued in the first section of this paper.

Such internal constraints will be of two general kinds: those mediated by the couplings in the system, and those manifested by the components of the system. We are, then, interested in the coupling structure of a system, but, for dynamic modeling purposes, only insofar as it allows us to determine the pattern of the variable constraints among the systems components. Note that knowledge of the existence of a real coupling is logically independent of knowledge of the variable constraint(s) which that coupling mediates: as above, the couplings impose a structure on the inter component constraints, but do not determine what those constraints are.

One strategy for modeling the <u>within</u> component constraints might be to represent the next lower level subcomponents and their interactive constraints. Clearly, however, this can only proceed a finite number of times, and at some level of subcomponent we must accept a component model expressed solely in terms of its externally manifested constraints --i.e., without reference to internal components or couplings. Models of units in terms of external constraints have at times been called 'object' models, in distinction from systems models (Goguen, 1970, Zadeh and Desoer, 1963), and, in this terminology, a systems model becomes a model of interconnection constraints among objects. In effect, the object models represent the within unit constraints, the coupling models represent the between unit constraints, and together they determine the general system dynamic constraints.

Components of an Organizational Model

As argued by Kate and Kahn (1965) and Kahn, et al. (1964), organizations would seem to be ideal entities for analysis from a systems perspective. The best choice of components for such a model is, however, far from obvious. Among several possibilities, perhaps the most promising and apparently the most favored is the social role: as a dynamic functional unit rather than a physical unit, the role offers the possibility of a direct model of organizational processes--the role is essentially a social psychological behavior model for an individual's interactions taken as the lowest considered component level.

We have found, however, that the concept of social role is still significantly confused between static and dynamic characteristics. At times it is used as a dynamic behavior or input-output unit, at times it is a static form of coupling between individuals; much of the time it seems to be simply ambiguous with respect to these characteristics.

Although many of these characteristics of the role concept have been discussed and explicated (Turner, 1962, Biddle and Thomas, 1966), we have nevertheless found it advantageous for our purposes to use the basic work group as our primary organizational component. The advantages have been threefold: work groups seem still to be easier to distinguish and identify than functional roles, data collection is easier since there are fewer components to be considered, and the work group as component makes a first approximation to the distinction between primary group processes and more formal processes--this allows us to represent the work group by a model of general constraints among its external variables, e.g., general work attitudes, interpersonal style, etc., thus avoiding the need for a systems model of group processes, and allowing us to focus more directly on the between group organizational processes. Note that the interconnections among work groups may be expressed in terms of the external role relationships of the group members, but that this is explicitly a coupling role concept, not a dynamic role concept.

The General Statistical Formulation of the Model

The model constraints are being expressed as linear algebraic equations. One important consequence of the restriction to linear models is that the various sources of constraint on a variable may be linearly separated. In particular, the constraints on a particular variable for a particular work group may be separated into those from other variables in the same work group and those from variables of different work groups. Thus we may consider the within workgroup model and the between work group model separately.

For a particular variable Y_i presumed to be constrained within work groups, that constraint may be expressed by $\underline{Y}_i = X_{i1}\,\underline{\beta}_{i1} + \underline{\varepsilon}_{i1}$. In this equations, the index i indicates that the equation is specific to the ith variable Y_i; the elements of the vector $\underline{Y}_i$ are the values of that variable Y_i for each work group; the second subscript of '1' indicates that this is a within work group equation; the matrix X_{i1} is composed of the column vectors of the variables presumed to constrain Y_i in the within group model; $\underline{\beta}_{i1}$ is the vector of unknown coefficients for those constraints; and $\underline{\varepsilon}_{i1}$ represents the constraint error.

The between work group constraints on Y_i (if any) may be similarly expressed as $\underline{Y}_i = AX_{i2}\,\underline{\beta}_{i2} + \underline{\varepsilon}_{i2}$. The vector $\underline{Y}_i$ is as before; X_{i2} contains the variables which constrain Y_i from other work groups (some of which might possibly also be in X_{i1});+ and $\underline{\beta}_{i2}$ and $\underline{\varepsilon}_{i2}$ are the between group counterparts to $\underline{\beta}_{i1}$ and $\underline{\varepsilon}_{i1}$. The matrix A, however, is new: it is the group coupling or communications matrix. The matrix A expresses the organization-wide group coupling structure--with zero entries for group pairs that are not coupled and non-zero entries for coupled groups--and its entries determine which groups will exert constraints on which other groups--that is, if group k (represented by the k_{th} element of $\underline{Y}_i$) is to be constrained by group k' (represented by the k'_{th} row of X_{i2}), then the element $a_{kk'}$ of the matrix A must be non-zero, i.e. , group k' must be coupled to group k.

We can now combine the expressions for the within group and between group constraints on a particular variable as follows:

$$\underline{Y}_i = X_{i1}\,\underline{\beta}_{i1} + AX_{i2}\,\underline{\beta}_{i2} + \underline{\varepsilon}_{i2} = \left[\ X_{i1}\ \ AX_{i2}\ \right] \left[\ \underline{\beta}_{i1} / \underline{\beta}_{i2}\ \right] + \underline{\varepsilon}_i = X_i^{*}\,\underline{\beta}_i^{*} + \underline{\varepsilon}_i$$

This last form of the equation is a form suitable for estimation and testing by least squares procedures, subject to the modifications discussed previously. Note again that this strategy for combining within-unit constraints with between-unit constraints in one expression is not

+ The introduction of lagged variables as entries in these X matrices poses special problems of estimation—discussed generally in the section on least squares procedures.

limited to work groups in organizations--it depends only on the linear separability of sources of constraint in any model whatsoever, a fairly weak restriction.

Adaptations of the within group and between group models described in Murray, et al. (1970) will be tested within the above general statistical formulation, with necessary elaborations to accommodate the longitudinal structure of the data.

BIBLIOGRAPHY

Anderson, T.W. An Introduction to Multivariate Statistical Analysis. New York: Wiley, 1958.

Berrien, F. K. General and Social Systems. New Brunswick, New Jersey: Rutgers University Press, 1968.

Bertalanffy, L. Von. General System Theory. New York: Braziller, 1968.

Bickhard, M.H. "Linear Least Squares Numerically." Unpublished Master's Paper, The University of Chicago, 1970.

Biddle, B.J. and Thomas, E.J. Role Theory. New York: Wiley, 1966.

Blalock, N. Theory Construction. Englewood Cliffs, New Jersey: Prentice-Hall, 1969.

Blalock, H.M. Jr. Causal Inferences in Non-Experimental Research. Chapel Hill: University of North Carolina Press, 1961.

__________and Blalock, A.B. Methodology in Social Research. New York: McGraw- Hill, 1968.

Blau, P.M. and Scott, W.R. Formal Organizations. San Francisco, Calif. : Chandler, 1962.

Buckley, W. Sociology and Modern Systems Theory. Englewood Cliffs, New Jersey: Prentice-Hall, 1967.

__________, ed. Modern Systems Research for the Behavioral Sciences. Chicago: Aldine, 1968.

Bunge, M. Causality. New York: World Publishing, 1959.

__________. "Conjunction, Succession, Determination and Causation." International Journal of Theoretical Physics, Vol. 1, No. 3 (1968), 299-315.

Campbell, D. T. and Stanley, J. C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1963.

Cartwright, O. and Zander, A. Group Dynamics. New York: Harper & Row, 1968.

Carzo, R. Jr. and Yanouzas, J. N. Formal Organization. Homewood, 111. : Dorsey Press, 1967.

Chisholm, R.M. Theory of Knowledge. Englewood Cliffs, N.J.: Prentice-Hall, 1966.

Dempster, A. P. Elements of Continuous Multivariate Analysis. Reading, Mass. : Addis on - Wesley, 1969.

Doob, J. L. Stochastic Processes. New York: Wiley, 1953.

Emery, F.E., ed. Systems Thinking. Baltimore, Md. : Penguin, 1969.

Ernst, G. W. and Newell, A. GPS: A Case Study in Generality and Problem-Solving. New York: Academic Press, 1969.

Etzioni, A. Complex Organizations. New York: Free Press, 1961.

__________, ed. Complex Organizations. New York: Holt, Rinehart and Winston, 1965.

Feigl, H. and Brodbeck, M., eds. Readings in the Philosophy of Science. New York: Appleton-Century-Crofts, 1953.

__________and Maxwell, G., eds. Current Issues in the Philosophy of Science. New York: Holt, Rinehart, and Winston, 1961.

Fletcher, R. and Powell, M. J.D. "A Rapidly Convergent Descent Method for Minimization." The Computer Journal, 06 (1963), 163-168.

Forrester, J.W. Urban Dynamics. Cambridge: M. I. T., 1969.

Glushkov, V.M. Introduction to Cybernetics. New York: Academic Press, 1966.

Goldberger, A. S. Econometric Theory. New York: Wiley, 1964,

Goguen, J. "Mathematical Representation of Hierarchically Organized Systems." Global Systems Dynamics. Edited by E.O. Attinger. New York: Wiley, 1970.

Hempel, C.G. Philosophy of Natural Science. Englewood Cliffs, N.J.: Prentice-Hall, 1966.

Hildebrand, F. B. Introduction to Numerical Analysis. New York: McGraw-Hill, 1956.

Homans, G. C. The Human Group. New York: Harcourt, Brace & World, 1950.

Householder, A. S. Principles of Numerical Analysis. New York: McGraw-Hill, 1953.
__________. The Theory of Matrices in Numerical Analysis. New York: Blaisdell, 1964. Johnston, J. Econometric Methods. New York: McGraw-Hill, 1963.
Kahn, R. L., Wolfe, D. , Quinn, R. and Snoek, J. Organizational Stress. New York: Wiley, 1964.
Katz, D. and Kahn, R. L. The Social Psychology of Organizations. New York: Wiley, 1966.
Klir, G. J. An Approach to General Systems Theory. New York: Van Nostrand, 1969.
Lange, O. Wholes and Parts. New York: Pergam on, 1965.
Luce, R. D. and Raiffa, H. Games and Decisions. New York: Wiley, 1957.
MacLane, S. and Birkhoff, G. Algebra. New York: MacMillan, 1967.
Mendenhall, W. The Design and Analysis of Experiments. Belmont, Calif. : Wadsworth, 1968.
Mesarovic, M.D., ed. Views on General Systems Theory. New York: Wiley, 1964.
Mills, T.M. The Sociology of Small Groups. Englewood Cliffs, N. J.: Prentice-Hall, 1967. Minsky, M. Computation: Finite and Infinite Machines. Englewood Cliffs, N.J.: Prentice - Hall, 1967.
Mitchell, G. Theory of Categories. New York: Academic Press, 1965.
Murray, J.R. et al. "Research on the Dynamics of Organizational Change." Unpublished draft, Industrial Relations Center, University of Chicago, 1970.
Nagel, E. The Structure of Science. New York: Harcourt, Brace & World, 1961.
Olmstead, M. S. The Small Group. New York: Random House, 1959.
Popper, K. R. The Logic of Scientific Discovery. New York: Harper, 1959.
Rao, C. R. Linear Statistical Inference and Its Applications. New York: Wiley, 1965. Rescher, N., ed. The Logic of Decision and Action. University of Pittsburgh Press, 1966. Rudner, R. S. Philosophy of Social Science. Englewood Cliffs, N. J. : Prentice-Hall, 1966. Shanno, D. F. "Conditioning of Quasi-Newton Methods for Function Minimization." Unpublished manuscript, University of Chicago, 1968.
Simon, H. A. Models of Man. New York: Wiley, 1957.
__________. The Sciences of the Artificial. Cambridge: M.I.T., 1969.
Slagle, J.R. Artificial Intelligence. New York: McGraw-Hill, 1971.
Theil, Henri. Principles of Econometrics. New York: Wiley, 1971.
Thibaut, J. W. and Kelley, H.H. The Social Psychology of Groups. New York: Wiley, 1959. Turner, M. B. Philosophy and the Science of Behavior. New York: Appleton-Century-Crofts, 1967.
Turner, R.H. "Role Taking: Process Versus Conformity." Human Behavior and Social Processes. Edited by A. M. Rose. Boston: Houghton - Mifflin, 1962.
Wiener, N. Cybernetics. Cambridge: M.I.T., 1961.
Zadeh, L. A. and Desoer, C₀ A. Linear System Theory. New York: McGraw-Hill, 1963.
__________ and Polak, E., eds. System Theory. New York: McGraw-Hill, 1969.