SDC

SP-3143

UPDATING SOME GROUND RULES FOR MAN-MACHINE SIMULATION

Robert J. Meeker and Gerald H. Shure

25 April 1968

SP-3143

SP *a professional paper*

Updating Some Ground Rules for Man-Machine Simulation*

Robert J. Meeker and Gerald H. Shure

April 25, 1968

SYSTEM
DEVELOPMENT
CORPORATION
2500 COLORADO AVE.
SANTA MONICA
CALIFORNIA
90406

*Research on which this report is based was supported by SDC and Contract Number DAHC15-67-C-0277, Advanced Research Projects Agency, U. S. Department of Defense.

A-1162 REV 4 65

Updating Some Ground Rules for Man-Machine Simulation

Robert J. Meeker and Gerald H. Shure

ABSTRACT

Laboratory gaming and simulation continue to be used as research and training tools of great value in many fields of social and organizational behavior. Discriminating researchers are aware, however, of the many fundamental problems that limit the valid applications of simulation studies. Laboratory simulation inevitably creates a problem of interpretation, i.e., of inferring correspondence between the simulation model studied in the laboratory and the "real" world. This is particularly true where the world created (simulated) in the laboratory attempts to simulate complex aspects of social reality. Traditionally, gaming for research and training has ignored many of these problems and dealt with others only superficially.

Behind the acceptance of this state of affairs is the almost axiomatic assumption that rigorous research and realism are inversely related, i.e., to gain an adequate basis for evaluation one must give up complexity in the laboratory; contrariwise, if realism is desired, research goals must be abandoned or sharply limited in scope. This paper takes the position that this assumption need to be reexamined in the light of new methodological and data analysis techniques. In developing our position we propose to show that the techniques and research designs borrowed from the highly abstract, sterilized, easily replicable experiments of the social scientist are, in part, responsible for the failure both to evaluate simulation methodology and to make significant research gains through simulation studies. With the availability of a large-scale, computer-based laboratory, techniques are now available for overcoming many of the serious limitations associated with the problems of data collection and analysis, particularly the difficulty of timely assessment and recording of data on large number of subjects in the precise detail and volume required; and secondly, the problem of usefully analyzing such an enormous and complex body of data. We propose to show that in our own research program the use of large-scale computer-based laboratory permits significant inroads to be made into methodological and validity problems of simulation, which remain relatively intractable under attacks of a smaller scale.

Updating Some Ground Rules for Man-Machine Simulation

Robert J. Meeker and Gerald H. Shure

Laboratory gaming and simulation are being used increasingly for research in many fields of social behavior, both for theory building and for exploring practical questions of policy and planning. But the validity and usefulness of such complex simulations are open to serious question, because of certain fundamental limitations. These are particularly noticeable where the world created in the laboratory attempts to simulate complex aspects of social reality; paradoxically, it is in these studies that the laboratory gaming approach is still considered, potentially, most useful. Thus the promise of simulation remains attractive even though there is an increasing awareness that it has been largely unfulfilled.

Behind the acceptance of this state of affairs is the almost axiomatic assumption that rigorous research and realism are inversely related. Complexity tends to be sacrificed to evaluation, or research goals abandoned for the sake of realism; thus, it appears that those drawn by the potentials of simulation must ultimately choose between the richness of validity and the constraints of tractability for analysis. To imply that the choice of the laboratory simulant narrows to an unwieldy verisimilitude or a sterile precision would be, of course, an overstatement; the distinctions are not that gross and the choice is rarely explicit, but there is undoubtedly a general polarization that tends to dichotomize gaming research into two camps.

For example, the International Simulation (INS) Game, one of the major examples of simulation for theory building, attempts to abstract critical system variables of the international scene and to model their interaction. It emphasizes operational definition and, where possible, the quantification of variables and the systematic recording of data. Replication of runs is used to demonstrate the stability of results obtained. In this game, players, generally students, take the roles of the leaders of interacting, hypothetical countries. INS games attempt to test a set of preformulated hypotheses taken from the political science literature.[1]

INS has been criticized on a number of grounds, a few of which are: The questionable significance of the system variables selected and the representation of their relationships; the omission of critical situational factors that operate in the real world (such as historical and geographical constraints); the use of naive, inexperienced players, driven perhaps by inappropriate motivation; and the unrealistic manipulation of simulated events for experimental purposes.[2] Because of these shortcomings, INS has not received seriously consideration by professional policy makers or advisers, either for theory building or as a basis for policy guidance.

Policy-oriented games, such as those conducted at RAND, attempt to avoid such pitfalls. Rather than being general, they tend to be intensive single-case studies of likely situations close to reality and involving real countries, as perceived by experts. Policy-oriented games typically use realistic political-military-geographical settings; they use seasoned, professional military and political experts as decision makers; they give attention to a wide range of relevant historical and current events; they use realistic action scenarios--the product of expert judgment--to guide the game; and they require players to prepare national-position papers, reviewed by a control team of experts who also manipulate the game in a way judged to be realistic for the simulated situation.[3]

In contrast to the INS example, the policy-oriented researcher is less concerned with testing theory than with gaining an understanding of decision-making processes that occur during particular situations, and he does not typically attempt to subject his detailed data to systematic recording and analysis. Data implicit in the player's position papers, the umpiring decisions, the player's values, perceptions, motivations, etc., in conjunction with given moves and the progress of the game, are not objectively assessed. The methodologies employed also are typically unevaluated. (For example, the conviction that expert players, rather than naive ones, are needed to achieve validity in such complex games, remains as much a matter of intuitive judgment as demonstrated fact.)

In summary, viewed from a research perspective, both the INS theory-building game and the RAND-type policy-oriented game suffer from fundamental limitations that appear critical for their further development and validity as research procedures. Neither game produces sufficient explicit data to allow a detailed (micro) analysis of ongoing behavior processes that may be essential for understanding the step-by-step development of the game, the predicted molar relationships, and the final results. What has not been fully acknowledged by INS researchers is that when a simulation game reaches a certain point of complexity, it is no longer feasible to attempt to demonstrate patterns of cause and effect exclusively by extensive replications of runs. This is true not only because of the prohibitively high costs of replication, but because of the sizable interrun variability that may be anticipated for even the most carefully executed and controlled complex game.

Both the theory- and policy-oriented approaches have their respective strengths and weaknesses; to combine the assets of both requires a reformulation of the data collection and analysis problem. Where extensive replication, alone, is not a feasible means of establishing reliability and validity of results (and even where it is), we believe it necessary to augment the data collected in the standard simulation experiment by embedding them in an extensive matrix of fine-grain observations. Thus we suggest that the simulation investigator, whether theory oriented or empirically oriented, will find it not only desirable to collect more data than are required to test preformulated hypotheses, but that it is necessary for him to do so in order to properly amplify or clarify his understanding of his data. Particularly when anticipated relationships among variables fail to materialize (or, stated less elegantly, when predictions are not confirmed), he will wish to check various possibilities among supplementary data that may account for his negative results. As has been already demonstrated for simpler gaming studies,[4] these details of response permit a more adequate reconstruction of the significance of the complex interplay of events and responses to these events. These systematically sampled data points allow a more adequate mapping of the micro processes and demonstration of patterns that underlie the predicted and unpredicted molar relationships.

Having stated these extensive demands for data collection, we also hold that in those simulations of social systems where human decision-making behavior is a critical, if not a central, determinant of the performance of the system,[5] it is not sufficient merely to record the overt actions taken; these data must be supplemented by the subjectively perceived significance of events and the decision maker's intentions. Elsewhere,[6] we have attempted to characterize the decision-making behavior under conditions where the problem situations develop piecemeal and spread out over time; and we have demonstrated the increased understanding of the decision process that is obtained if one can also obtain the immediately held, moment-to-moment perspective of the decision maker. While most individuals would be willing to grant the relevance of such data, the problem is to find an efficient means of obtaining these subjective data that permits us to establish the "presently actualized" perspective of decision makers. Where exercises or operations extend over any sizable period of time, and where a large number of events require decisions, the ability of the decision maker to reconstruct at a later time his original and altered assessments for even moderately complex situations is necessarily limited and is very likely to undergo modifications with the passage of time and changing configurations of circumstance.

In an attempt to solve this problem, we use the speed and efficiency of the computer to collect, on-line, both actions taken and associated introspective data from decision makers as they are engaged in simulation experiments. In these studies the computer simultaneously monitors 24 participating subjects, detects all situations about which further information is desired, and then selectively (and with minimal disruption) displays questions relevant to what has just happened to each subject. The subject records his answer in a form the computer can store and interpret for subsequent analysis.

These procedures permit us to ask the decision maker what he is doing at times closer to the moments of critical significance for the decision; and as a result, these procedures considerably minimize loss and distortion of pertinent information that may reside only momentarily in the head of the decision maker. The difficulties of obtaining and interpreting these kinds of "subjective" data are necessarily very great and many problems of collection and interpretation remain unsolved. Nevertheless, the procedures have already yielded decisive information in a number of laboratory studies of negotiation processes.[7]

If we grant the logic of our argument--that for optimal utilization of simulation research, both more and greater variety of data need to be collected and that the on-line use of the computer to run an experiment can now provide the experimenter with the ability to record more details of experimental processes and finer gradations of response than hitherto possible--we still have another major hurdle. What do we do with this presumably rich data base that has been so ably generated and recorded in our computer simulation? The problem is further compounded by the nature of the data: Typically, they are time-ordered or sequenced-ordered, irregular in occurrence, frequently redundant or irrelevant, hierarchical, and of variable length and format.

The task of classifying, grouping, and summarizing these data so as to identify summary and configurational indices that attempt to satisfy the criteria of reliability, specificity, validity, and relevance requires an interplay of intuition and trial-and-error that can become exceedingly costly in time and effort. With even modest amounts of data, the task of iteratively sharpening and improving variable definitions can become impossibly burdensome.

The brute-force attack, even with available statistical computer programs, is rarely a satisfactory way of exploring the abundant, computer-recorded data. In a typical experiment of our own, we have had as many as 1,000 items of information for each of hundreds of subjects. If we did nothing more than calculate intercorrelations for each of these items of data, without considering combined indices, we would generate approximately one-half million correlation coefficients or cross-tabulations. The resulting stacks of computer printouts for such analyses are not infrequently measured in lineal inches. Techniques such as factor analysis are not particularly effective for inducing or discovering a fundamental order in the data, especially where reassessments frequently call for dividing the data in different ways, for omitting various subclasses of data from comparisons, for using different observational data or values in operational definitions, and, generally, for a great deal of data manipulation (recombining, regrouping, and recalculating) before new assessments can be made.

These kinds of data analysis problems require a new sort of computer assistance--data manipulation programs that are designed to aid the inductively oriented researcher in exploring relationships that may obtain among complex sets of data. We have undertaken the development of two such programs, TRACE--Time-Shared Routines for Analysis, Classification and Evaluation,[8] and IDEA--Inductive Data Exploration and Analysis.[9]

While a detailed account would not be appropriate to our present purposes, these programs (and more centrally, TRACE) are sufficiently novel in intent and operation that a brief description is appropriate in order to appreciate their special utility for simulation research.

TRACE differs from typical data management systems in that the primary objective is to derive new variables from existing ones, rather than being concerned primarily with data retrieval and/or summary reporting. It differs from standard analysis programs in the magnitude and complexity of its data manipulation capability in that it is primarily logical and algebraic, rather than statistical; and in that the system assumes the responsibility for data base construction and maintenance.

TRACE is operated conversationally and performs the desired data manipulations--recombining, regrouping, and recalculating--in a time-shared mode of computer operations, with the use of computer-connected teletypes and CRT displays. The on-line capability of the program permits immediate feedback about the relative utility of newly derived indices, so that the investigator may modify the steps in the analyses as he goes along. TRACE thus permits an effective interplay between the investigator's conjectural and judgmental skills and the computer's capacity for rapid and accurate data processing.

For a figurative understanding of the program's operations, one might think of it as a sophisticated automated data clerk who has access to all the experimental records and is prepared to calculate, count, sort, classify, cross-tabulate, generate simple statistical indices, retrieve information and keep records according to the user's directions and specifications. In conversing with the program, the user refers to variables by name with the structure of his data base implicitly understood; thus in instructing the program he can ask directly for, say, "the average decision-making time that each chief of state used in resolving all domestic type problems." The capability for direct questioning presupposes, of course, that the constituent variables are well defined within the data record; but this requirement is substantially relaxed in actual use, since the program provides the user with the capability of using variables that have been previously derived.

For any derivation, then, the user can first of all specify the level of aggregation or grouping of the data along any dimension of the data record--thus, in the example noted previously, instead of each chief of state, the user might, optionally, have selected each decision maker, each nation, or each alliance bloc as the unit of analysis. Secondly, the user can restrict the data in an analysis to any specifiable subset--thus, in the example, the restriction to domestic problems might, optionally, have been eliminated and all problems, or foreign problems or the first ten problems encountered, or any other subset of special concern in the analysis. Thirdly, the user has considerable latitude in terms of the complexity of the functions to be performed--in the example, the request for a simple mean might be extended to a derivation request of 100 terms defining a new measure whose values will be based on logical and arithmetic combinations of values obtained from a set of existing variables. These combined rules permit operational definitions to satisfy a variety of contingent conditions that permit highly complex search patterns.

Subsumed under the derivation procedures is another program feature of particular utility for researchers with time-ordered and sequenced data: Variables can be treated as a string of symbols that can be subdivided in analysis. Each subdivision can be treated as a separate value, or elements in the string can be aggregated into larger units, and indeed the entire string can be treated as a single datum. This capability permits the derivation program to be used as a device for detecting sequence patterns in the data--by sliding a view-box or a "window" across successive sets of entries and sensing for particular patterns.

Once variables have been derived, their efficacy can be quickly assayed. Subprograms permit the user to examine the distribution of the newly derived variables and to relate each to other criterion variables. One such program offers a number of criteria for partitioning each variable into intervals or classes. Another provides a means of evaluation--to see whether there are anomalous distributions that will require redefinition or adjustment; or, in the bivariate distribution, to make preliminary assessments of a measure's descriptive or predictive potential. A number of statistical indices and simple procedures for identifying outliers are available along with a number of other features for aiding the user to augment and check his insights and further refine his measures as he reexamines his data.

What, in effect, we are saying throughout this paper is that the usefulness of simulation studies of complex social environments with human participants can be greatly if not fundamentally increased if we recognize the following facts:

1. More detailed information is required (on the states of the environment, and on the behavior, both objective and subjective, of the players in that environment)

2. The ability to implement a greatly expanded inductive approach is now technically feasible through the use of the computer as a data collection tool

3. The associated opportunity to "pray over" extensive data sets is also available

4. Reanalyses of data collected in simulation studies, like those collected in the physical science experiment, are not adjunct procedures to be employed as back up procedures but are central aspects of research strategy.

Let us quote John Tukey on this issue of reanalyses and the conception of experimental results as indicators rather than conclusions.

> "We need to face up to the need for iterative procedures in data analysis. It is nice to plan to make but a single analysis, to avoid finding that the results of one analysis have led to a requirement for making a different one. It is also nice to be able to carry out an individual analysis in a single straightforward step, to avoid iteration and repeated computation. But it is not realistic to believe that good data analysis is consistent with either of these niceties. As we learn how to do better data analysis, computation will get more extensive, rather than simpler, and reanalysis will become much more nearly the custom."[10]

When we suggest, then, that the ground rules for simulation research should be updated, we are, in effect, suggesting that the issues that opened this paper should be reexamined in the light of recent methodological developments. We do not offer any panacea or starkly simple solution; on the contrary, our alternatives share all the characteristics of working principles--not entirely polished and not tightly systematic--but, for all that, they have the working principles' saving grace of being grounded in experience. What we propose then is nothing as grand and formal as a separatist philosophy of science for simulation researchers, but rather a program for incorporating new means for achieving rigor while retaining a commitment to a meaningful degree of realism.

REFERENCES

1. Guetzkow, H., Alger, C. F., Brody, R. A., Noel, R. C., and Synder, R. C.; "Simulation in International Relations"; Englewood Cliffs, N. J.: Prentice-Hall, 1963.

2. In particular, the validity of interrelationships of programmed variables in the INS model has been criticized (see Chadwick, R. W., "An Empirical Test of Five Assumptions in an Internation Simulation, About National Political Systems", General Systems, XII, 1967, 177-192; Pendley, R. E., and Elder, C. D., "An Analysis of Officeholding in the Internation Simulation in Terms of Contemporary Political Theory and Data on the Stability of Regimes and Governments", unpublished manuscript, Northwestern University, November 1966, and others); the unrealistic manipulation of simulated events in INS also has been criticized (see Rosencrance, R. N., and Mueller, J. E., "Decision-Making and the Quantitative Analysis of International Politics", in London Year Book on World Affairs; and DeWeerd, H. A., "Political-Military Scenarios", paper presented at a conference on Modern Strategic Analysis, held by Security Studies Project of UCLA on January 13, 1967, available as RAND document P-3535).

3. Goldhamer, H., and Speier, H.; "Some Observations on Political Gaming"; Behavioral Science, 4, 1959, pp. 183-191.

4. Meeker, R. J., and Shure, G. H., "Real-Time Computer Studies of Bargaining Behavior: The Effects of Threat Upon Bargaining", AFIPS Conference Proceedings, 1964 Spring Joint Computer Conference, pp. 115-123; and Shure, G. H., Meeker, R. J., and Hansford, E. A., "The Effectiveness of Pacifist Strategies in Bargaining Games", Journal of Conflict Resolution, 1965, 9, pp. 106-117.

5. There obviously are many simulations in which the human participants are not of intrinsic interest, but are employed as an economic substitute for computer subprograms where the latter are too expensive or difficult to write; under these conditions, it is assumed the variation in the behavior of the human participants remains within acceptable bounds so that their performance does not distort the evaluation of other components or the system as a whole.

6. Shure, G. H., and Meeker, R. J.; "Probing Behind the Human Decision"; Proceedings of 15th Military Operations Research Symposia (MORS), 1965, 35-39.

REFERENCES

7. Shure, G. H., Meeker, R. J., Moore, W. H., Jr., and Kelley, H. H.; "Computer Studies of Bargaining Behavior: The Role of Threat in Bargaining"; SDC document SP-2196, February 9, 1966, 172 pp.

8. Shure, G. H., Meeker, R. J., and Moore, W. H., Jr.; "TRACE--Time-Shared Routines for Analysis, Classification, and Evaluation"; in Spring Joint Computer Conference Proceedings, 1967, 525-529.

9. Press, L. I., and Rogers, M. S., "IDEA--A Conversational, Heuristic Program for Inductive Data Exploration and Analysis", in the Proceedings of the 22nd ACM Convention, 1967, 35-40; also Shure, G. H., Press, L. I., and Rogers, M. S., "Man-Computer Derivations of Tree Structures from Multivariate Data", American Psychologist, 1967, (in press).

10. Tukey, J. W.; "The Future of Data Analysis"; Annals of Mathematical Statistics, 33, 1962, 1-67.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author) System Development Corporation Santa Monica, California	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3 REPORT TITLE

Updating Some Ground Rules for Man-Machine Simulation

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

5 AUTHOR(S) (First name, middle initial, last name)

R.J. Meeker
G.H. Shure

6 REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
25 April, 1968	12	10

8a. CONTRACT OR GRANT NO. ARPA DAHC15-67-C-0277	9a. ORIGINATOR'S REPORT NUMBER(S) SP-3143
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10 DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

11 SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13 ABSTRACT

This paper proposes to show that the techniques and research designs borrowed from the highly abstract, sterilized, easily replicable experiments of the social scientist are, in part, responsible for the failure both to evaluate simulation methodolgy and to make significant research gains through simulation studies. With the availability of a large-scale, computer-based laboratory, techniques are now available for overcoming many of the serious limitations associated with the problems of data collection and analysis, particularly the difficulty of timely assessment and recording of data on large number of subjects in the precise detail and volume required: and secondly, the problem of usefully analyzing such an enormous and complex body of data. We propose to show that in our own research program the use of large-scale computer-based laboratory permits significant inroads to be made into methodological and validity problems of simulation, which remain relatively intractable under attacks of a smaller scale.

DD FORM 1 NOV 65 1473

Unclassified

Security Classification

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Man-machine simulation						
Computer-based laboratory						
Laboratory gaming						
International Simulation Game (INSO						

Unclassified
Security Classification