AD-A048 147 OHIO STATE UNIV RESEARCH FOUNDATION COLUMBUS F/G 5/4

THE INTER-NATION SIMULATION PROJECT: A METHODOLOGICAL APPRAISAL--ETC(U)

1974 S J THORSON DAHC15-73-C-0197

UNCLASSIFIED RP-25 NL

| OF |

AD A048147

END

DATE FILMED

1 -78

DDC

1.0 1.1 1.25 1.4 1.6 1.8 2.0 2.2 2.5 2.8 3.2 3.6 4.0

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

WORKING DRAFT

1

The Inter-Nation Simulation Project:

A Methodological Appraisal*

Stuart J. Thorson
Polimetrics Laboratory
Department of Political Science
The Ohio State University

Prepared for delivery at the 1974 Annual Meeting of the International Studies Association, St. Louis, Missouri, March 18-21, and is forthcoming in Quantitative International Politics: An Appraisal, Sage Publications.

* *Support for this work was provided in* part by the Advanced Research Projects Agency (DAHC 15 73 C 0197).

Research Paper No. 25

Definitional discussions, especially in short papers, may appear to be carried on at the expense of "doing what the paper is about." However, I prefer to view the clarification of definitional problems as initial investments which yield high (and immediate) rates of return in the form of both increased clarity and conciseness. Therefore the first section of this paper will involve stating as precisely as possible the questions to be discussed in the following sections. In addition, though no less important, the next section should serve to provide a general context in which the specific topics I will be treating become interesting.

I Problems and Context

The task of this paper, as is evident from its title is to appraise the methodological practices used in the Inter-Nation Simulation Project (INS). Such a task almost requires quoting the remarks with which R.F. Harrod (1938) began a "methodological essay:"

> Exposed as a bore the methodologist cannot take refuge behind a cloak of modesty. On the contrary, he stands forward ready by his own claim to give advice to all and sundry, to criticise the work of others, which, whether valuable or not, at least attempts to be constructive; he sets himself up as the final interpreter of the past and dictator of future efforts.

Such observations are especially important as one attempts to "assess" such a major and obviously constructive project as INS. The enormity of the technical output generated in INS related research can be begun to be sensed as one reads the annotated bibliography of Leserman (1972). That a political scientist might undertake to critique INS from a methodological perspective is itself an indication of the successful accomplishment of Guetzkow's goal of developing a "college of simulators" within political science. Indeed, one measure of the success of an innovative project such as INS might well be the variety and quality of criticism it draws.

Since this paper involves, in part, "doing methodology," it will be useful to clarify how I will be using the term. This is especially important since many behavioral scientists appear to use the two terms "methodology" and "technique" as synonyms. I would like to distinguish the two and suggest the following. Methodology deals with questions of justification. "The method of a science is, indeed, the rationale on which it bases its acceptance or rejection of hypothesis or theories (Rudner, 1966:5)." Techniques, on the other hand, are ways of generating information which is used in evaluating a hypothesis or theory. Thus methodological positions will influence choice of technique. For example, a Hempelian scientist would probably argue that, for methodological reasons, the technique of crystal ball gazing is an inappropriate technique for the scientist. Techniques then are ways of generating evidence and methodology provides the rules according to which the evidence is admitted and, if admitted, evaluated.

The distinction I am drawing between technique and methodology is important for several reasons. First, it suggests that we can not critique the use of a particular technique without first specifying the methodological perspective from which our critique is directed. Second, and perhaps a corollary of the first, techniques which are inappropriate from one methodological perspective can become appropriate under another. Thus in appraising the techniques used in INS I must be concerned with questions of methodology choices. Indeed, as will be argued, different methodological choices may be made depending upon the purpose(s) of the investigation. Third, and perhaps least important for this paper, distinguishing technique from methodology allows the possibility that not being able to use all the techniques of, say physics, does not preclude the use of a methodology appropriate to physics. To summarize, all three reasons support the position that question of technique cannot be totally separated from the philosophical questions surrounding choice of method.

Since whatever else INS may be viewed as having done, it certainly involves the use of simulation in the investigation of international relations phenomena, it is important to define what is meant by the term simulation. In doing this, it will be helpful to discuss the related concepts of model and theory. After providing explications and definitions for these terms, I will then attempt to pose the specific questions the remainder of this paper will address.

Harold Guetzkow has defined a simulation as

> "An operating representation in reduced and/or simplified form of relations among social units by means of symbolic and/or replicate component parts (Guetzkow, 1959, p. 184)."

The use of the adjective "social" restricts the class of simulations, but to generalize the definition one need only remove social. Indeed, this definition suggests that there are three elements which need be specified in talking about a simulation. First there is the system to be simulated, S. Second there is the representation of that system, M. Third there are the statements according to which the representation is described and/or "operates," T. These statements might be in Fortran, PL1, English, Chinese, etc. We can think S,M, and T as being sets. That is, S is the set of all referrents (or systems to be simulated); M is the set of all representations of S; and T is the set of all statements according to which the representations "operate."* Notationally we can write:

$$S: \{ s_1, s_2, \ldots . s_n \}$$

$$M: \{ m_1, m_2, \ldots . m_o \}$$

$$T: \{ t_1, t_2, \ldots . t_p \}$$

Any specific simulation can then be identified with an ordered triple which tells us which member of S is being represented by what member of M in terms of what statements in T. These distinctions are use are useful because they point out that under

---

* This discussion is based in part upon Zeigler,(1970)

the Guetzkow definition changing any one of the three elements gives us a new simulation. Thus in predicating statements of INS we must be as careful as we can in specifying each of the three components. All too often the literature seems to equate INS only with an element of M. As will be argued below, this has the implication of ignoring the impact changes in the $s \in S$ might have upon the resolution of methodological and technical questions.

However before I can make these arguments I must first explicate more completely what elements of S and M look like. Elements of S are difficult to visualize in any concrete way since, by visualizing them they become represented. The elements of S are the "realities" discussed in Guetzkow (1968) or, equivalently the "reference systems" discussed in Hermann (1967, p. 220). There is no reason to require that these realities be in any sense completely known. Indeed it is precisely because they are complex (and perhaps partially unknown) that we choose to represent them as a part of a simulation effort (an elegant argument for this point is found in Naylor, 1971, pp. 2-10).

An adequate representation consists of specifying the <u>objects</u> making up the representation <u>and</u> the <u>relations</u> which are defined on these objects. That is, a representation consists of a well specified world. This notion of representation is, I believe, the one intended by Guetzkow when he writes, "Simulations in international relations attempt to represent the on-going international system or components there of, such as world alliances, international organizations, regional trade processes, etc. (1969, 285)," and by Hermann, "A simulation or game is a partial representation of some independent system (1967, 216)".

A representation is a mathematical structure (a collection of objects and relations) and one can ask questions of possible relations between the representation and the referrent reality. For example, is there a homomorphism between s_i and m_j; an isomorphism; or, more generally, what sorts of morphisms ("generalized" mappings) might one be willing to assert to obtain between s_i and m_j. Loosely speaking these might be termed questions of external validity. Answers to these questions will require an examination of possible purposes of INS and will be considered later in this essay.

Simply representing a referrent reality (i.e., picking an $m \in M$) is not enough to provide us with a simulation. The representation will often be too complex to study directly. Therefore we will be interested in seeing what sorts of behaviors are produced (generated) by the representation. This objective requires developing a theory of the representation. This theory will take the form of statements "about" the representation. In the case of a simulation, the theory may be viewed as a set of "operating rules." The language in which these operating rules or statements are written will depend in part upon the mechanism chosen for use in the simulation. In an all machine simulation, the statements might all be in Fortran, or

Dynamo, or Simscript, etc. In a simulation combining human subjects and a computer, the operating rules may be partly in Fortran and partly in some natural language.

Once again, in any specific simulation effort, we will be led to ask questions as to possible relations to assert between elements of T and the elements of S and M. Here, however, our concern is with relations between structures (m_i and s_j) and sets of sentences (t_k). Or, more generally, one concern is with the relations between sentences and objects and relations "referred to" by those sentences. Such relations may be termed semantic (Tarski, 1944). An example of such a semantic relation is truth. Are the sentences true of the representation? Are they true of the referrent reality? Whether or not particular sentences (or sets of sentences) are accepted as true is a question of methodology. What it means to assert them to be true is a (non-trivial) question of semantics.

Thus far this discussion has been focused upon simulation. Perhaps the reader has sensed that what is involved in doing simulation is quite similar to what is involved in doing theory in general. Therefore it will be useful to quickly relate what has been done so far to the commonly encountered terms theory and model. Accomplishing this will enable me to bring in more general methodological points as this essay progresses.

The particular set of sentences or statements (i.e., $t \in T$) making up the simulation are roughly equivalent to a technical sense of theory. They are sets of sentences which are asserted to be true of some world. Further, these sets of sentences will generally (always if an artificial programming language is used) have some preassigned logical framework or "calculus axioms" (e.g., first order predicate calculus, probability theory, etc.). These axioms allow us to investigate the implications of subsets of our set of theory statements.

A model for a theory is that "thing" which makes the sentences in a theory true. In theorizing we generally want to order or account for some aspects of a perceived reality or referrent system. Thus we represent reality in terms of some posited objects and relations. Whether or not these posited objects and relations indeed represent reality is of course in many senses moot and is certainly contingent upon both our perceptual system and our ability to make and hold to distinctions.

However, as we have seen, a collection of objects and relations is a mathematical structure and not a theory. We must write down some sentences describing (i.e., which are true of) this structure. These sentences I have termed a theory. The underlying structure I will call a model for that theory.

More specifically, a mathematical structure m is a set of elements (objects), $A = \{a_1, a_2, \ldots\}$, together with a set of relations of order i, $P_1^{i_1}$, $P_2^{i_2}, \ldots$, and may be expressed

$$m = \langle A; P_1^{i_1}, P_2^{i_2}, \ldots, P_n^{i_n}, \ldots \rangle.$$

A formal language L in which properties of M can be expressed will consist of formulas generated by a specified set of rules, say the predicate calculus, from an alphabet consisting of relation symbols (R_1, R_2, . . .), variable symbols (x_1, x_2, . . .), connectives ($\neg, \vee, \wedge$, . . .) and quantifiers ($\forall, \exists$). Since functions and constants are special kinds of relations, function symbols (f_1, f_2, . . .) and constant symbols (c_1, and c_2, . . .) will also be used in L. The language L will be assumed to be first order, that is, its variables range over the elements of A (as opposed to ranging over the subsets of A, or sets of subsets, etc.). Sentences in L are formulas containing no free variables.

Let T be a set of axioms in a language L. If ϕ is a mapping of constant symbols occurring in T into the set of objects A, and also a mapping of relation symbols occurring in T into the set of relations in M, then M provides an interpretation of T under ϕ. If this interpretation results in the sentences in T being true, then M is said to satisfy T and M is a model of the axiom set T. A model for a set of axioms then, is a mathematical structure which is used to interpret the axioms in such a way that the axioms are true.*

The above discussion may appear needlessly abstract. However, its generality enables us to do several things. First it makes explicit a relationship between the simulation enter-

---

* This discussion is taken from S. Thorson and J. Stever, (1974).

prise and the theory enterprise. To be specific, under these definitions, simulations become a subset of theory. This means we can employ (as is suggested in Guetzkow (1968)) the kinds of critical methodological tools used to analyze scientific theory to analyze simulation efforts such as INS.

Second, I have said that one component of any simulation is an explicit representation ($m \in M$) of a referrent system. Thus simulations belong to the subclass of theories with explicit models. By a result from model theory, we know that if an axiomatic deductive theory has a model, then it is logically consistent (i.e., it is non-contradictory). This point, while subtle, is of extreme importance. It was not until the nineteenth century, for example, that model theory was employed to show that negating the parallel lines axiom of Euclidean geometry did not lead to an inconsistency (assuming Euclidean geometry itself to be consistent).

Third, by making the representation (model) explicit, it is possible to efficiently investigate results of slightly (or grossly) perturbing the representation. This kind of sensitivity testing is very difficult to do for theories without explicit models. That INS was used in such a way is documented by the number of INS "variants" which have been produced (e.g., Bremer (1970), Smoker (1968), Abt and Gordon (1969), and Leavitt (1971)).

Having defined simulation and related this definition to that of theory and model it should be apparent that a methodological critique (under my use of methodology) of INS must address questions of the *utility* of INS in the development of international relations theory. To be fair to the INS project, such an appraisal must attempt to perform this evaluation within the context of the projects stated theoretical goals. Therefore, I will organize the remainder of this essay around the general problem of validity with respect to a variety of stated theoretical objectives.

II Methodological Critique

One of the great strengths of the INS project may be seen as its continuing and serious concern with problems of validation. This concern has been reflected in such papers as Alger (1963), Chadwick (1967), Crow and Noel (1965) Guetzkow (1966, 1967), Hermann (1967), Hermann and Hermann (1967), Kress (1966),McGowan (1972), Nardin and Cutler (1967), Noel (1963), Raser, Campbell,and Chadwick (1972), Robinson et al (1966), and Verba (1964). Given the methodological focus of this essay, perhaps two of the most important of these papers are Guetzkow (1968) and Hermann (1967). The Hermann paper is something of a classic and is often cited both within and outside of the behavioral sciences as a standard piece on the validation of simulations.

Hermann (1967, 217) argues that "....validity is not a singular issue..." and that "...we may more accurately refer to multiple validity issues." He then goes on to identify three components of validity issues.

> First the validity of an operating model is affected by the purpose or use for which the game or simulation has been constructed. What may be a relatively valid operating model for one objective may be strikingly unsatisfactory for another. Second, model validation can be expected to vary according to the type of validity criteria employed. Third, the validation issues will be significantly altered depending on whether human participants are introduced into the model (1967: 217).

Guetzkow (1969) identified three general purposes for simulations in the study of international relations:

> Simulations may serve in three ways as formats through which intellectuals may consolidate and use knowledge about international relations: (1) Simulations may be used as techniques for increasing the coherence within and among models, enabling scholars to assess gaps and closures in our theories; (2) Simulations may be used as constructions in terms of which empirical research may be organized, so that the validity of our assertions may be appraised; (3) Simulations may be used by members of the decision-making community in the development of policy, both as devices for making systematic critiques, through "box-scoring" its failures and successes, and as formats for the exploration of alternative plans for action (1969, 286).

These three purposes might be labelled "programmatic guide to research," "description," and "policy" respectively. That is, in the first the concern is with integrating "islands of theory" and organizing "the division of labor more coherently among the scholars working within international affairs (1969: 288)." In the second, the emphasis is upon the degree of correspondence between the representation and the reference system, i.e., description. In the third, primary focus is upon policy planning and the specification of alternative futures. Policy planning may be viewed as involving the specification of probability distributions over consequences of alternative (feasible) policies. For the purposes of this essay it will be helpful to separate the "policy" objective from what might be termed the "design" objective.

In design we are less concerned with identifying impact of alternative (presently) feasible policies and more concerned with identifying new structures for the achievement of particular goals. This notion of design is compatible with what Guetzkow (1966, 1969) has termed "constructin alternative futures". As he points out, this concern with design is one of the oldest raditions in political science. And, he argues, simulation is an important tool for the design theorist. One reason for this importance is, as we have already seen, due to the relative case with which the impact of alternative representations (models) can be investigated.

A fifth purpose for simulation (and INS in particular) which is often cited (e.g., Guetzkow (1959)) is that of education. INS has been used to teach principles of international politics in high schools, colleges and universities and various governmental bureaucracies. The few studies I have seen of the effectiveness of INS versus more traditional techniques (e.g., case studies, Robinson (1961)) have not found INS to be clearly "superior". However,given the lack of agreement on appropriate educational objectives (even within the cognitive domain) the question of the relative utility of INS type simulations for students with certain educational objectives is still an open one.

Leaving aside the educational purpose, let me next examine

INS along each of the first four objectives in somewhat more detail. These four were:

1) Programmatic guide to research
2) Description
3) Policy
4) Design

Programmatic guide to research: This purpose is potentially one of the most important for large scale simulation projects. Reasons for this importance have been brought out in another context by Allen Newell (1972) in a very interesting analysis of contemporary experimental psychology. He argues "that the two constructs that drive our current experimental style are 1) at a low level, the discovery and empirical exploration of (discrete empirical phenomena) and 2) at the middle level, the formulation of questions to be put to nature that center on the resolution of binary oppositions." This characterization could be equally well made of the contemporary empirical study of international relations. A look at recent journal articles or convention papers illustrates the concern with discrete phenomena. Further these investigations are often driven by and imbedded in such binary distinctions as internal/external, conflict/cooperation, rational/irrational, large/small, open/closed, developed/underdeveloped, etc. Will this approach to science (Newell terms it "playing twenty questions with nature") work? It may. However, "reality" may be too complex to yield to this approach. That is, we may not be able simply to add up answers to these "simple" questions to get a general theory of international relations. An alternative to playing twenty questions with nature is to construct "complete processing models," i.e., large scale simulations such as INS. Such simulations allow us to examine particular phenomena (e.g., nuclear proliferation (Brody, 1963) or public goods and alliances Burgess and Robinson (1969)) as part of a general problem structure (INS).

INS appears to have been moderately successful in terms of this objective. Numerous studies have been done using INS as a representation within which to investigate specific hypotheses. However, before we can begin to assess the utility of these studies it will be necessary to consider the related objective of description.

Description: In the introductory section of this paper I identified three elements in a simulation - referrent system (S), representation or model (M), and statements (T). The question of descriptive adequacy is a question of the S-M relationship. If both the referrent system (reality) and the model are viewed as "black boxes" with inputs and outputs, the problem of descriptive adequacy can be posed in terms of the morphism (mappings) which are preserved between S and M.

Brodbeck (1959) in her essay on models and theories suggests that there should be an isomorphism between s and m. An isomorphism is a term used to denote a mapping between two structures such that there is a one to one correspondence between the objects and relations of the first and the second structure. Yet, as Guetzkow (1968, 207) points out, such a requirement is far too strict. I would even go beyond this and argue that such a requirement is generally undesirable. A model which was isomorphic to reality would be as intractable as reality itself. "...simulations...- like other models - are always a simplification of their reference system (Hermann, 1967: 217)." Thus we generally ask that the relation between the reference system (s) and the model (m) be a homomorphism (Guetzkow, 1968: 207). Here, at an intuitive level, rather than requiring a one-one corresponce, we allow many elements and relations of s to be mapped into those of m.

However, and to my knowledge this has not been explicitly considered in INS related research, there are a variety of morphisms which might be asserted to obtain between s and m (e.g., see Zeigler, 1970, 1971). Three such morphisms can be termed "behavior preserving," "function preserving," and "structure preserving." The weakest of these is behavior preserving. Here the concern is only that equivalent inputs in s and m produce equivalent outputs. Function preserving morphisms preserve not only input-output relations, but also internal state changes (see Arbib, 1969 for a discussion of state). Finally the most restrictive morphism - structure preserving - preserves (in addition to input-output relations and state transition functions) the manner in which these relations and functions arise out of local coordinate functions (see Zeigler, 1970; pp 6ff). While I will not discuss structure preserving morphisms in this paper, the set theoretic view of simulation developed in the beginning of this essay is general enough to pursue such investigations.

The first two of these morphisms - behavior preserving and structure preserving - may be formalizations of what Guetzkow meant when he wrote:

> Some homomorphy may exist among outputs as well as between the very processes which result in such outputs. As we analyze the correspondences between simulations and "realities" sometimes an internal process, like the representation of the decision-making within foreign offices, helps produce an outcome of some validity, such as the constellation of internation alliances. At other times less often because of lack of appropriate research an internal process will be judged to be of some validity because the very process itself has some congruence with corresponding processes in the reference data (1968, 207).

A homomorphism among only outputs would be similar to the behavior preserving morphism, while a morphism which preserved outputs as well as "internal processes" would be a function preserving morphism.

Most of the Hermann (1967) essay seems to be concerned with identifying criteria for being able to assert a behavior preserving morphism between s and m. In the section discussing criteria for assessing the "fidelity with which a model produces aspects of reality," he suggests five validity standards:

1) Internal validity

2) Face validity

3) Variable - parameter validity

4) Event validity

5) Hypothesis validity

It will, I think, be useful to consider briefly each of these for each has implications for the method by which simulations are constructed.

"Any exogenous inputs introduced during the course of the game are held constant across all trials or runs. The unexplained variance between these intended replications would provide a measure of reliability or what Campbell (1957) calls 'internal validity.' When the structured simulation properties are held constant, the smaller the between-run variance, the greater the internal validity is assumed to be (Hermann, 1967: 220)." This notion of internal validity or reliability seems to me to be somewhat misleading. First of all note that the concern here is not with a relation between s and m. Rather an interest is purely in properties of m (or perhaps relations between m and t).

Hermann seems to be asserting that it is desirable for two runs of the same simulation with equivalent input to show equivalent output (response). With input fixed, he wants to minimize between run variance. However if the simulation is stochastic (i.e., probabilistic) either due to the explicit use of pseudo-random number generators or due to "noise" resulting from the use of human subjects (as in INS), the response surface of the simulation will itself be a random variable which will have associated with it a variance etc. One reason for constructing the simulation may be to estimate the variance associated with the response. In fact, a similar sort of argument may be found in Guetzkow)1963; 117). There is no general a priori reason to desire low between run variance.*

This criticism may appear trivial, but I think that it generalizes too much of the INS validation effort. That is, reality itself (s) is seen as highly internally valid in the Hermann sense and it therefore is expected to have very low variance. Since, by hypothesis, reality is low variance, the external validity of INS ought depend in large part upon its ability to reproduce reality. This position should become clearer as we discuss Hermann's remaining four criteria.

The second such criteria, face validity, "is a surface or initial impression of a simulation or game's realism (1967: 221)." As Hermann points out, this criterion is rather vague and is generally useful only in early stages of simulation development. However, recent efforts (Thorson and Phillips, 1974; Richardson, 1974) suggest that face validity may be

---

* It might be counter argued that in the case of stochastic computer simulations the pseudo random number seed is also an input value. Therefore we would want to have low between run variance with the seed fixed. However this argument loses force when the problem of internally validating the pseudo random number generator itself is considered (see Mihram, 1972 pp 18-146).

capable of being rendered more precise by identifying the mental images of individuals who "work with" the process being represented.

The notion of variable-parameter validity "involves comparisons of the simulations's variables and parameters with their assumed counterparts in the observable universe (1967: 222)." In terms of one earlier discussion, we are concerned that the objects of m correspond to objects in s. There are at least two problems with the use of variable-parameter validity in evaluating INS. First there is the problem of aggregation. Even a complex simulation such as INS deals with variables at a highly aggregated level. Thus we might not expect the INS variables to have simple "real world counterparts." In fact, unless we are willing to assert a structure preserving morphism between INS and "reality" we <u>would</u> not expect such a correspondence. The earlier quotes from Guetzkow and Hermann suggest that no structure preserving morphism is posited.

If we are not willing to assert a simple correspondence between INS variables and those of the real world, then the problem of measurement becomes critical. Measurement theory deals with how numbers can be associated with attributes or appearances of objects in such a way that the properties of the attribute are represented as numerical properties (see Krantz et al 1971). The problems of identifying the measurement structure necessary to discuss variable-parameter validity in highly aggregated simulations have, to my knowledge, not been discussed. My guess is that simple additive structures would not suffice. Yet without confronting the aggregation and measurement problems, it seems to me that the idea of variable-parameter validity cannot be applied to INS in any but an intuitive manner.

The fourth standard discussed by Hermann, event validity, uses " 'natural' events as criteria against which to compare outcomes occurring in the simulation (1967: 222)." For example, Hermann and Hermann (1967) used INS in an attempt to simulate the outbreak of World War I. The primary purpose of their study was to evaluate the validity of INS. "One means of investigating this question is to ascertain if a simulation produces events similar to those reported in a historical situation (1967: 401)." Thus one of the validity criteria being used was event validity.

However, for the reasons discussed under internal validity, such a use of event validity requires a commitment to a low variance external world. What if the "true" probability of the specific chain of events leading to World War I is .3? Would we then want to say that a simulation which "reliably" reproduced these events was valid? This same point is made by Verba (1964: 513) when he notes that stochastic "forces" may be operating both in the simulation and in the real world. Thus it would seem that high event-validity is neither necessary nor sufficient to having a "valid" simulation.

The fifth and last criterion discussed by Hermann is hypothesis validity." If X is observed to bear a given relationship to Y in the observable universe, then X' should bear a corresponding relationship to Y' in a valid operating model (1967: 223)." This criterion would seem to be identical to the behavior preserving morphism discussed earlier. Again, the use of hypothesis validity requires making considerable measurement assumptions. These assumptions should be explicated.

The last four of Hermann's five criteria are all concerned with assessing the descriptive validity of simulations. Each of these criteria deals with certain posited correspondences between the refereent system and the simulation. The strongest of these correspondences would appear to be the behavior preserving morphism of hypothesis validity. I have suggested that applications of these criteria to INS may suffer from a deterministic (i.e., low variance) view of "reality" and, given the aggregated nature of INS, a too simple view of measurement.

These criticisms of the methodological assumptions underlying the evaluation of the descriptive adequacy of INS suggest why INS has not been more successful as a guide to programmatic research in the development of scientific theories of international relations. By this I do not mean, of course, that INS has not served as a catalyst for a number of very important substantive studies in international relations. It most clearly has.

However one of the luxuries of the arm chair critic is to suggest what might have been done differently. My guess is that had there been more attention paid to questions of aggregation and the related problems of measurement as the simulation was being constructed, later researchers would have found it much easier to imbed specific research questions in the general INS structure. Some work we have done at Ohio State with Forrester's World Dynamics simulation strongly indicates that once a complex simulation (especially one with significant non linearities) is constructed, it is too late to disaggregate the concepts and relations to consider specific hypothses using simple measurement structures. The tendency then is to do experiments on the simulation itself and to compare only broad behavior patterns in the simulation with those in the referrent system. For certain purposes, of course, this is perfectly adequate. However if there is a need for a simulation to serve as a "complete processing model" with which to investigate the complexity of international relations, then there is a need to consider early problems of measurement and aggregation and disaggregation. Such a purpose would seem to be what Guetzkow had in mind when he wrote that simulations "...permit the coherent amalgamation of subtheories into interactive, hosistic constructions of great complexity (1969: 206)." Further he (Guetzkow, 1968: 210) clearly recognizes the measurement problems I have outlined.

Policy: A third general objective according to which INS might be evaluated is "policy." There are numerous ways in which INS has been relevant to the policy community. These include pretesting of alternative policies on INS, providing a monitoring system, and training policy makers to be aware of the complexities underlying the impacts of policies by involving them in INS. As many of the papers coming out of the INS project have pointed out, the "validity" of INS has not been sufficiently established to rely on it very heavily in the actual policy making process. Indeed, in order to so use INS we would probably want to be able to assert at least a function preserving morphism between the simulation and the referent reality. Yet, as was argued earlier, the major validation efforts have centered upon establishing only a behavior preserving morphism (i.e., preserving input-output relationships).

With regard to the policy objective, it might have been useful to have adopted a more explicit control theoretic structure. This would have made it far easier to address questions of system optimazation using existing techniques (e.g., Box, 1954, Box and Hunter, 1958, Draper, 1962, or Gardiner et al, 1959). Had this been done it would be (at least theoretically if not computationally) possible to investigate within INS:

(a) the relative importance of alternative policies, if different environmental conditions, or if differing parametric specifications as they affect the similar response at some point T in simular time; and,

(b) that set (or combination) of policies, environmental conditions, and parametric specifications which will provide, in some sense, the optimal similar response at time T (Mihram, 1972: 402).

Answering (or even posing) such questions would probably require reconceptualizing INS in response surface terms (this might not have been all that difficult since several of the "experimental designs" employed in INS reports approached doing this). More importantly, it becomes necessary to attach some sort of objective function to the simulation so that alternative outcomes can be ordered with respect to their "desirability." The problem of identifying such functions for social systems is most difficult (for an analysis of the problem see Raiffa, 1969). Yet if simulation is to be seriously used in policy <u>selection</u> it would seem that such problems must be addressed.

<u>Design</u>: Whereas with the policy objective the concern was with identifying and implementing feasible strategies to meet some goal(s), design problems deal with identifying and describing various mechanisms for the achievement of goals. The distinction I am making here between policy and design is analogous to the distinction between the values of variables (including parameters) and their structure. Policy changes are changes in the level of variables and design changes are changes in the structure relating the variables.

Simulation is a very powerful technique (i.e., means of generating data) for the design theorist. The low cost of computer computation makes it possible to examine numbers of complex mechanisms in a variety of "environments." Moreover, the use of simulation allows the investigator to deal with non-linearities, time lagged feedback and other complications numerically where analytic solutions may either not exist or, if they exist, be beyond the symbol manipulation skills of most behavioral scientists.

Notice too, a key distinction between the policy objective and the design objective. In the case of policy the representation (i.e., $m \in M$) is taken as being fixed. Our concern is with the impact of parametric and variable value (i.e., level) changes. However, in the case of design, the representation itself is the datum and our concern is with identifying desirable (perhaps in a constrained sense) elements of M. Thus experiments designed to optimize existing simulations will not generally serve the purpose of design as I have outlined it here (though, of course, it may give us an idea of the "best" that particular m can do).

Moreover, in the case of design, simulation can be quite properly be viewed as a data generating technique. The data generated are, of course, the performances of various m M. Under this view of design, questions of appropriate methodology again become relevant. Here, for example, we are less concerned with "correspondences" between m and s. Rather we are interested in developing preference orderings over the elements of M. As Guetzkow (1966) suggests, some of the methodological positions of "traditional" utopian thinkers may be applicable.

III Conclusion

The title notwithstanding, this essay has been more a critical review of selected aspects of the methodology of INS than it has an "appraisal." According to the O.E.D., to appraise may be rewritten "to estimate the amount, quality, or excellence of." The narrow scope of this essay does not permit an appraisal. Moreover, such a task requires a more experienced appraiser than I. However, and this may be more revealing anyway, I would be willing to bet that in the year 2000 it will be commonly acknowledged that the state of international relations theory owes a great debt to INS related activities. Already, in 1974, one cannot help but notice the number of rather sophisticated projects employing simulation (and employing it as a matter of course). Compare this to 1964 or 1954. Had Harold Guetzkow not begun the Simulated International Processes project, a large number of us would be unable (for a wide variety of reasons) to do what we are today. Assuming that we value that which we are doing, how can we do other than consider INS a success?

References

Abt, C., and Gorden, M.(1969) "Report on Project TEMPER," pp. 245-262 in D. Pruitt and R. Snyder (eds) Theory and Research on the Causes of War. Englewood Cliffs, N.J.: Prentice-Hall.

Alger, C. (1963) "Use of the Inter-Nation Simulation in Undergraduate Teaching," pp. 150-189 in H. Guetzkow, et. al, Simulation in International Relations. Englewood Cliffs, N.J.: Prentice-Hall.

Arbib, M.A. (1969) Theories of Abstract Automata, New York: Prentice-Hall.

Box, G.E.P. (1954). The Exploration and Exploitation of Response Surfaces: Some General Considerations and Examples, Biometrics 10, 16-60.

Box, G.E.P. and Hunter, J.S. (1957), Multifactor Experimental Designs for Exploring Response Surfaces, Annual Mathematics Statistics 28, 195-241.

Bremer, S.A. (1970) "National and International Systems: A Computer Simulation." E. Lansing, Michigan.: Ph.D. dissertation, Dept. of Political Science, Michigan State University.

Brodbeck, M. (1959) "Models, Meanings and Theories" in L. Gross (ed) Symposium on Sociological Theory , New York: Harper-Row.

Chadwick, R.W., (1967) "An Empirical Test of Five Assumptions in an Internation Simulation, About National Political Systems", in General Systems Yearbook, 12: pp. 177-192.

Crow, W. and Noel, R.(1965)"The Valid Use of Simulation Results," La Jolla, California, Western Behavioral Science Institute.

Draper, N.R. (1962). "Third Order Rotatable Designs in Three Factors: Analysis," Technometrics 4, 219-234.

Gardiner, D.A., Grandage, A.H.E., and Hader, R.J. (1959). "Third Order Rotatable Designs for Exploring Response Surfaces, Annual Mathematics Statistics 30, 1082-1096.

Guetzkow, H. (1959) "A Use of Simulation in the Study of International Relations." Behavioral Science 4 (July): 183-191.

Guetzkow, H., et al. (1963) Simulation in International Relations: Developments for Research and Teaching, Englewood Cliffs N.Y.: Prentice-Hall.

Guetzkow, H., (1966) "Comment on Professor Deutsch's Paper," in *A Design for Political Science: Scope, Objectives, and Methods*, J. Charlesworth (ed), The American Academy of Political and Social Science: Philadelphia, pp. 185-191.

Guetzkow H., and Jensen, L., (1966) "Research Activities on Simulated International Processes," *Background* Vol. IX, No. 4 February: pp. 261-274.

Guetzkow, H., (1968) "Some Correspondence Between Simulations and Realities in International Relations" in M. Kaplan (ed) *New Approaches to International Relations*, New York: St. Martin's Press, pp. 202-269.

Guetzkow, H., (1969) "Simulations in the Consolidation and Utilization of Knowledge about International Relations," in D. Pruitt and R. Snyder (eds) *Theory and Research on the Causes of War*. Englewood Cliffs, N.J.: Prentice-Hall. pp. 284-300.

Harrod, R.F. (1938) "The Scope and Method of Economics," in *Economic Journal* Vol. 48 September, pp. 383-412.

Hermann, C.F. (1967) "Validation Problems in Games and Simulations with Special Reference to Models of International Politics" in *Behavioral Science* Volume 12 May pp. 216-231.

Hermann, C., and Hermann, M. (1967) "An Attempt to Simulate the Outbreak of World War I," in *American Political Science Review* Vol. 61 June, pp. 400-416.

Krantz, et al., (1971) *Foundation of Measurement*, New York: Academic Press.

Kress, P.F., (1966) "On Validating Simulation: With Special Referance to Simulation of International Politics," Evanston, Illinois: Northwestern University Stenciled, p. 26.

Leavitt, M.R. (1971) "A Computer Simulation of International Alliance Behavior," Ph.D. dissertation, Dept. of Political Science, Northwestern University.

Leserman, (1972) "Simulated International Process Sipfil: Bibliography of Technical Outputs with Abstracts" Evanston, Illinois: Northwestern University.

McGowan, P. (1972) "Some External Validities of the Inter-Nation Simulation," Evanston, Illinois: Northwestern University, mimeo, p. 100.

Mihram, G.A. (1972) Simulation: Statistical Foundations and Methodology, New York: Academic Press.

Nardin, T. and Cutler, N. (1967) "A Seven Variable Study of the Reliability and Validation of Some Patterns of International Interaction in the Inter-Nation Simulation," Evanston, Illinois, Simulated Inter-national Processes Project, Northwestern University, December.

Naylor, T. (1971) "Computer Simulation Defined" in T. Naylor (ed) Computer Simulation Experiments with Models of Economic Systems, New York: Wiley p. 1-10.

Newell, A. and Simon, H.A., (1972) Human Problem Solving, Englewood Cliffs, N.J.: Prentice-Hall.

Newell, A. (1973) "You Can't Play 20 Questions with Nature and Win: Projective Comments on the Paper of this Symposium," in William Chase (ed) Visual Information Processing, New York: Academic Press.

Noel, R. (1963) "Evolution of the Inter-Nation Simulation," in H. Guetzkow, et. al., Simulation in International Relations, Englewood Cliffs, N.J.: Prentice Hall, pp. 69-102.

Raiffa, H., (1964) "Preferences for Multi-Attributed Alternatives," RM-5868 - DOT/RC, Santa Monica: The RAND Corp.

Raser, J.R., Campbell D., and Chadwick, P. (1970) "Gaming and Simulations for Developing Theory Relevant to International Relations," in General Systems Yearbook 15: pp. 183-204.

Richardson, J.M. (1974) "Quantifying 'Soft' Variables: Difficult but Essential," in Proceedings of the Fifth Modelling and Simulation Conference, Instrument Society of America: forthcoming).

Robinson, J. (1961) "A Research Design for Comparing Simulation with Case Studies and Problem Papers in Teaching Political Science," Proceedings, Conference on Business Games as Teaching Devices, New Orleans, April 26-28, pp. 123-129.

Robinson, et al., (1966) "Teaching with Inter-Nation Simulation and Case Studies," in American Political Science Review, 60 March, pp. 53-65.

Rudner, R.S. (1966) Philosophy of Social Science, Prentice-Hall.

Smoker, P. (1968) "An International Processes Simulation: Theory and Description. Evanston, Illinois: Simulated International Processes Project, Northwestern University.

Snyder, R.C., (1963) "Some Perspectives on the Use of Experimental Techniques in the Study of International Relations," in H. Guetzkow et all (eds) Simulation in International Relations, Englewood Cliffs, N.J.: Prentice-Hall, pp. 1-23.

Tarski (1944) "The Semantic Conception of Truth," in Philosophy and Phenomenological Research, IV, pp. 341-375.

Thorson, S.J., and Phillips W.R., (1974) "Simulation for Policy Planning," in Proceedings of the Fifth Modelling and Simulation Conference, Instrument Society of America, forthcoming.

Thorson, S.J., and Stever, J. (1974) "Classes of Models for Selected Axiomatic Theories of Choice," in Journal of Mathematical Psychology Vol. 11, No. 1 February, pp. 15-32.

Verba, S. (1964) "Simulation, Reality, and Theory in International Relations," in World Politics Vol 16, No. 3 April, pp. 490-519.

Zeigler, (1970) "Toward a Formal Theory of Modelling and Simulation I" University of Michigan.

Zeigler, B. (1971) "Toward a Formal Theory of Modelling and Simulation II", University of Michigan.