UNCLASSIFIED

63-2-5

[redacted]

(SP Series)

SP-1047

Arms Control: The Search for an Acceptable Research Methodology

by

Robert H. Davis

17 December 1962

SYSTEM DEVELOPMENT CORPORATION, SANTA MONICA, CALIFORNIA

PORTIONS TO BE PRESENTED AT THE INTERNATIONAL ARMS CONTROL SYMPOSIUM, THE UNIVERSITY OF MICHIGAN, ANN ARBOR, MICHIGAN, DECEMBER 17-20, 1962

TABLE OF CONTENTS

I. INTRODUCTION

For centuries men have been exploring, without great success, alternative approaches to the control of arms. Here and there one can point to apparently lasting and significant arms control measures such as the Rush Bagot Agreement of 1817 which disarmed the Canadian frontier. In spite of these efforts, very little real progress has been made. War and the threat of war are still major instruments of foreign policy.

Sensible men might have despaired and given up the search for an alternative to war long ago if the consequences were any less frightening. Despite the air of pessimism and doubt which surrounds the field, scholars from virtually every academic discipline have given the matter their consideration and have voiced their opinions. Whether driven by guilt or simply the nagging conviction that there must be a solution, scientists have been particularly active in the quest for a stable peace.

While this interest is both encouraging and gratifying, the concrete results of our efforts have not been particularly noteworthy. In view of the frustration and sense of futility which sometimes sweeps over those of us devoting our time to the study of arms control, it is not unreasonable to ask why we have not done better. If we can isolate the sources of our failure, it may be possible to devise remedies.

One source of failure is, of course, rooted in the nature of man and the societies he has built. Since the earliest times, the contribution of social and psychological forces to armed conflict has been recognized, and any discussion of this aspect of the problem falls outside the scope of this paper. A second source of failure is more germane; it seems reasonable to suggest that our failure to find solutions may be attributable, in part at least, to our inability to develop acceptable and meaningful methods for studying the problem of arms control.

While I do not wish to minimize the contribution of scientists and scholars to the organization of this field, particularly in the past decade, the hallmarks of scholarship and science are frequently missing from arms control proposals--the rigorous attention to detail, the exploration of past relevant research, the collection of supporting data, the construction of detailed research designs, and the judicious application of appropriate methods. Instead, one often finds apparently serious proposals for solving these momentous problems in "Letters-to-the-Editor" columns, in three- or four-page articles, and even in paid advertisements. Under the phrase "acceptable methods," therefore, we should include the concept of discipline--for discipline would appear to be at the heart of the matter.

Surely our scientific training and tradition today does not support an undisciplined approach which leaves all the "engineering details" to others. If the problem of arms control could be divorced from technology--as in the case with membership in a church or political affiliation--then the failure of some of us to attend to engineering details would be understandable. But the primary justification for the scientist being on the scene at all is that he, presumably, has more to offer than just another suggestion. Out of his unique training and experience comes an appreciation for the complexity of nature and the need for careful control and research design. The challenge before us is not to offer miscellaneous suggestions, but to match the skills of the scientists to the problems of arms control.

II. SOME EXPERIMENTAL CONSIDERATIONS

Our search for a satisfactory match between the scientist and the problem of arms control should begin, I believe, not with the skills of scientists but with the nature of reality. It is a commonplace observation that the world we live in is enormously complex. Where lawful relationships do exist, they are frequently disguised from the casual observer. We seek to unravel the simplest thread only to find that it binds together an enormously complex array of interdependent events. The myth that the world might be otherwise has been exposed by many philosophers including Leibnitz (17), and Whitehead (23); by psychologists, notably Koffka (14), Kohler (15, 16) and Lewin (18); by mathematical physicists such as Einstein (6); and by mathematicians such as Wiener (24).

Today, many of us subsume these ideas under the concept of a system, and treating reality from a system viewpoint helps one to appreciate the complexity of the world. How should we go about studying such complex phenomena? There is much to be learned by considering the experimental methods of both the physical and social scientists, particularly techniques for controlling error variance.

In studying a given system, there is always the danger that we will change it into an entirely new and different system by contaminating it with variables which are not normally a part of it or by abstracting away from it essential elements. Contamination of the experimental environment is a commonly recognized danger. The danger has been recognized by most physical and many social scientists. In physics, for example, the problem was clearly demonstrated by Heisenberg (9); in social and industrial psychology, it has also been demonstrated and sometimes called the Hawthorne Effect (21). The control of experimental variables is nothing more than the effort to preserve and maintain the integrity of essential system elements so that meaningful data about a system can be collected and reasonable inferences drawn.

As the complexity of the system being studied increases, experimental control in the most rigid sense becomes impossible. Two kinds of steps are then taken in an effort to preserve the system for controlled observation: <u>aggregation</u> and <u>distillation</u>.

Aggregation is the term commonly used to describe the clustering of a number of factors into a single selected variable. Thus, for example, Clark Abt (1) has recently described an arms control game played on a computer in which a very large number of crucial variables were aggregated. According to Abt, his strategic model consists of the following six elements:

1. Major actors participating in global interations; Western, Eastern (Communist) and Neutral blocs.

2. Actor capabilities to modify themselves and each other; military, economic, diplomatic, psychological, scientific, cultural, and idealogical.
3. Actor motivations; psychological, cultural, ideological.
4. Actor decision-making; ideal-sensing, reality-sensing, ideal-to-real discrepancy-measuring, resource-allocating, and inter-bloc bargaining elements.
5. Master time control coordinating the dynamic interactions of the operative elements.
6. Master geographic control coordinating the spatial dynamic interactions in 15 global areas.

In addition, Abt's game includes "submodels" for military, ideological, cultural, psychological, and other factors. If one seeks to predict real-world events by techniques of this kind, the assumption that political and military events can be meaningfully studied when aggregated into all-inclusive variables of this kind is, of course, subject to considerable argument. Whatever the merits or demerits of this approach may be, the method of aggregation is widely used.

The second method I shall call distillation. It consists of systematically removing variables in order to create an abstract environment. It is the classical method of 19th Century physics. Reichenbach (20) describes it as follows: "As long as we depend on the observation of occurrences not involving our assistance, the observable happenings are usually the product of so many factors that we cannot determine the contribution of each individual factor to the total result. The scientific experiment isolates the factors from each other; the interference of man creates conditions in which one factor is shown at work undisturbed by the others, thus revealing the mechanism of the complex occurrences that happen without man's interference. The falling of a leaf from a tree, for instance, is a complex occurrence in which the gravitational force competes with aerodynamic forces resulting from the flow of air under the sliding leaf, which moves down in a zigzag course. If, on the one hand, we exclude the air by letting the leaf fall in an evacuated space, we see that with respect to gravitation, its fall is the same as that of a stone... By means of the artificial occurrences of planned experiments, the complex occurrence of nature is thus analyzed into its components."

Using distillation, isolated independent variables can be manipulated and their interactions determined. Unfortunately, the frequency of interactions between political, social, and psychological variables is so great as to place a limit on the usefulness of this approach for the experimental study of any socio-psychological processes.

Psychologists are particularly fond of applying this method to the study of human behavior. An example taken from a recent article by de Rivera, published in the American Psychologist (5), illustrates the nature of this method and its weaknesses as well. In this article de Rivera describes the following experiment undertaken by two of his students.

Consider the following situation. Two subjects are seated at a table. They each have a stack of chips representing their military resources. In the center of the table is a third stack of chips representing available arms. Subjects may take a chip from this third stack or discard one into it. Each subject also has a "red" button. The experimental conditions are as follows:

1. If either subject pushes the "red" button, the one with the fewer chips loses and must spend an amount of extra time in the laboratory equal to the number of chips remaining. Punishment is thus directly proportional to the amount of armament held by the loser, and an arms race is encouraged by the fear that no matter how many chips a subject has, he may still lose.
2. A loud, ticking clock constantly reminds subjects of another variable, i.e., "time is running out." The game may end at any minute, and when it does, both sides lose and must devote extra time in the lab in proportion to the number of chips each has remaining. The inevitability of "disaster" as a function of time encourages "disarmament" or the casting-out of chips.

In reporting the results, de Rivera states: "Of 12 pairs of subjects, only 2 pairs managed to disarm without a war. Most subjects went into an arms race. If one subject tries to disarm gradually, the other often becomes suspicious and pushes his button as soon as he is ahead."

This experiment is a good example of the use of distillation in the social sciences. It is a method which allows reasonably tight control of variables and is inexpensive. On the other hand, it is difficult to relate these results to the "real" world except by analogy.

Distillation and aggregation may be used in combination. But whichever of these two techniques is used singly or in combination, the ultimate question is still the same: Has the "real" system been so violated by the experimental technique that there is not a reasonable degree of isomorphism between model and reality? Or in simpler terms, does the system being measured have sufficient relationship to the one in which we are interested to allow us to make useful predictions about the real world?

The answer to this question is commonly assessed by determing (where possible) the reliability and validity of the experimental technique used. Establishing reliability involves determining the extent to which a technique provides consistent measurements. Establishing validity involves the determination of the extent to which a technique measures what it purports to measure. A technique may enjoy great reliability but very little validity. A technique is never completely reliable or unreliable, completely valid or invalid; they are only more or less reliable, and more or less valid.

III. FIVE RESEARCH TECHNIQUES

Although it may seem incongruous, it is gradually becoming apparent that methods currently being used to study war may profitably be applied to the study of the control and prevention of war. Problems of peace and war, military and political strategy, have many characteristics in common including complex interactions of elusive social-psychological variables, experimental control difficulties, etc. During the past decade, various military organizations have attempted to apply new techniques to the solution of old problems. While most of these techniques are firmly rooted in the past, going back at least to the 17th Century, recent developments have had a marked effect on the possibility of their more general applicability, particularly the use of high speed digital computers. Let us examine both the techniques and their historical antecedents in an effort to isolate some of their advantages and disadvantages.

In addition to the more traditional experimental methods of which de Rivera's study is an example, there are at least five general techniques available today for the study of such complex social, political, and economic problems as those of war and peace. These are:

1. Individual and Group Planning
2. Scenarios
3. Crisis Games
4. Symbolic Simulation
5. Environmental Simulation

Individual and Group Planning. Prior to the 17th Century, strategic planning was carried on by individuals and/or groups. It is a commonly noted fact that individuals and organized groups tend to be partisan; their perceptions are determined by their past experience, motivations, and social groups; and it is often difficult for them to project themselves into the role of others. Indeed, under some circumstances, projection is very dangerous. For the individual, "thinking like one's opponent" threatens the logic of one's own position and disturbs the entire fabric of one's thought. Individuals, therefore, tend to reorganize their ideas to fit their existing modes of thought and not vice versa. Projection may be dangerous in the group situation as well; the individual who thinks like "the opponent" in the presence of his peers may be seen by those around him to be in sympathy with outsiders. Thus, individual and group planning is often controlled by psychological and social pressures which encourage the selective and biased consideration of particular ideas.

One strategy for minimizing such selectivity is for the group to encourage "role playing" and for the individual to force himself to see the world from the point of view of others. Role playing is at best only a partial solution for its success depends primarily upon our ability to see ourselves as others see us and to see others as they see themselves--a very difficult thing to do.

A second problem with individual and group planning is the inability of individuals to clearly visualize alternatives and constraints. In the early history of war gaming, this problem led to the development of chess-type boards and sand tables. Today, the need to visualize alternatives has resulted in the construction of complex models (both physical and abstract) for simulating force structures, economies, etc. Let us consider for a moment the evolution of some of the early ideas which influenced the subsequent development of war, political, business, and peace games. (See Figure 1.)

Efforts to motivate officers to "role play" realistically led to the development of other aids for lending realism to games. Elaborate rules were developed for the calculation of such factors as troop movement, fatigue, effects of fire, and casualties. At first, many of the calculations were purely theoretical, but gradually the inventors of new games turned to empirical data derived from actual battles to construct their tables. Games following this tradition generally fall into the class known as Rigid Kriegspiel. In view of the lack of sophisticated data processing equipment available at the time, it is not surprising that Rigid Kriegspiel was seldom played according to the rules and was frequently disliked by officers required to play it. In describing Livermore's American Kriegspiel, for example, Sayre (22) states: "It may be confidently stated that Colonel Livermore's system is the best of its class, but it cannot readily and intelligently be used by anyone who is not a mathematician, and it requires, in order to be able to use it readily, and amount of experience, instruction, study, and practice about equivalent to that necessary to acquire a sure knowledge of a foreign language."

Its critics pointed out that the rules were not only tedious and difficult to learn, but caused inordinate delays in the conduct of the game so that both continuity and interest were lost. Many of the objections to Rigid

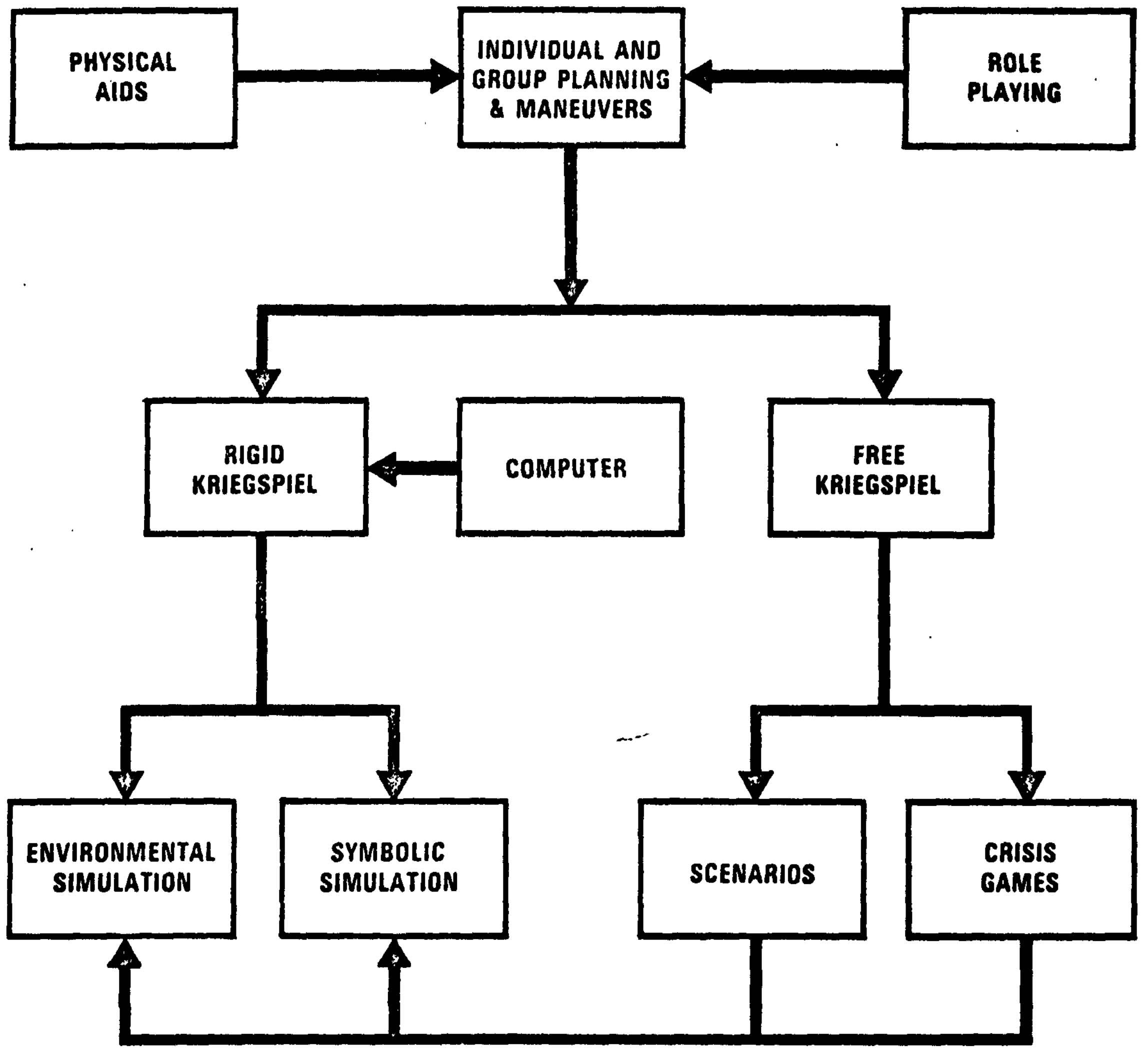

FIGURE 1. Some Alternative Methods for Studying Arms Control and their Antecedents

Kriegspiel have now vanished as a consequence of the development of the general purpose digital computer, and computerized war games tend to be in the tradition of Rigid Kriegspiel. New objections, however, have risen in their stead, having to do with the assumptions and values buried in the programs of the computer.

Free Kriegspiel, which is generally identified with Meckel and Von Verdy, was a protest against the formality of Rigid Kriegspiel. Its proponents argued that Rigid Kriegspiel was "too great a strain on the patience." Free Kriegspiel relied on the judgement of experienced officers for the conduct of the game and decisions regarding rates of troop movement, casualties, etc. Its opponents pointed to the inevitable problems associated with getting reliable and valid judgments under such conditions, but Free Kriegspiel took hold and was the more widely played of the two until after World War II.

These two streams of thought (Rigid and Free Kriegspiel) have led to the development of the remaining four approaches to the study of strategic planning. In discussing these, the distinctions and similarities between the Free and Rigid Kriegspiel are well worth keeping in mind. Arguments for using empirical data and fixed decision rules parallel those of the Rigid Kriegspiel; arguments for employing experts, however, do not have to do with expediting the game since, as indicated, such arguments are no longer as valid as they were before the invention of the digital computer. Instead they are related to the general problem of preserving the range, complexity, and richness of variables when problems of this scope are studied.

Scenarios. Even if one is able to achieve a degree of objectivity and see the world as his opponent sees it; and even if one is able to visualize most of the alternatives, a third serious problem presents itself for the individual and group planners. Namely, what conditions and constraints determine the point of departure for planning and the subsequent play of the game? How is the universe of discourse to be controlled? One very effective way of doing this is through the use of scenarios. Scenarios are often used as game aids or they may be used as a self-contained technique for studying a particular problem. As aids, they provide a point of departure and focus the attention of participants on a particular set of problems throughout the problem-solving period.

But in addition, the task of writing a detailed scenario can, in and of itself, be a very useful and productive exercise. Herman Kahn has described two scenarios of this kind in his most recent book (12). When seen as an aid to thought, particularly for the development of new hypotheses which might not otherwise be noted, the technique can be quite useful. Like the more sophisticated games in which they are often used, scenarios raise numerous questions regarding the interactions of complex social, psychological, political and military factors. Because a scenario describes a hypothetical series of events, possible consequences of a given course of action become more apparent and planners of events are forced to deal with problems which, for the social and psychological reasons previously discussed, an individual for a group might other wise avoid. The violent and emotional uproar created by Kahn's efforts to trace out in this way the consequences of various kinds of conflict in his book, "On Thermonuclear War" (11), illustrates how powerful these social-psychological forces can be when accepted modes of thinking are seriously disturbed by unconventional ideas.

Crisis Games. Individuals acting alone can develop scenarios. Crisis games, on the other hand, require two or more "actors" who play the role of aggregated real-life counterparts, such as the U.S.A. or the U.S.S.R., NATO or the Warsaw Pact Countries. A key element in crisis gaming is role playing. At the simplest level, no effort is made to simulate the environment in detail.

Both the Germans and Japanese are known to have used crisis games prior to World War II. Eric von Manstein reports in some detail on a political-military game played by German officers and civilian diplomatic officials prior to World War II. Von Manstein is quoted by Goldhamer and Spier (7) regarding the value of the game: "We had the impression also that the gentlemen from the foreign office to whom such playing-through of possible conflicts seemed to be completely novel, were thoroughly convinced of the value of the game."

Rear Admiral Zacharias, who lived for many years in Japan prior to World War II, reports that similar crisis games were played by Japanese military leaders (25), and further evidence regarding the conduct of such games was introduced at the Tokyo War Crimes Trials.

One of the most serious problems in connection with such games lies in the efforts of individual actors to portray highly aggregated entities such as nations, military blocs, or even nature. This criticism, as we have already noted in connection with the work of Abt (1), can also be leveled against similar mathematical games played without human participants. Advocates of the analytic approach sometime cite their clearly defined assumptions as almost _prima facie_ evidence for the superiority of mathematical and computerized games over crisis games employing human actors. The belief that the validity of one approach _vis a vis_ another can be decided by "the elegance of the method" is not particularly compelling.

Furthermore, one should not underestimate the degree of precision demanded by crisis games. When experts are assembled together and given sharply defined roles, explicit statements of their assumptions will be required to defend their arguments and actions (7). Out of these confrontations frequently come new ideas and hypotheses which may later be examined by individual or group analysis.

Symbolic Simulation. Simulation for the purpose of studying political and social phenomena involves the construction of models and the movement of models through time. When one uses _symbolic_ simulation, the models are either logical or mathematical. If the construction of models for symbolic simulation involves the incorporation of strictly mechanical and stochastic processes, precise definition of variables is generally possible. It may, however, involve symbolic representations of human behavior or human organizations in which case the problem is far less manageable. Although it is very difficult to establish criteria for determining when symbolic simulation is appropriate for studies of this kind, something can be said about when it is not appropriate to use symbolic simulation.

However powerful the method or machines, real-world alternatives cannot be predicted or evaluated if we are unable to define precisely the relevant variables and their interactions, and operationally relate them to the real world. Cause and effect relationships must be known. Human interactions pose a particularly serious problem in this regard. Where human behavior in a system is routinized and rigidly bounded, the use of symbolic simulation may be appropriate. To the extent that human behavior is not so routinized or bounded, it becomes less predictable and doubt arises regarding the utility of symbolic simulation.

It is sometimes argued that symbolic simulation provides new "insights" for those involved in the construction and employment of the models, and it seems reasonable to suppose that this is, indeed, the case. If this is the purpose, however, the criteria for evaluating symbolic simulation change. If the objective is "training," rather than predictive efficiency, one would want to know, for example, who is being trained, whether transfer willbe positive or negative, and whether this method is more efficient than others.

Environmental Simulation. When the de Rivera experiment was described (Section II), it was emphasized that experiments of this kind used a process called "distillation" to reduce the number of variables being studied in the hopes of isolating general relationships. With respect to the distillation process, it was also stated that interactions in human affairs are so pervasive as to place limits on the usefulness of this approach for developing generalizations about the real world. While the classical experimental method involves distillation or the abstracting away of complexity, environmental simulation involves efforts to preserve the complexity of the environment. In environmental simulation, the experimenter seeks to maintain the richness and variety of the environment and allows human participants to respond to a situation which resembles as nearly as possible "real-life." Both the more traditional psychological experiment of de Rivera and environmental simulation involve measuring decisions made by humans--one in an abstract environment which bears little resemblance to the real-life situation, the other in a more complete replica of its real-life counterpart.

There are a number of examples of this kind of simulation including studies by Guetzkow (8), Kennedy (13), and Chapman et al (3). An arms control study using environmental simulation was conducted by Carpenter and the author some months ago.

This study highlighted some of the differences between the more traditional experimental approach, as exemplified by de Rivera's experiment, and environmental simulation.[1] It was designed to manually test the feasibility of building a more complex computerized version of the arms control game. But before describing the game in detail, it may be worthwhile to pause and consider the kind of problem it is designed to study.

Both the U.S.A. and U.S.S.R. are required to make regular decisions regarding the allocation of their resources. In doing this, each nation responds to intelligence about the other. Any arms control agreement signed in the near future would presumably contain provisions for changes in the behavior of both countries (arms controls) and a means of determining whether such changes had indeed taken place (inspection). Numerous information channels would provide the two nations with information about one another. The treaty would probably allow some degree of formal inspection, perhaps by on-site observers and/or by aerial over-flights or other techniques. In addition, intelligence would be received from more informal and sometimes

1. The description of an arms control game which follows is taken in large part from a paper by P. B. Carpenter and R. H. Davis entitled, "Arms Control Simulation: An Empirical Approach." This paper was presented at the meetings of the American Psychological Association in September 1962 at St. Louis, Mo. In an effort to construct an economical version of the game, many of the required models were greatly simplified; details of a more complete and somewhat different version of this game may be found in Davis et al (4).

even routine channels such as attache reports, newspaper articles, tourist observations, etc.

When these assumptions are tied together, the general pattern which emerges is one of dynamic interaction. Since arms control agreements, of necessity, would be implemented over a period of years, the picture is one of two or more nations gradually modifying their economies and military forces as a consequence of their confidence, or lack of confidence, in the intentions of other nations. If one could construct suitable data base representing the military, production, and manpower resources of each of the countries, it would be possible to study the kinds of decisions made by national leaders and their consequences as a function of various treaty provisions. One key provision would deal with the inspection resources permitted to each side. In addition to treaty-provided inspection, other more informal sources of intelligence do exist. The optimal use of these two kinds of intelligence resources is, of course, very critical. Thus, for example, if the treaty allowed 20 on-site inspections and 7 aerial over-flights per year, decisions regarding use of these intelligence allotments, along with an analysis of informal inputs, would have a direct bearing on the question of whether or not cheating would be detectable and possible counter-actions that might be taken. An adequate game--particularly one with a large enough data base--could be used to study the kinds of decisions made for various treaty provisions and the consequences of these decisions.

During the course of this game, experimenters are interested in two kinds of behavior: the way in which each side modifies its data base as a consequence of intelligence received, and the way in which each side expends its treaty-provided intelligence resources.

Since the construction of a data base describing in detail the military, industrial, and human resources of both the U.S.S.R. and the U.S.A. would be a very expensive undertaking, we decided to devise a more modest version consisting only of military force units.

A Blue team (U.S.A.) was organized. This team consisted of three subjects, all of whom were charged with making decisions which they felt were in the best interest of their country: a national leader, a military leader, and a civilian leader. The military leader had, in addition, the responsibility for keeping the country safe from an expected or a surprise attack; the civilian leader had the additional responsibility of spurring the economy through tax cuts or other measures all of which depended upon disarmament for funding.

In order to conserve manpower and to provide experimental control, decisions of the Red team (U.S.S.R.) were scripted to follow precisely the provisions of a disarmament proposal. That is, Red complied with all treaty provisions. Since there were no direct contacts between the two teams, the passive nature of Red and its compliance strategy were unknown to Blue. The general structure of this game is illustrated in Figure 2. Observers monitored the game from an adjacent room. Using special aids and a desk calculator, they

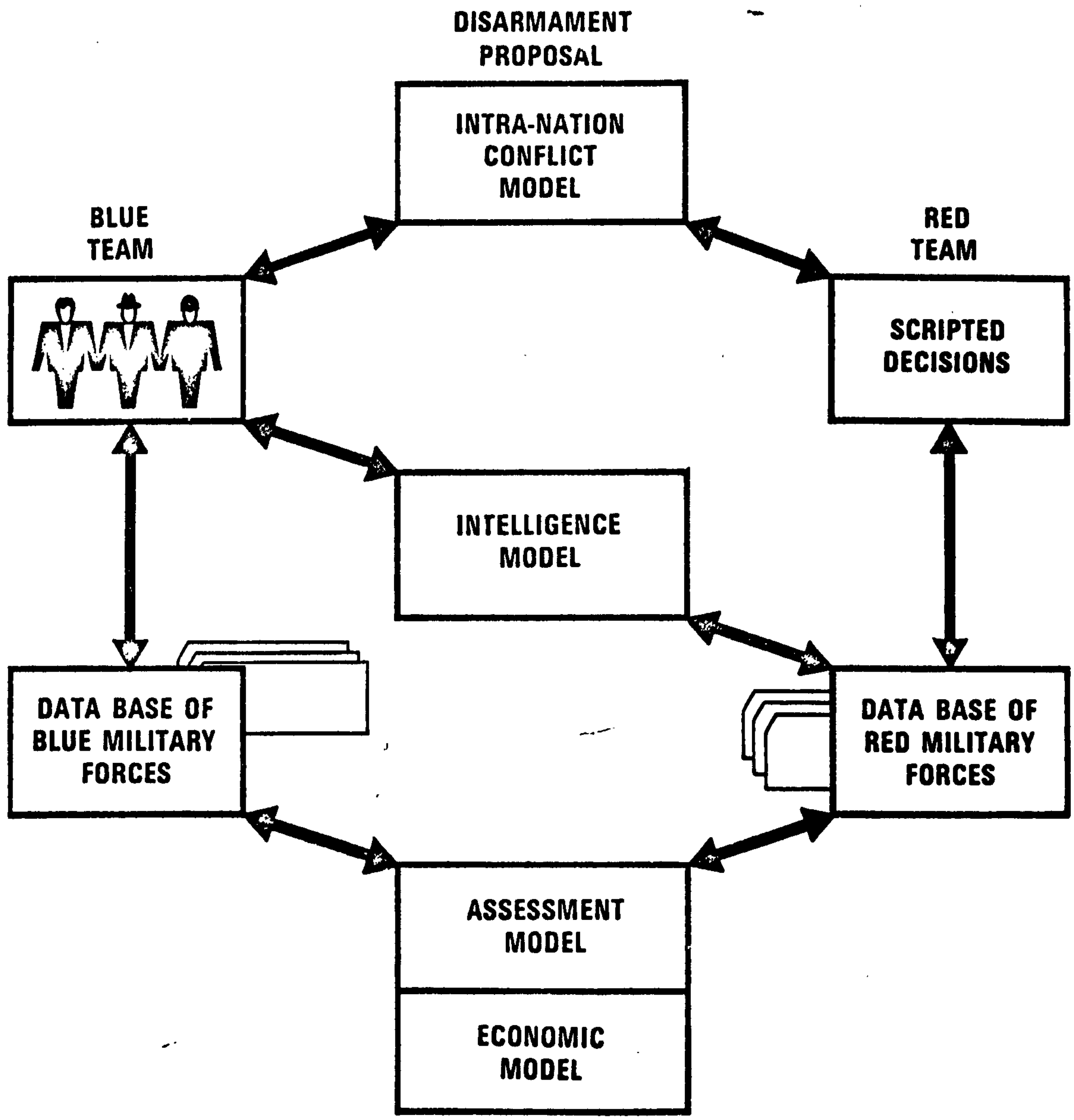

FIGURE 2. Structure of the Game

computed game results on a cycle-by-cycle basis.

The Blue team had the following basic tasks to complete during each cycle:

1. To review and evaluate the latest intelligence reports;
2. To review strategy;
3. To specify changes in the composition of their armed forces;
4. To assess the effects of changes on their strategic and limited war capabilities vis-a-vis Soviet capabilities.

The models and game aids used were designed to create a high degree of realism and to facilitate the play of the game. These included a disarmament proposal, inventories of U.S. and Soviet armed forces, an aggregated economic model, an intra-country conflict model, and inter-country conflict model, assessment models of limited war strength and strategic war strength, and finally, an intelligence model.

The disarmament proposal imposed on this game was a highly abstracted version of the proposal that President Kennedy presented to the U.N. in September 1961. This proposal called for reduction of weapons and of nuclear delivery systems through three phases.

The U.S. weapon systems inventory contained a list identified by name or number, e.g., the 101st Airborne Division. Identification numbers did not necessarily refer to their real-life counterparts. These systems were aggregated to the division, wing, squadron, or capital ship level. Acquisition and maintenance costs for each system were listed and entered into the operation of the economic model.

The Soviet weapon system inventory contained a list identified by number to the same degree of aggregation as for the U.S. weapon systems.

The economic model was aggregated, dealing with the Blue team's Gross National Product, its rate of growth, and the tax base.

The intra-country conflict model established the roles of the three players on the U.S. team as described above. In addition, a payoff pool was created to enhance both team motivation and individual competition within a team. Players could enlarge the payoff pool at a specified rate by disarming or by using money saved from disarmament to stimulate economic growth at a rate faster than expected. The national leader automatically received 1/3 of the payoff pot. The military and civilian leaders divided the remainder of the payoff pot according to how well each played his role. The more military force retained by the team, the larger the military leader's share of the pot and the less received by the civilian leader. Conversely, more and faster disarmament increased the civilian leader's payoff and reduced the reward for the military leader. Even under these mild motivational circumstances, the civilian leader seemed to regard the competing military leader as a warmonger, and the military leader seemed to see the competing civilian leader as soft on national security.

Preparing subjects to participate in the game took a half day of training. One of the concepts most difficult to communicate was the notion that this was a non-zero sum game. Possible outcomes were relative stability and moderate to high payoff to both teams; or instability, nuclear war, and virtually no payoff to either team. Description of the inter-country conflict model may clarify this point.

The outcome of the inter-country conflict model was partly determined by the actions of players and partly stochastic. Subjects had major control over the occurrence of strategic war in that the probability of strategic war increased with each game cycle that the arms level increased or remained constant. Strategic war also would occur in the event of gross differences in force structure (instabilities). This was left undefined to subjects, but was arbitrarily set at a 100% difference in force capability. The occurrence of a strategic war had the immediate effect of reducing the value of all game assets, military and economic, by 50%. (The payoff pot was also reduced by 50%). Then the remaining strategic war capabilities were subtracted from one another. This difference was added to the victor and subtracted from the loser. Thus, the cost of strategic war was inordinate, but victory, while Pyrrhic, remained more desirable than defeat. Occurrence of strategic war also ended the game after these prior calculations had been made.

The occurrence of limited war was not under the control of the team. Limited wars occurred by chance and at any time. The cost of limited war was only 10% of the military assets. Limited war also required a confrontation between the limited war strength of the two teams with the stronger team gaining and the weaker team losing the difference between their respective limited war strengths.

The assessment of strength inherent in different military postures is very difficult and remains contentious. Resolution of this issue became critical to the game because the key concepts of stability and instability necessitated an assessment procedure. Our solution (2) to this problem was to have 100 weapon system experts rank the relevant weapon systems for utility, once for strategic war and once for limited war. Scaling techniques were then used to place these weapon systems on equal interval scales, one for strategic war and one for limited war. These scales were the basis for our military assessment models (Figures 3 and 4). After taking into account differing utility for any particular weapon as the numbers of that weapon increased, summing the values of all the weapons in the inventory provided a single value for either strategic or limited war capability.

In designing this game, a number of possible independent variables were taken into consideration. These included the disarmament proposal, with attendant details such as the level of disarmament and the possible phasing of it; intelligence with its attendant details such as quantity, quality, frequency, and source; and subject instruction with regard to strategies, tactics, treaty observations, and intra-team conflicts. These variables were held at constant values throughout our several games except for the amount of intelligence available to the team.

The intelligence model provided two different kinds of intelligence, covert and treaty directed. The covert intelligence reports occurred every other cycle and contained mostly general information. Treaty directed information, the second kind of intelligence, was the only independent variable which

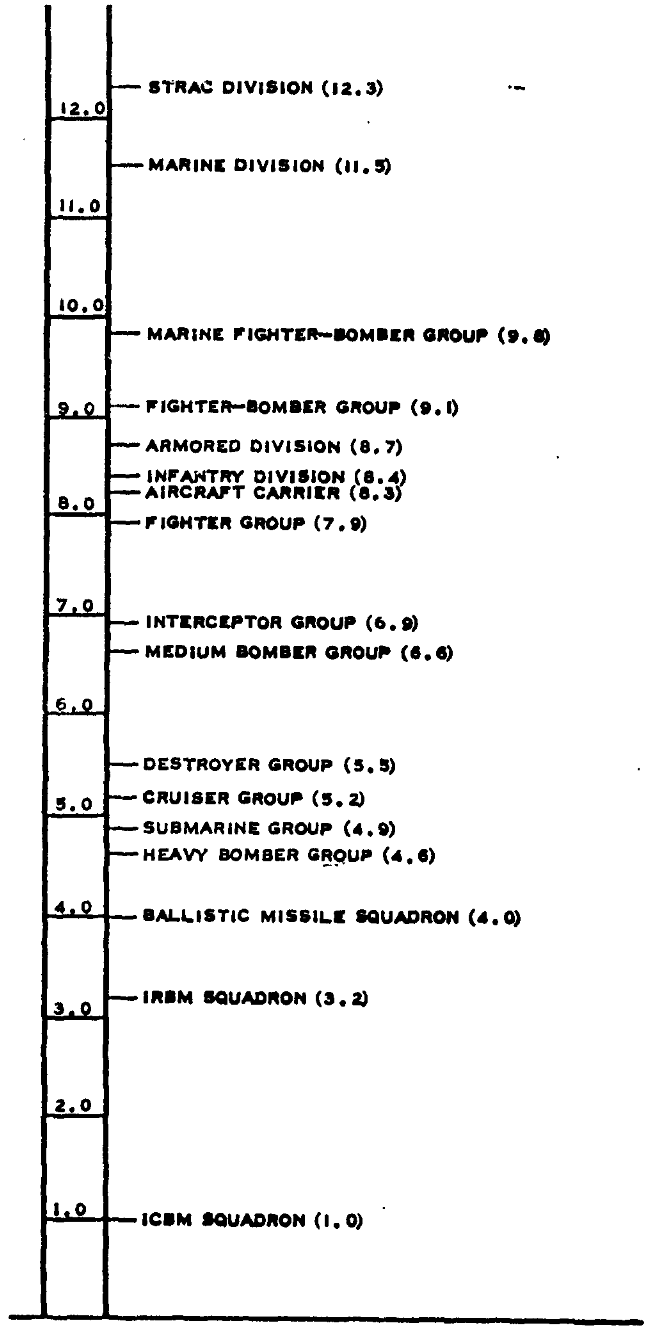

FIGURE 3. Scale Values for 17 Weapon Systems for a Limited War Ranked by 100 Experts

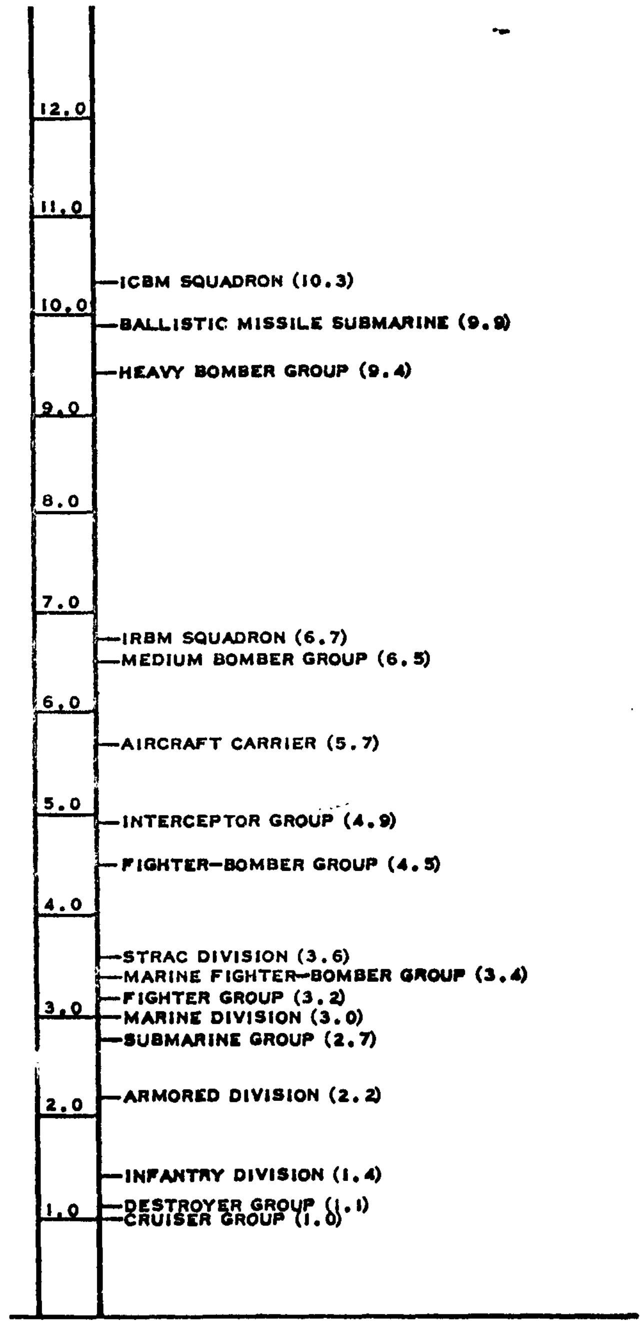

FIGURE 4. Scale Values for 17 Weapon Systems for a Strategic War Ranked by 100 Experts

was varied. One run of the game employed a high information condition which identified 10 active and 10 disarmed units from the Soviet armed forces. Another run of the game employed a low information condition which identified only five active and five inactive units from the Soviet armed forces.

The game described here represents an evolutionary rather than a revolutionary development. The present version is at least the third. Our last effort with this version was to run several feasibility exercises. Analysis of these exercises indicates that this version of the game is internally consistent, has "face" validity, at least, and is quite playable. Furthermore, it does not require extensive facilities or a computer, although use of a computer would permit faster play with more complex and adequate models.

The present version of the game was run twice, once with high information available and once with low information available, so it is apparent that the results cannot be generalized although they do provide stimulation for thought.

The top line of Figure 5 describes the effects of the disarmament of the low information U.S. team on their strategic war capability. This team, with its limited information, remained quite distrustful of the Red team, particularly of their limited war capability. They adopted a strategy of disarming very slowly overall by reducing strategic capability moderately and replacing much of the lost strategic capability with limited war capability.

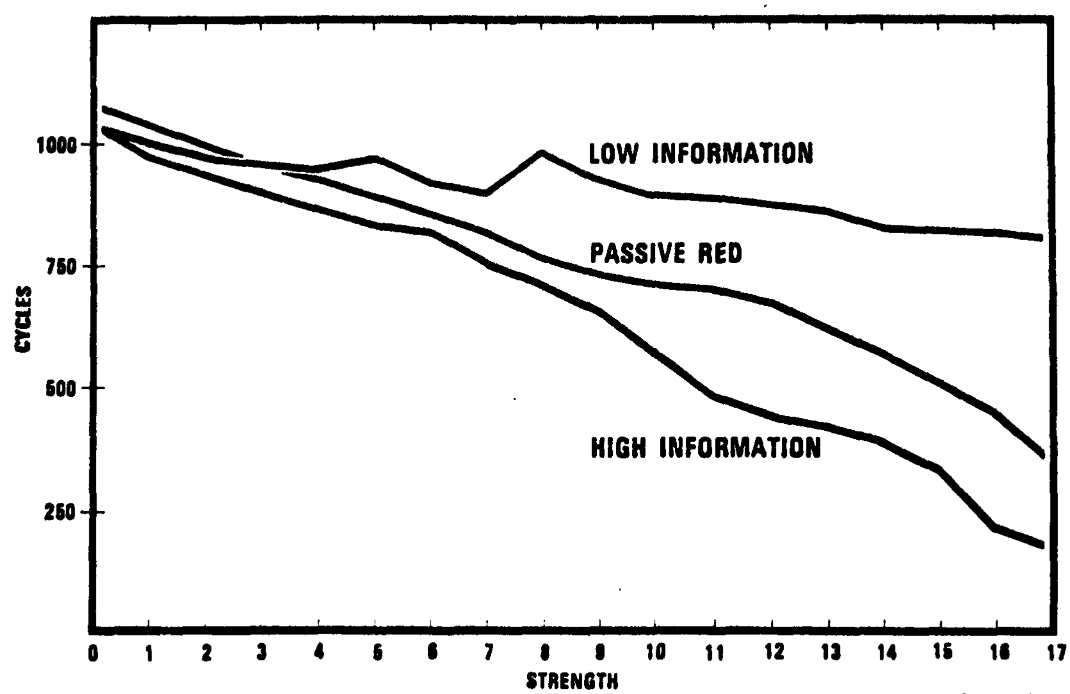

FIGURE 5. Strategic War Strength Plotted as a Function of Game Cycles for Passive Red and Active Blue under Two Information Conditions for Blue (High and Low)

The bottom line of Figure 5 describes the effects of high information on the U.S. team's disarmament of their strategic war capability. This team became convinced that the Red team was disarming honestly and in accordance with treaty specifications. The U.S. team became concerned that its good intentions might not be recognized by the Red team. To meet the objective of communicating their sincerity, this team disarmed much more rapidly than was called for by the disarmament proposal. According to the provision of the inter-nation conflict, these conditions led to strategic war.

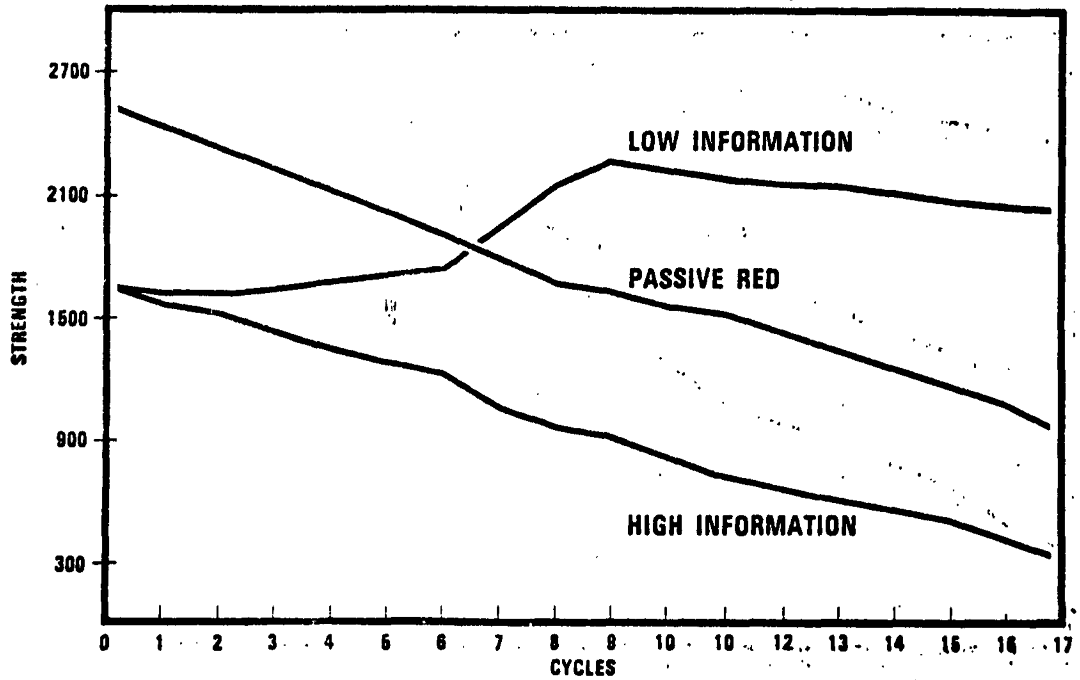

FIGURE 6. Limited War Strength Plotted as a Function of Game Cycles for Passive Red and Active Blue under Two Information Conditions for Blue (High and Low)

The top line of Figure 6 describes the effects of the low information U.S. team's disarmament on their limited war capability. Continued suspicion and intransigence on the part of this team resulted during cycle 17 in a military posture that was twice Soviet capability for strategic or limited war, and thereby was responsible, according to our inter-country conflict model, for a strategic war.

The bottom line of Figure 6 describes the effects of the high information U.S. team's disarmament on their limited war capability. By coincidence again on cycle 17, the consequences of this team's strategy resulted in the

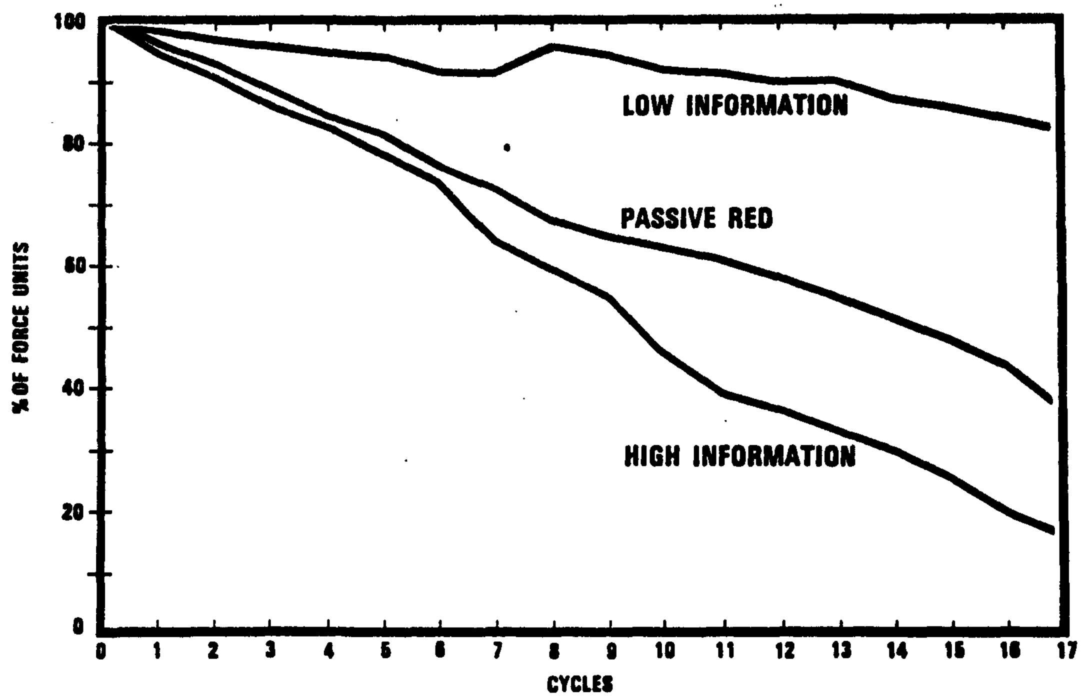

FIGURE 7. Percent of "Force Units" (Data Base Items) Left to Passive Red and Active Blue under the Two Information Conditions (High and Low)

Soviet strategic and limited war capabilities both being twice the respective U.S. capability. Again the consequences were by definition a strategic war.

These data can also be plotted in terms of percentage changes in military force units, such as divisions, squadrons of fighters, etc. The percent of disarmament that occurred for the low information U.S. team is shown on the top line (Figure 7). The bottom line describes the percent of disarmament that occurred for the high information team.

It is interesting to note that the results of this game, like those of de Rivera, point to the "inevitability" of strategic war. It is important to keep in mind, however, the fact that the game described was conducted without replications and focused on the question of feasibility--not validity. The validity problem remains unresolved for both games.

- - - - - - - - -

The relationship of all of these methods, scenarios, crisis gaming, symbolic simulation, and environmental simulation to one another has already been discussed (Figure 1). Free Kriegspiel has led to the development of scenarios and crisis games as self-contained approahes to the problem. Rigid Kriegspiel has led to environmental and symbolic simulation. When scenarios and crisis role playing are seen as aids, rather than methods, they may both be used during environmental simulation; and scenarios may be used during symbolic simulation.

IV. THE EMOTIONAL ELEMENT

The previous sections have implied that acceptability and meaningfulness depend almost exclusively upon logical and objective consideration. This is probably never the case.

An emotional element influences every judgment we make. William James (10) has used the term "live" and "dead" to distinguish between hypotheses which conform to our view of the world and those which do not.[2] We tend to dis-

---

2. James considered a hypothesis as "anything that may be proposed to our belief."

believe or reject facts, theories, methods, and hypotheses which violate our established modes of thought. In the struggle to deal with the complexities of international relations, even scientists--trained in the difficult art of objectivity--tend to simplify problems by forcing them into established and accepted patterns of thought--a tendency which Osgood (19) calls "psycho-logic." To Western man, for example, the question of the divine origin of Mohammed holds little interest; it is, in James's words, a "dead hypothesis." Similarly, for many contemporary scientists the ideas of "extra-sensory perception," "reincarnation," and even "enduring peace" are "dead hypotheses." For many of us, the applicability of a particular method, such as simulation or modeling, to the study of so complex an issue as arms control is a "dead" hypothesis.

One way of determining the degree to which a hypothesis is alive or dead in the mind of the man who entertains it may be to measure the time he is willing to devote to it. If such a time scale has any veracity at all, many of us apparently have far less confidence in the viability of our arms control suggestions and ideas than we would care to openly admit.

V. CONCLUSIONS

What conclusions can be drawn from this array of potentially useful methods for studying arms control? Are any of them really satisfactory? The answers to these questions depend almost entirely upon our ability to resolve the validity issue and this has not been done. For the purpose of prediction and evaluation, at least, no one of the techniques described is, in my opinion, adequate; not one of them can be operationally traced back to the bedrock of reality and their utility for prediction and evaluation is therefore questionable. Does this mean we should abandon all attempts to game such complex political problems? Not at all. For one thing, there is no reason to believe that the more traditional approaches to planning are necessarily better than simulation and gaming. Although simulation will probably never supplant traditional methods of individual and group planning, it can suppliment them.

The validity problem remains the most difficult one of all and because it is so crucial, we cannot ignore it much longer. As a minimum, it should be possible to test one method against another in order to determine whether they produce similar results. Furthermore, it seems reasonable to suppose that symbolic simulation could be tested against environmental simulation. It should be easier to predict the results of an environmental simulation using analytic techniques than the outcome of a similar set of events in the real world. Inputs to environmental simulation games are precisely defined by comparison with the real world. Assessment and other models exist and are readily available for symbolic simulation. Surely predicting the outcome of a laboratory run under such conditions should be easier using symbolic simulation than predicting events in the real world where inputs and models can only be surmised.

But all of this assumes a coordinated effort; it assumes that someone, somewhere has the resources and authority to study the relationship of isolated research projects to one another in order to determine their reliability, their validity, and their hierarchal relationships.

We should undertake studies of this kind not for the answers they will provide today, but because such studies may lead to better methods for solving problems tomorrow. Indeed, the great value of methodology, per se, is that it is devoid of specific contents; it is a means to a solution which--once perfected--can be used over and over again for a variety of problems.

Although our understanding of the stars began with speculation and contemplation, tools, techniques, and methods made a systematic science out of stargazing. To develop the tools and methods needed for the scientific study of arms control will be expensive; but it is difficult to point to a more deserving or crucial problem in the world today. Surely, we can find the necessary resources and ingenuity to solve it.

REFERENCES

1. Abt, C. C. Computer applications to arms control. Digest of Technical Papers, 1962 ACM National Conference, New York: Lewis Winner, 1962, 86.

2. Carpenter, P. B. & Davis, R. H. Perceived Utility of Seventeen Weapon Systems for Strategic or Conventional War. Falls Church, Va.: System Development Corporation, TM-WD-726/000/01, 1962.

3. Chapman, R. L., Kennedy, J. L. et al. The system research laboratory's air defense experiments. Management Science, 1959, 5, 250-269.

4. Davis, R. H., & Carpenter, P. B. et al. Arms Control Simulation. Santa Monica, Calif.: System Development Corporation, TM-(L)-633, 1961.

5. De Rivera, Joseph. Teaching a course in the psychology of international relations. American Psychologist, 1962, 17(10), 695-699.

6. Einstein, Albert, & Infeld, Leopold. The Evolution of Physics. New York: Simon and Schuster, 1938.

7. Goldhamer, Herbert & Speier, Hans. Some observations on political gaming. World Politics, 1959, 12(1) Reprint.

8. Guetzkow, Harold. A use of simulation in the study of inter-nation relations. Behavior Science, 1960, 4(3) Reprint.

9. Heisenberg, W. The Physical Principles of the Quantum Theory. 1930.

10. James, William. Essays in Pragmatism. In Alburey Castell (Ed.) The Hafner Library of Classics, New York: Hafner Publishing Co., 1952, No. 7.

11. Kahn, Herman. On Thermonuclear War. Princeton, N.J.: Princeton University Press, 1960.

12. Kahn, Herman. Thinking about the Unthinkable. New York: Horizon Press, 1962.

13. Kennedy, John L. The system approach: organizational development. Human Factors, 1962, 4(1), 25-52.

14. Koffka, K. Principles of Gestalt Psychology. New York: Harcourt Brace, 1935.

15. Kohler, W. Gestalt Psychology. New York: Liveright, 1929.

16. Kohler, W. Dynamics in Psychology. New York: Liveright, 1940.

17. Leibnitz, Gottfried Wilhelm von. The monadology. In Rand, Benjamin (Ed.) Modern Classical Philosophers. Boston, N.Y.: Houghton Mifflin Co., 1936, 199-214.

18. Lewin, J. Principles of Topological Psychology. (Trans. by Heider, F. & Heider, G. M.) New York: McGraw-Hill, 1936.

19. Osgood, Charles E. Reciprocal initiative. In Roosevelt, James (Ed.) The Liberal Papers, Garden City, N.Y.: Doubleday & Co., Inc., 1962, 155-228.

20. Reichenbach, Hans. The Rise of Scientific Philosophy. Berkeley & Los Angeles: University of California Press, 1951.

21. Roethlisberger, F. J. & Dickson, W. J. Management and the Worker. Cambridge, Mass.: Harvard University Press, 1939.

22. Sayre, Farrand. Map Maneuvers and Tactical Rides. The Army Service Schools, Springfield, Mass.: Springfield Printing and Binding Co., 1911.

23. Whitehead, Alfred North. Science and the Modern World. New York: The Macmillan Co., 1950.

24. Wiener, Norbert. The Human Use of Human Beings. Garden City, N.Y.: Doubleday & Co., Inc., 1954.

25. Zacharias, Ellis M. Secret Missions. New York: Paperback Library, Inc., 1961.

UNCLASSIFIED

System Development Corporation,
Santa Monica, California
ARMS CONTROL: THE SEARCH FOR AN
ACCEPTABLE RESEARCH METHODOLOGY.
Scientific rept., SP-1047, by
R. H. Davis, 17 December 1962,
39p., 6 figs., 25 references.
Unclassified report

DESCRIPTORS: Armament.

Discusses potentially useful methods of
studying arms control. States that
validity issues have not yet been resolved,
and that the techniques described are not
adequate, in the author's view, for

UNCLASSIFIED

UNCLASSIFIED

prediction and evaluation purposes.
Reports that simulation and gaming
offer excellent supplements to
traditional arms control planning
methods. Advises testing symbolic
simulation against environmental simulation
in order to examine the validity of these
methods. Suggests a coordinated effort
to study the relationship of isolated
research projects to one another to
determine their validity and their
hierarchical relationships.

UNCLASSIFIED