# CRISIS GAMES 27 YEARS LATER: PLUS C'EST DÉJÀ VU

Robert Levine, Thomas Schelling, William Jones

May 1991

P-7719

RAND

Papers are issued by RAND as a service to its professional staff. Their purpose is to facilitate the exchange of ideas among those who share the author's research interests; Papers are not reports prepared in fulfillment of RAND's contracts or grants. Views expressed in a Paper are the author's own and are not necessarily shared by RAND or its research sponsors.

RAND, 1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

**PREFACE**

The short pieces that make up this paper were written as an internal colloquy at RAND in the summer and fall of 1964. The exchange began when I bridled at what I saw as some inanities and misinterpretations in crisis games then ongoing, and resorted to the internal print process to discuss my views. (This was before the era of e-mail, when it all would have been electronic, and unlikely to be saved.) Tom Schelling and Bill Jones bridled at my bridling, and the discussion was joined.

Upon returning to RAND in recent years, I again found myself embroiled in crises over crisis gaming. I dug out the old exchange and suggested that the new gamers look at it. It proved interesting and perhaps useful to them, so this paper is intended to make it more generally available.

The five pieces are all as originally presented, except for some minor editing to remove references to RAND documentation formats long obsolete. They retain the spirit of an uninhibited exchange among people who were (and still are) friends as well as colleagues. The five are:

Robert Levine, *Crisis Games for Adults*
Thomas Schelling, *An Uninhibited Sales Pitch for Crisis Games*
William Jones, *Crisis Games for Adults and Others*
Robert Levine, *Crisis Games: A Rejoinder to Tom Schelling and to Some Extent to Bill Jones*
Robert Levine and Thomas Schelling, *Possibly the Last Word on Games*

Tom Schelling is, of course, the father of deterrence theory and many other things, including four sons from observing whom he drew many of his theories. He was then a Professor at Harvard and a RAND consultant; he is now a Professor at the University of Maryland and a

RAND consultant. Through the period, he has continued to write about gaming, his most recent piece being "The Role of War Games and Exercises," in Ashton Carter, John Steinbruner, and Charles Zraket, *Managing Nuclear Operations*[1]. In 1964, Bill Jones had recently retired as a Colonel from an Air Force career that culminated in running the main Air Force war gaming facility; he has now retired from RAND as well. I left RAND shortly after this exchange, and returned several times; I am now in my third incarnation, after doing other things in policy analysis.

Robert A. Levine
Santa Monica, California
April 1991

[1]The Brookings Institution, Washington, 1987

## CONTENTS

PREFACE ..... iii

Section

I. INTRODUCTION ..... 1

II. AN UNINHIBITED SALES PITCH FOR CRISIS GAMES..... 22

III. CRISIS GAMES FOR ADULTS AND OTHERS..... 39

IV. CRISIS GAMES: A REJOINDER TO TOM SCHELLING AND TO SOME EXTENT TO BILL JONES..... 47

V. POSSIBLY THE LAST WORD ON GAMES ..... 57

# CRISIS GAMES FOR ADULTS*

Robert A. Levine

"But listen Abner, if you ever come up with a game for two or three couples, keep us in mind."

-- conclusion of Bob Newhart's routine on Abner Doubleday's attempt to sell a new game called "baseball" to a game manufacturer.

## I. INTRODUCTION

This document is critical of gaming as a tool for the study of the management of real-world political-military crises. The criticism is not unmixed; it is suggested that some games may have some uses for limited purposes. But for two reasons it may be that even games which are most carefully designed to avoid some of the pitfalls are, on balance, questionable in their net value. The first of these reasons is well-recognized; gaming is an expensive analytical technique, particularly in terms of a very scarce resource -- the time of analytical personnel -- and carefully designed games are even more expensive than careless ones. The second reason may be less familiar, however. It is that games are seductive. Their excitement and the logical problems of structure they present seduce those who intend to use them economically into using them elaborately and frequently; their surface plausibility seduces those who enter them skeptically, "merely looking for hypotheses," into leaving them with conclusions. Thus, although the criticisms in this document are not primarily directed at games as stimulators of the imagination, games for the creation of high international theory, games as devices to acquaint

---

*Since this essay is critical of games it should be pointed out at the start that the author has personally participated in very few, and these were generally considered to be poor as games. The criticisms here are not based on this experience.

analysts with their colleagues, games as devices to acquaint analysts with themselves, games for training of analysts or decision-makers; or, for that matter narrowly military war games, implicit throughout is the caveat that such non-policy games tend to accidentally produce conclusions for political-military policy or to preempt resources which should be producing such conclusions by other methods. To the extent that they do so tend (and some examples will be given) the criticisms here apply to them as well as -- perhaps even more than -- games designed to produce such conclusions.

The arguments here, however, are directed primarily at gaming for research into political-military crisis policy. Lacking concurrent gaming of a crisis at the time the crisis is actually going on, such gaming can have two purposes. It can be used to assist in posture decisions -- the placement of forces and the writing of general doctrine for these forces -- which will constrain future action in crises; and it can be used to assist in contingency planning -- advance determination of the alternative lines of action in various crises which may be possible within a given posture. It is contended that gaming is a clumsy and dangerous tool for posture research but it may be slightly better for planning research. Many of the same drawbacks apply to both, but one of the chief criticisms of gaming is that it tends to confuse plausibility with probability. And probability is important in determining the weight of a budget-limited posture, but planning is cheap enough so that it is useful to plan for anything which might be plausible.

## II. THE DIFFICULTIES -- ERSATZ HISTORY AND ERSATZ PEOPLE

It is not a valid criticism of a mode of analysis to state that it

simplifies; analysis must separate the relevant from the irrelevant. Analysis, however, can be reasonably criticized if, after separating out the irrelevant, it brings in the artificial -- and if these artificial factors bias the results. Crisis gaming introduces two biasing artificialities: ersatz history and ersatz people.

The term "ersatz" history is a more pejorative but more descriptive phrase than the common term "scenario." Scenarios used without games are subject to many of the comments listed here, but writing scenarios is both cheaper than gaming and less likely to cause hypotheses to be mistaken for conclusions. Although crisis games are not necessary for scenarios, however, political-military scenarios are necessary for crisis games. The scenario provides the background which explains why a crisis exists in 196x, and it describes the objective military and other conditions which constrain the actions which can be taken in the crisis.

And the trouble with crisis-game scenarios is that because they are used to establish plausibility of the crisis when plausibility really wasn't in question, they tend to substitute an implicit estimate of probability for such plausibility. There are very few crises which cannot be made to appear plausible -- possible -- with just a little reflection. Tell me that the Russians took over a section of Northern Finland six months ago without NATO being particularly disturbed, and that they are going to use this as a springboard for an invasion of the North Cap of Norway, and I can believe that it could happen. Tell me that in a war between Greece and Turkey the Russians suddenly dropped their fear of escalation and intervened militarily on one side, and I can believe that it could happen. Further, given that it could happen, this in itself provides a rationale for investigating, by gaming or other techniques,

what kind of forces we would need to oppose the Soviet action -- but if gaming is used, it is military war gaming, not political-military crisis gaming.

Plausibility of almost any crisis, then, is not in question. But so many crises can be made to appear plausible that plausibility furnishes no guide to how a political-military posture should be weighted to meet various plausible crises or classes of crises. The key to posture decisions, given limited budgetary and other resources, is not plausibility but probability -- relative probability in a crude statistical sense. An over-all posture should have something to do in case of any plausible crisis, but the bulk of it must be weighted against probable crises; if none seems more probable than any other, then this in itself implies some sort of equal weighting.

But ersatz history, by highlighting certain plausible crises makes them appear probable relative to the ones which, randomly, were not chosen, and ersatz history provides no way of objectively evaluating the relative probabilities of the different crisis demands on a single force posture. To take one example, Harvey Averch and Marvin Lavin, in their RM-4202-PR, Simulation of Decisionmaking in Crisis: Three Manual Gaming Experiments, report that "Our games consistently suggested that being able to expand local military capabilities rapidly is extremely important to political decision makers. The play revealed many specific requirements for sudden orders-of-magnitude expansion of logistical flow to a crisis theater, for quick redeployment of European and ZI-based ground and air units and their command and control systems, for effective ad hoc arrangements providing air defense and reconnaissance, and so on."*

* RAND Research Memorandum, August 1964, pp. 30-31.

But the trouble is that the existence of possible demands for force redeployment to theaters where forces are meager should be self-evident without gaming; at least it should be obvious that scenarios starting with trouble in Northern Norway or the Greek-Bulgarian border are rather likely to lead to a call for such redeployment. What games cannot do, however, is establish the weight of such demands for scarce resources as compared to the weight of demands, say, for forces on NATO's Central Front. And by making the situations played appear concrete in the minds of the players and the analysts, the games provide the chosen crises with a spurious subjective probability which can easily bias posture decisions made under budgetary limitations.

The problem is similar to the one highighted by Roberta Wohlstetter for Pearl Harbor. For twenty years the debate raged about why we hadn't read the signals of the attack on Pearl Harbor. Her book* pointed out that we had read the signals, but we couldn't distinguish them from the noise; we could not tell which of the many possible Japanese actions were more likely than others. We were prepared for none of them, but if we had held a crisis game, we might have been well prepared -- for a surprise attack on Manila on December 21, for a seaborne landing above Singapore in November, for sabotage on the Panama Canal in January, perhaps even for an air attack on Hawaii on December 7.

Two contradictory "real world" examples are sometimes used to argue for crisis gaming. One is that before Pearl Harbor, General Matthew Ridgeway was severely criticized for announcing a scenario for a war game as fantastic as the United States Pacific Fleet having been neutralized at the very outset of a war. The other is that nobody predicted the Cuban missile crisis before it happened. The implication drawn from

---

* Pearl Harbor: Warning and Decision, Stanford, Stanford University Press, 1962

the first is that therefore we should take games (and scenarios) very seriously; the implication drawn from the second that we should have lots of games so as to turn up randomly the improbable cases like Cuba. But the true implications, if any, might be almost the opposite. The fact that Ridgeway was not taken seriously seems to show that more is needed than the plausibility of a scenario to show that a certain sort of event should be prepared for; one needs an analysis showing not only that these things are possible, but also that the enemy's incentive to do them may even make them probable.

And, so far as the Cuban missiles are concerned, even hindsight does not show that we should have expected the crisis. Analysis a year and a half after the event is much more along the lines of, "For what possible reasons did they do that stupid thing?" rather than "Of course. We should have expected it all along." A game in this case would have shown its improbability; a game was not needed to show its possibility. Just as Ridgeway did not separate the Pearl Harbor signals from the noise, a Cuba game would not have separated the possible missile crisis from the noise of the near-infinity of plausible crises.

Indeed, the one recent exception to the rule of crisis games not being useful for policy planning tends to establish the rule. This is the 1961 (post-wall) gaming of the Berlin crisis. These exercises, reported to be highly useful, had a characteristic necessarily different from practically all other games; the scenario was real history, not ersatz. Unlike the Cuban missile crisis, the 1961 Berlin crisis had so much history from 1958 on that every historical, geographical, and military factor had been studied; unlike Cuba, the 1961 Berlin crisis moved slowly enough that there was time to game. What had happened is that out

of the infinite richness of possible history, real history had, by chance or otherwise, narrowed the possibilities down to what had happened. Some additional ersatz history was of course necessary to carry the game forward from the date of game commencement, but the branch points here were many fewer than they would have been, say, for a 1961 Berlin crisis gamed in 1958. I do not know whether anyone predicted the wall before it was built; certainly nobody did publicly, and I rather doubt that it was predicted in any classified study. For similar reasons it seems unlikely that a 1967 Berlin exercise carried out in 1964 would have even the same order of magnitude of utility as the 1961 Berlin crises studied in 1961. In fact, when such a 1967 Berlin game was recently put on, perhaps the most interesting phenomenon was that although the game directors tried to specify ersatz history and environment that distinguished 1967 from 1961 (e.g., assumed different strategic balance), the game participants intuitively went back to real history, ignored the differences, and played it under about the 1961 conditions.

The ersatz people problem (which is peculiar to games, rather than scenarios) is somewhat different. As representations of decision-makers, the players of a game may not be consistently biased in given directions, but it is difficult to say what they represent or what differences with the real world should be allowed for. The difficulties are threefold: game-players never seem quite sure whether they are supposed to be positive or normative representations of decision-makers; game-players generally have both an imperfect knowledge of the decision-makers they are representing, and even more important, a very imperfect intuition of the way in which the decision-makers feel the various pressures on them; and game-players do not have the time to think systematically and

objectively of the ways they differ from real decision-makers, but rather run all the factors through an *ad hoc* and poorly programmed human computer.

The first difficulty, the confusion between positive and normative representation, is probably the most easily solvable, at least in principle. Without initial agreement on the point, some players are going to act as they feel decision-makers would, some as they feel decision-makers should, and most players are going to use some self-generated and unspecified combination of the two. Under this condition, it becomes rather difficult to tell what combination is illustrated by the game results. If all the players were to act as they think the real decision-makers *would*, then what would be shown (imperfectly) would be how the initial posture specified by the scenario constrains crisis decisions. This posture then might be compared to others. If all the players were to act as they think the decision-makers *should*, then a comparison might be made among the results of this "prescribed" set of actions and other sets, prescribed or predicted. If some players act one way, some the other, and some in between, however, it becomes difficult to see what might be shown -- even imperfectly.

Possible positive-normative inconsistencies might be taken care of by careful instruction at the beginning of a game, although since the various players' perceptions of both the "shoulds" and the "woulds" are likely to vary, this may not be very simple. But once consistency here has been obtained, the next problem, that of accurate representation of the decision-makers, arises. In a sense, the problem is simpler if it has been decided that the players should act consistently normatively, since then there is no call for accurate representation. The players would then act

as the decision-makers should; ignoring real-world constraints would make the results recognizably beyond any possible realm of reality; and the whole thing would be discarded.

For this reason, the more frequent decision is for the players to try to act as they think the decision-makers would. And ordinarily, game-players are just not very good at it. The problem is easily recognized (and is usually recognized) for the Red teams. That we simply know very little beyond generalities about what motivates the Communists is ordinarily admitted; simulated actions may resemble those which would be taken by real Russians, or they may not -- no one knows for sure. This can be a killing fault for one sort of policy game -- that which attempts to simulate international organizations, negotiations, and the like, and come up with generalized conclusions applicable to arms control, the UN, etc. Such games provide conclusions applicable to groups of college students, but seldom much else. For crisis studies, however, the difficulty of simulating Russians is not so serious; the subject under study is Blue policy, not Red, and so long as the Blue team is tested against a variety of Red policies, realistic simulation of the Reds is relatively unimportant. Indeed, it might be asked, why a Red team at all; why not just play Blue against a malevolent Control?

But in some crisis games, it appears that the players know as little about United States' decision-makers as Russian. And although this is completely not true for the more sophisticated games and players, it does remain the case that there is obviously some variation between real decision-makers and their play counterparts, and nobody knows how much or in what direction. To cite one example, we know that a United States President is going to be extremely reluctant to initiate the use of nuclear

weapons; we also can feel reasonably sure that there exist possible circumstances in which he would initiate their use. The extreme cases of likely use (massive Soviet invasion of Europe) or likely non-use (defoliation of South Vietnam) are easy to figure out, but at which set of circumstances does the boundary between use and non-use lie? The present state of the art of determining this boundary can be characterized by pointing out that individuals who have looked rather hard at the subject cannot agree either on where it is or whether it can be changed by trying to erode the importance the "firebreak" seemingly has in the minds of American decision-makers. Perhaps the art can be advanced by detailed study -- by psychological study of Presidents and political study of the institution of the Presidency, as well as military study of the effects of use or non-use of nuclear weapons. Perhaps factors in the Presidential decision can be discovered which are now unknown to the President, since it seems unlikely that he can finally know until faced with the decision. But for now, these things are unknown. In an objective and careful study of crises they can be allowed for, or explicit assumptions can be made. But in a crisis game, one player with his own subjective imperfect ideas about the Presidency (and with the degree of imperfection unknown) makes a decision, and all results proceed thence.

The fact is that the President and the Presidency are both very complex, and a game simplifies them by snap decision. If this introduces a known bias, it may be one in favor of the obvious as compared to the subtle, and if the obvious is what we are after, there are cheaper ways to go about it. But a game tries to solve the problem of the real subtleties that govern the world (particularly, perhaps in crises) by using real people as play-acters -- ersatz decision-makers -- and hoping that

the real subtleties of the real people will somehow match the subtleties of the decision-makers. There is no reason to believe that they will. The kind of man who is elected to high office in the United States is quite different in his makeup from the kind of person who becomes a high-grade game-playing systems analyst (whether the analyst is in uniform or mufti).

It may be that the elective office-holder can be studied and understood, if not by the systems analyst, at least by the political scientist, but knowing him and being him are two different things. Many economists (including myself) long ago observed that an economist -- even a good economist -- is not a businessman. Similarly, a political scientist is not a politician, nor for that matter is a Sovietologist a Russian. Indeed, the major advantage that narrow war games have over crisis games is that at least the active or retired military officers who play war games are faithful representations of themselves. Similarly, it may be that in a crisis game played by government officials at a very high level, these officials may be faithful representations of themselves, although even here some doubt can be introduced because the pressures on the officials during a game can only be ersatz. In any case, it is difficult to see how the key element in U.S. decision-making in any important crisis -- the President -- can be reproduced. And in no case does it seem likely that analysts and researchers can mystically reproduce in their own persons the decision-making officials.*

---

*One approach which has been tried is to use as game-players, rather than those who may understand what motivates decision-makers, those who may, in their personalities, reproduce decision-makers. Harold Guetzkow, a seminal crisis gamer, is quoted (by Arthur Herzog, in his forthcoming A Search for Peace) as saying: "We want to get the right

A major advantage the analytical essay has over the game (and, for that matter, the computer representation has over the human one) is that assumptions must be made explicit. Ersatz people with unknown biases can produce only questionable results.

## III. GAMES AS GENERATORS OF HYPOTHESES

Many of the above contentions are readily admitted by those who advocate the use of games as a crisis research device. Their claim is that, although biases such as those discussed invalidate games as generators of research conclusions, they nonetheless are valuable as stimulators of hypotheses which can then be investigated more rigorously. But the questions to be asked are, in the light of the costs of gaming, what the quality is of the hypotheses, and whether, in fact, the results of games are treated as only hypothetical.

The hypotheses generated by games can be put into two categories: those having to do primarily with the actions, feelings, or beliefs of a single side; and those having to do more directly with the interactions of the two sides. Looking first at the reactions of a single side, the only relevant reactions here are those of Blue, since Red in these games is used mainly as a foil. And it is doubtful here that a game can produce meaningful Blue reactions and beliefs other than those which were

---

personalities not only because we want to be historically accurate but also so we can deliberately put in the wrong personality to see what would happen then. We are painstaking. All the details of our simulation are accurate, and we tested 800 people in one game to get the right 24." Guetzkow works primarily with high school and college students, and it is fairly easy to see from the outside that much more than personality-matching is needed to get accurate simulation from them. Yet there is little reason to assume more accuracy from simulation by analysts who do understand much of the problem -- but still from the outside.

input by one or another of the Blue players initially. The Blue players' feelings, for example, about what United States objectives in Europe are, how important they are, where we can afford to compromise or allow the enemy to save face and where not, etc., can be reported as game outputs, but they seldom differ from the beliefs the players brought to the game. And to the extent they do differ for an individual player, they seldom present anything novel which might be called a contribution of the game to the universe of analytical knowledge. Rather, the game may present an arena for the holder of one existing belief on U.S. objectives to convince the holder of another. The result may thus be a new distribution of beliefs among people, but hardly a different set of beliefs for the group as a whole. At least the burden of evidence must be on the advocates of games, because it does seem doubtful, on the face of it, that a game could create hypotheses about Blue attitudes which did not exist before.

And, looking at the evidence, take for example, the report of Averch and Lavin on the "attitudes" of their games. They head it by saying that:

> What was most important to BLUE in all three games was (1) providing public and private evidence of the American evaluation of its own power position and its willingness to maintain its general rights and obligations, and (2) maintaining the status quo in Europe or returning to the status quo ante.*

But that the major objective of Blue in Europe is to maintain its own rights and the status quo, while it may not be self-evident, is certainly not a novel idea; it has been the basis of NATO policy at least

---

*Op. cit., p. 15. Averch and Lavin are cited frequently here not because they provide the worst example of a discussion of crisis gaming, but because they provide one of the best, in terms both of depth of analysis and modesty of claims for results. Also, their work is handy.

since the 1949 Berlin Blockade. It would be interesting if, somehow, the Blue teams had decided, because of something which happened during the game, that their real objective was the liberation of the satellites. But the improbability of such a decision consequent upon game events makes the point that attitudinal inputs equal attitudinal outputs.

Of course, this is not completely fair to gamers. The hypotheses that are supposed to be stimulated are those having directly to do with bilateral interactions rather than unilateral attitudes. But here too the burden of proof of such stimulation must be on the gamers. Since it is agreed that games cannot test hypotheses,* the drawing from a game of a statement about policy already in common currency elsewhere can have little value; a useful game-stimulated hypothesis must be a novel hypothesis. And again, although it is impossible to list all hypotheses stemming from all crisis games and demonstrate that none of these are novel, an example of the best kinds of hypotheses stemming from recent games is suggestive.

Escalation is par excellence, an interaction phenomenon of the kind for which games might be useful. In his "Escalation in the Game World and the Real World," Seyom Brown discusses the results of the Averch-Lavin games as they affect escalation.** He discusses them in the most sophisticated way, arguing from the premise, as shown in his title, that it must be the variations of the games from reality which

---

*At least it is agreed by responsible gamers. E.g., "... we believe crisis games not only have poor predictive qualities, but also are doubtful sources, of themselves, of conclusions on crisis phenomena." (Averch and Lavin.) Those who, in spite of the ersatz history and ersatz people problems listed above, as well as other problems mentioned by Averch and Lavin, would deny the validity of their statement are, frankly, not worth discussing.

** Internal Working Paper, June 1964.

are instructive. His substantive conclusion is that the critical variables in escalation are: the value of the objective fought for; the local balance of force; the over-all military balance; and the risk-taking propensities of top decision-makers. He then goes on to point out that the relative weights assigned to these determines in an important way resource allocations to various military systems and strategies. Now, these are valid and perceptive statements, but it does Seyom no credit to say that he deduced them from the subject games. They are just not very novel; they come through pretty clearly, for example, in much of Seyom's work done before the games, and they have been analyzed in considerable detail by Morton Halperin in his Limited War in the Nuclear Age. Yet, to repeat, these are the good hypotheses "generated" by good games.*

Even if it is admitted that games may not after all be productive of new hypotheses, however, it may nonetheless be contended that they have value for other purposes -- for the spreading of certain old ideas into quarters where they had not previously been well-received, for example. But the seductiveness of gaming is such that it is all too easy to turn hypotheses into conclusions and small gross value into negative net value. In one recent discussion of gaming, the usual disclaimer of

---

*Seyom's methodological conclusion is somewhat more curious:

"The game 'unrealism' should therefore be looked at as a very important device for examining the consequences of being wrong in our pet premises about the real world, whether they be on Soviet behavior or the nature of escalation generally. It raises the following pertinent question:

"Suppose in the next few years the Soviet learning curve rises very rapidly so that they begin to behave as 'intelligently' as our RAND gamsters. If this is coupled with actualization of current projections of the strategic balance, what then?" (pp. 5-6.)

This might be paraphrased, "We should game because the real world might grow to look like our ersatz construct." This strikes me as being a rather long shot.

See Morton Halperin, Limited War in the Nuclear Age, New York, Wiley, 1963.

conclusions was presented at the beginning; at the end, the language was "our games confirmed that ..." That this is not the language of hypotheses was pointed out, and the language was changed, but it remains true that the gamers, responsible men one and all, were caught up in the spirit of the thing and needed to be reminded that they were merely hypothesizing.

Similarly, but more dramatic, consider the example of one of the leading workers in the field of theoretical games, Harold Guetzkow. In 1959, he wrote:

> Our simulation is an operating representation in reduced and/or simplified form of relations among social units by means of symbolic and/or replicate component parts ... The concreteness to be embodied in this description of the inter-nation simulation must not be taken seriously. As the work proceeds, revisions necessarily will make the representation more adequate ... It is believed that inter-nation simulation will be of heuristic value in clarifying our theories of international relations.*

But compare to Guetzkow in 1963, when he evidently assumed that his representation was adequate to aid directly in policy design; he told an interviewer:

> I call ourselves the idea-builders. The trouble with working in international relations is that you're handicapped because you can't experiment with history, so we have tried to create it, using a combination of people and computers ... We are painstaking. All the details of our simulation are accurate ... We would like to be able to predict the outcome of future events, and we've done some studies along these lines. One was on the proliferation of nuclear weapons, the Nth century problem, and our game showed that the world was different after the spread of nuclear arms ... The outcome was very like what is happening today

---

*Harold Guetzkow, "A Use of Simulation in the Study of Inter-Nation Relations," in Behavioral Science, July 1959, pp. 184-188. Emphasis in the original.

> with France -- with alliances gone to pot ... One third of our worlds had thermonuclear war.*

It is exceedingly difficult for someone with an interest in policy to remain on a high theoretical plane. Indeed, although most of the criticisms of this paper do not apply directly to games for the creation of international relations theory, such as Guetzkow's, games for individual or group stimulation or mental discipline, or games for training analysts or decision-makers, even such games can be way-steps on the paths of temptation, as witness Guetzkow. If a game result appears in language which can be read as a confirmed policy conclusion, there is a substantial chance that it will be so read.

## IV. WHAT THEN?

It has been asserted here that ersatz history and ersatz people introduce biases into games which invalidate them as research tools for the study of crises, that the hypotheses produced by such games are seldom novel, and that these hypotheses are all too easily confused with conclusions. Destructive criticism has its place, and it is not necessarily incumbent upon the critic to suggest alternatives, but particularly since the problem of to game or not to game is in part one of allocation of research resources, "What then?" becomes a reasonable question to ask. What is offered here in answer are some ideas which have not been fully developed, but are nonetheless (hopefully) suggestive.

To begin with, it may be useful to separate the long-run and the short-run study of crises. The long-run study, as conceived here, may

---

*Quoted in Arthur Herzog, A Search for Peace, New York, Harper, forthcoming, pp. 270-271 of the draft.

be quite long-run, and thus not too helpful for the next few crises. For the long-run study, the implication of the primary criticism of gaming presented here is that the method should avoid using ersatz history and people. And the obvious way to do this is to study real history and people. That "history repeats itself" is a phrase which can be understood or misunderstood on several levels. The ordinary understanding is that history repeats itself in big lumps -- that looking at a whole historical situation, a crisis in this case, we can expect a rather close repetition sometime in the future. An example is the translation of the early twentieth century arms race which brought about the irreversible mobilizations of August 1914 into the current "arms race" and the possibility of irreversible escalation. The belief in this kind of big-lump repetition is buttressed by examples where some sort of repetition can be shown. Warner Schilling, for example, discusses the 1919 attempts to negotiate the abolition of submarines or at least the exemption of merchant ships as targets, and compares the lack of success to the current lack of success in attempts to negotiate abolition of nuclear weapons or at least exemption of cities as targets.*

The trouble is, however, like the trouble with the use of games to establish plausibility and probability -- coincidences (big-lump repetitions) sometimes do occur, sometimes they do not, and nobody can tell when they will or will not. As a result, history repeating itself in the big-lump sense can be interesting in the way that Bill Stern's stories of sports coincidences ("And those two little boys later grew

---

*Warner R. Schilling, "Weapons, Doctrine, and Arms Control: A Case from the Good Old Days," Journal of Conflict Resolution, Vol. VII, No. 3, pp. 192-214.

up to be Babe Ruth and Lou Gehrig") were interesting, but it is not very useful for research.

In another sense, however, perhaps history does repeat itself. If it is broken down into fine enough parts, perhaps there is historical repetition. The Cuban missile crisis and the Berlin wall crisis will never repeat themselves exactly, but some of the small parts may. Some examples might be the American President's desire to at least start the confrontation at as low a level of threat and violence as possible in both cases, and the strength of Khrushchev's desire to maintain viable Communist hegemony in Cuba and Berlin at high costs. These examples may still be in too big lumps, and the offhand interpretation suggested here may be completely incorrect. If so, they illustrate the point that breaking history down into fine repetitive lumps is not to be done casually, substituting a poorly programmed ad hoc human computer called an analyst for a poorly programmed ad hoc human computer called a game player. Rather, it must be done with full and detailed research by someone who is expert enough to know the sort of thing he is looking for.

A possible addition to such a micro-analysis of real history as a substitute for ersatz history would be a micro-analysis of relevant people as a substitute for erstaz people. Obviously the President of the United States is going to be a key element in any crisis involving the United States; equally obviously some of the Presidential reactions are going to be peculiar to the individual incumbent and others are going to be peculiar to the office. In such conditions, statements like, "What m a President in time of crisis is the effect on his personal prestig giving up on previous commitments," are not very helpful. But per a careful depth interview designed by, at a minimum, a good poli

scientist and a good psychologist, could elicit both consistencies having to do with the office and differences having to do with the individual. Interviewing an incumbent President in such a way might be too sensitive to be accomplished, but under the heading of classified and personally confidential national security research might not Messrs. Truman and/or Eisenhower be willing? And in addition, there is a small class of individuals in the United States who have undoubtedly thought in the most serious terms of what they would do in various situations if they were President, a class including such people as Stevenson, Nixon, Rockefeller, Humphrey, et al. Might they be willing to discuss such things? If, as suggested above, the mind of the Presidential politician is at the same time they key element in crisis management and the least-understood element by crisis analysts, some such technique might be useful.

What is tentatively suggested for long-run crisis research, then, is the breaking down of history, personality, and political institutions into repetitive elements. Once this analysis is done, it might then be possible to synthesize -- to put back together new combinations of these elements (together with similar Soviet inputs) and find out what makes a crisis move, and what are key elements which might be shifted through pre-crisis posture, policy, and planning. The suggestion of this sort of process should not be mistaken, however, for an avowal that history is determinate if it is broken down far enough. Maybe in a philosophical sense it is, but the process here would, at best, lead to a distribution of possible outcomes, which could then be studied. If the elements are fine enough and numerous enough, it might even be conceivable that some sort of computing machine procedure would be devised for

recombining. If probability estimates of the elements can be made, then perhaps probability weights can be computed for the output alternatives, thus getting around the plausibility-probability difficulties discussed above.

The result, then, might be using the human mind for that at which it seems best -- observing evidence, analyzing it, and looking for consistencies; and using the machine for digesting and recombining large quantities of data. The result might be synthetic history, but it seems far prefereable to what has been termed here ersatz history, which involves using the human mind carelessly rather than carefully, even in jobs for which the mind is best, and also using it as an ad hoc computer in jobs for which a real computer might be better.

The sort of procedure outlined above is obviously far too long and complicated to help much in the immediate study of on-going or soon-coming crises. For these, simple ad hoc tools may be all that are available. There is no reason why gaming should not be among these tools. But perhaps if used, it should be used under the absolute caveat:

Game if you will. But in presenting policy results, don't tell anyone that you gamed. Present it in essay, model, or other analytical form without mentioning the game. If it is convincing in this form, then the game has been as good an instrument as any. If it is necessary to fall back on game "evidence," however, then the whole process is of very doubtful validity.

## AN UNINHIBITED SALES PITCH FOR CRISIS GAMES

Thomas C. Schelling

I almost always agree with Bob Levine and write him a short note to say so. This time I disagree so completely that I shall exploit my RAND-consultant status and babble at length about Bob's recent internal working paper "Crisis Games for Adults," and ask that mine be be given the same circulation as his. The trouble with Bob is well described on pages 3 and 4 of his own paper: Bob has worked up a few scenarios to establsh the plausibility of some deplorable phenomena and then goes on "to substitute an implicit estimate of probability for such plausibility." All the bad things he described could happen, they sometimes do. My experience is that they usually do not.

Actually I would settle for the final underlined paragraph of Bob's paper, at the bottom of page 21. Let me repeat it, without his underlines: "Game if you will. But in presenting policy results, don't tell anyone that you gamed. Present it in essay, model, or other analytical form without mentioning the game. If it is convincing in this form, then the game has been as good an instrument as any. If it is necessary to fall back on game 'evidence,' however, then the whole process is of very doubtful validity."

As a first approximation, that is a terribly important and valid comment about the results of games. (I have been saying it too, for a good many years.) I would modify it in a few ways. First, many of the results are not "policy results." Second, much of the discussion takes place among the participants themselves, to whom the game provides a wealth of illustrative material that they have in common and that aids quick communication among them. Third, there are a few kinds of hypotheses that may become "established" on the evidence provided by games, and many more for which a single valid illustration is an adequate rebuttal; a "wrong" principle is often one that is widely held in absolute form.

Since Bob is raising practical arguments, not just theoretical questions, let me begin with a mundane rebuttal. Bob refers in

several places, including his very first paragraph, to the high cost of these games "particularly in terms of a very scarce resource -- the time of analytical personnel." I have now participated in something like 9 crisis games, and have been a bystander or consultant in a few others; confining myself just to the games I was associated with, which have probably involved about 250 valuable people, I assert that the cost of these games is invaribly negative. (I am speaking now only of the games that I have witnessed.) The participants have been distributed along a value scale that has people like Bob Levine and me at one end and cabinet officers at the other; the time involved has been on the order of three full days for most participants, with the exception of very high-ranking people who may have devoted time equivalent to a full day. I will let Bob specify any cost function he wishes; and I will make the assertion not only for the aggregate of games but give high odds on it for any single game. A simple test, but not the only one I have in mind, is that any participant asked the day after the game, a month after the game, or two years after the game, whether it was worth all the time and trouble will say, with something like 90 per cent probability, that it was unquestionably worth his while.

In fact, leaving aside the <u>important</u> results of the game, the ones for which the game is organized, nearly any such game will yield by-products worth the cost in people's time. Let me mention three kinds.

First, the games are intensely stimulating; people are very active; ideas and conjectures get tossed around and analysed by a highly motivated group of people; a great deal of expertise is collected in a single room, expertise that is not often collected together; and people discover facts, ideas, possibilities, capabilities, and arguments that do not in any way depend on the game but nevertheless emerge in it. Somebody discovers that in some country the supply of jet fuel is ten times greater than he realized because the cooking fuel is kerosene and domestic storage supplies are large if security measures are taken to preserve them. Again, somebody discovers that an action around the world timed to coincide with a

high-level television appearance is constrained by the the fact that people turn off their sets by midnight and do not turn them on much before breakfast. One discovers that certain precautions could be taken without much notice early in a crisis if they had already been established as routine, but if precedents had not been deliberately created they constitute "unprecedented" acts and excite attention. Or somebody discovers acute jurisdictional problems among the three sectors of West Berlin that are surprising precisely because they would never have occurred to him and he would never have thought to ask. Or somebody discovers important things about the way SAC ground alert is organized that utterly contradict what he had taken for granted all along and put, in the short run, certain "obvious" responses completely beyond reach. Or one discovers that there are sensitive diplomatic problems in putting American overseas forces on alert without consulting the host government and equally sensitive diplomatic problems in consulting the host government. One discovers that certain operational plans have embarrassing byproducts that could not be known to the people who drew up the plans, and the people to whom they could be embarrassing had no part in drawing up the plans. There is absolutely no reason why these little "lessons" depend on a game; as a matter of fact, they do not. I am simply asserting that the game, as a social and intellectual occasion, tends to be highly productive of little things of this sort, "little" by comparison with grand policy results but not so little in relation to a three days' investment.

Second, people probably learn more about the geography, the distribution of population, the telephone system, the recent history, the political personalities, the diplomatic entanglements, the weather, the street layout, the armed forces, the political and ethnic groups, and all the other "tourist" information about a country, by going through a game of this sort than by any cram course that could be devised for a comparable two or three day period. This is a hell of a way to advertise a game; but it is an important by-product. If somebody were going to be responsible for some operations in the Pacific Islands, or were going to be Deputy Chief of Mission in

Finland, or going to run an AID program in Cyprus, just putting him into a game for three days focused on the area he is going to would teach him more than he could get by any kind of briefings, lectures, reading program, or other program of self-improvement.

As a third kind of by-product, let me just advert to friends who have spent a year at a war college, considered it worthwhile, and felt the greatest benefit was that they became intimately or at least casually acquainted with a number of people that they might have occasion to work with or rely on in the future. These games can be an intense experience in joint problem solving; one might suppose that they would create lasting antagonisms, the contempt that goes with familiarity, or an exacerbation of existing differences. Experience strongly suggests the contrary. I hate to brag about the games on grounds that they create good fellowship, but they do provide (in only 3 days or so) an intense common experience in circumstances where a bit of humor is allowable and mistakes do not lead to real disaster, and typically lead to remarkably good relations afterwards. I used to be worried about putting military and civilian people together in a game environment, where each would confirm the other's worst expectations. After a few games I quit worrying about it.

All the above is preliminary, and just on the question of how high the costs are. As I said, I am sure that the costs are below zero (smaller than the by-products) and need not concern us. The question is whether the major results are positively good or positively bad. They do not "preempt resources which should be reproducing such conclusions by other methods," as Bob claims on page 2. Most people have about 362 days left over in the year to try the other methods. Few people make a career out of such games, few enough not to worry us about resource allocation.

Let me make a couple of other preliminary remarks. One is that any technique for generating hypotheses or selecting contingencies for analysis, any detailed investigation of particular contingencies, research into historical events, or techniques of wool gathering, brainstorming, chatting over martinis or debating in a formal meeting, can lead to overemphasis of the particular, the dramatization of cute

ideas, the hardening of positions through personal controversy, the belaboring of entertaining but irrelvant analogies, the development of vested interests in particular positions, or the drowning out of good ideas by a real or imagined majority view that would make the idea unpopular. All analytical techniques, all research methods, all stimulants to the imagination are dangerous. This includes games. But games are not much worse in this regard than the other techniques.

Bob's notion of "ersatz" history is an interesting one. One can, I suppose, get carried away by a "realistic" sequence of events that never actually occurred. But one can get equally carried away by a sequence of events that did occur. There are some people who think that President Eisenhower was overly impressed with his experience in World War II, and less well equipped to be president in the nuclear age. All kinds of people have been through some one experience that left too strong a mark on them, an experience by which they judge all the problems that arise. One can go off the deep end by taking seriously a single simulated limited war in the Far East that is generated by a team of gamers, but statistically one can make the same error by studying the Korean War in detail. It, too, was just one observation drawn from a potentially larger universe; and if the Korean War enjoys the advantage that it did occur and is known to be plausible and free from historical contradiction, it is still only one occurrence and people who draw historical conclusions from it have to be careful. The same thing with games.

Bob raises a problem about "ersatz" people. How serious the problem is depends on what the game is after. In the games I have associated with his confusion between the positive and normative representations does not seriously arise. Everybody is instructed to play normatively, not to mimic or predict decisions but to play the best they can. Of course, there is an assumption here that the players' individual values are not wholly inconsistent with the values that might be held by real decision makers. But this just does not prove to be a serious problem. The imposition of troublesome "realistic" constraints can often be left to the control team in a way

that frees the red and blue teams to identify themselves with the situation, not with officials whom they represent, and to do the best they can. So I report that this is not terribly troublesome, at least it is not if the players are properly instructed and it is not hard to instruct them properly.

Let me turn now to what I think these crisis games can accomplish. There is one thing they can do that only a game can do, and I shall emphasize that in a moment. It has to do with what Bob calls the "interaction" between teams. There is another thing that games are not alone in doing, but awfully good at doing. That is demanding careful sequential analysis of plans, decisions, events, and intelligence. Very few plans or situations seem to be subjected to a process of "walking through," of dress rehearsal. This is particularly true of plans and contingencies that are political-military, i.e., that involve military considerations and movements, diplomatic considerations and communications, intelligence activities and interpretation, and the coordination of activities over time as well as among agencies. Crisis games typically subject the players to a continuous process over time in which they are both making decisions and living with prior decisions, in only partial control of their environment, committing themselves to actions that have lead times, reaching decisions on the basis of intelligence that is only partially available when they cannot wait for more. People sensitive to a variety of responsibilities collaborate, applying the criteria that are relevant to their own interests, making estimates that reflect their own kinds of knowledge, and putting themselves in a mood to worry about probabilities rather than just a list of possibilities. They really live through a simulated crisis and not only learn things about their plans and their predictions but learn something about the nature of crisis. Some of the things they learn may be trivial when reduced to print. Bob says on page 15 that some of the things that come out of a game are not very novel. But their meaning is sometimes quite novel.

I can easily say that the use of military force in a crisis is typically "diplomatic" or "political" as much as "military," and most people will agree with me. I can say that policy makers' conception

conception of their own objectives may change in the course of a crisis, and people may agree that this has been known for generations. I can say that often for want of a nail a shoe is lost and for want of a shoe a horse is lost, and people will not credit me with a new idea. I can say that some messages get drowned out by noise, and that policy makers become so obsessed with their own perception of events that they may completely misconstrue the way the other side perceives the situation, and this will not sound like a new discovery in the social sciences. These things, if true, can go in one ear and out the other; but if one lives through an intense experience he may learn "lessons" like these in a way that means something to him, in a way that he can talk about with other people who have been through the experience, and in a way that gives him some feeling for the depth or intensity or scope of the problem, or the principal dimensions along which to look for it. Games might be superfluous if real crises occurred much more frequently. I believe there is a striking similarity between the kinds of lessons that get learned from real crises and those that appear to come out of these crisis games. Either kind of experience carries the danger of exaggeration, of course; but if one wants experience, that is the risk one takes.

Aside from this matter of "feeling," one can sometimes discover phenomena that afterwards can be rationalized without reference to the game. Let me give you a favorite example. It is a common observation that each team consciously considers how prudent or bold particular actions will seem to the other team and often chooses actions according to where they fall on the scale of boldness or prudence. It is interesting that teams typically <u>overestimate</u> the boldness of particular moves they make. (I am not ready to say that on the whole a team appears less bold than it is, but <u>individual</u> moves typically appear less bold than they were <u>intended</u> to appear.) This is usually true for both red and blue. The phenomenon is observed by watching both teams simultaneously, i.e., by reading the documentation on both sides, and in the post mortem after the game. This is an important phenomenon if it occurs in real life. Its occurrence in the game does not prove that it occurs in real life.

But once it is observed in the game, one can inquire what accounts for it, and possibly find something that he can then identify with real life.

I believe I can account for a good deal of it. As follows. Typically the moves that a team decides on are selected from an array of bold moves, cautious moves, wild moves, appeasing moves, firm moves, hesitant moves, startling moves and expected moves. The eventual plan or decision may involve a score of components, each of which in some vague way can be measured along a scale that runs from dove-like to hawk-like. And typically the batch of moves is arrived at by argument among some hawks, some doves, and some birds in between. The end result has a certain average "feeling" as to where it is on the scale. The hawks and the doves on the same team may not have the same idea of where it falls on the scale, but each has some idea of where it falls. It is this "average" feeling for the entire configuration of moves that usually falls short, in the adversary team's interpretation, of what was intended or expected.

Now this batch of moves usually contains some that are bold and some that are prudent, i.e., there is a frequency distribution along this boldness-prudence scale. One can also distribute all the moves along another scale, a scale that goes from "definite" to "contingent." We put at the top of this scale those that are definitely decided to be taken no matter what, at the bottom of the scale the steps to be taken in contingencies that are unlikely to arise, and in the middle those steps that will be taken in contingencies that are likely to arise. Call this the "definiteness" scale. Then there is a third scale on which one can distribute the individual actions that comprise the entire plan, and temporal scale. Some moves have immediate consequences; some have delayed consequences; and some moves are themselves delayed in the team's plan.

Now compare these scales. One finds a correlation. The prudent moves are immediate and definite, the bold moves are delayed or contingent. So those actions that are on the prudent side get revealed quickly to the other team. Those that are on the bold side often

do not occur at all, because the contingencies do not arise or occur with a delay. The result is that what the adversary sees it is not the plan but a biased sample, a sample biased in time and definiteness, and therefore in "prudence." If the entire plan were shown to the adversary, he might get a different impression; but the adversary does not get the private plans and discussions of the team, he gets only evidence of what happens. If I firmly resolve to bear your insults with patience but to shoot you dead if you slap my face the likelihood is that I will acquire a reputation for being a meek character.

This is not a "policy result" but a useful principle. It may not always be applicable but is worth keeping in mind. I do not see how it can do much harm, even if it is only partially correct or even incorrect. It is a principle that people can easily comprehend; and people who have been through a game can check for themselves whether their experience confirms that it may be true. As far as I know this principle is "original" and, though it might have been discovered without games, it actually was discovered in crisis games.

There are a number of similar things that emerge from games. They are hardly "policy results" and I think for that reason less likely to do harm if they are misunderstood or if they are in fact wrong. The above illustration is the kind of thing that I believe the games produce, and not "policies" of the sort people usually expect to come out of a game.

Let me make another observation. That is that the most striking observations that come out of a game usually have little to do with the scenario or even the countries involved or the part of the world in which it all takes place. I recall visting a game on its final day, one that I did not directly participate in, and listening to the blue team argue among several alternative courses of action. It was a rather critical point in the game and the principle that emerged that was given a good deal of emphasis in the critique after the game. I was at first astonished, then amused, and finally impressed (and somewhat delighted) with the fact that the same principle had received great emphasis in very similar language in a game that I had observed

whose locale was 10,000 miles away where the terrain and the climate, the adversary, the allies, the nature of the weapons, and the degree of U.S. involvement, were just as different as they could be. The same principle then emerged in another game in still a third part of the world, and not because I made it emerge. I think it is generally the case that the lessons learned are not so particular as to depend on the locale of the crisis, the scenario chosen to initiate the game, the individual participants, or even the character of the crisis.

Bob is worried that something may hold true for the northern cap of Scandinavia and be quite false for NATO's central front, leading to a maldistribution of emphasis and a maldistribution of NATO's resources. I think the contrary proposition is what emerges from these games: two games conducted by different people at different times and different places, with no overlapping, one concerned with Spitsbergen and the other with East Germany, will tend to produce much the same kind of lessons, observations, principles, or whatever you choose to call them. I cannot say that I have systematically tested this proposition by unbiased means, but I have compared the transcripts of critiques of games in very different places, I have compared notes with people who have been in a good many games, I have listened to people discuss the games and what they got out of them, interviewed many of the people, and occasionally followed up post-game planning and other actions, and am of the opinion that Bob Levine has made a very plausible conjecture that just happens to be substantially wrong. At least it is wrong for the kind of games that I have taken part in. I shall be delighted, of course, if someone wants to accuse me of improving games by participating, but even then Bob would have to amend his criticisms to apply only to bad games, not good ones.

Let me turn now to what I think it is that "game-organized" research or planning can do that nothing else can. To begin, let me say what I take to be the critical feature of a game, the thing that makes it a "game." This is that at least two separate decision centers are involved, neither of which is privy to the other's

planning and arguing, neither of which has complete access to the other's intelligence or background information, neither of which has any direct way of knowing everything that the other is deciding on. I am willing to call it a game if it is done by questionnaire, by people sitting at consoles, or by teams sitting in separate rooms, whether it lasts five minutes or five months, whether it involves continuity over time or not. The point is that a "game" as I understand it involves two separate decision centers (and is technically non-zero-sum). What this mode of organization can do that can not otherwise be done is to generate the phenomena of understanding and misunderstanding, perception and misperception, bargaining, demonstrations, dares and challenger's, accommodation, coercion and intimidation, conveyance of intent, and uncertainty about what each other has already done or decided on. There are some things that just cannot be done by a single person or by a team that works together.

One is to judge how "obvious" something is. An analogy is the "hidden face" in the picture. If I draw a face with a hidden picture there is no way for me to tell how hard it is to see the face except to show the picture to somebody. I can't not see the face because I put it there, and the hidden face has the quality that once you've seen it is awfully hard to recapture your innocence and not see it. Alternatively, one can play jazz variations on a musical theme, trying to make the theme almost unrecognizable but not quite, so that the listener has the pleasure of hearing an elusive theme just barely available in the "noise" of the rest of the music. There is no way to tell whether your theme is properly hidden among the alternative tones except by trying it on somebody. (With experience you can learn; but the experience requires hearing other people do it or trying it out yourself on other people.)

It is the peculiar element of collaboration, communication, and bargaining, that is involved in any crisis game, that cannot be captured by "straightforward" unilateral analysis. I may think that if I attack PT boats in four harbors of North Vietnam a few hours

after PT boats have attacked some of my own ships the pattern or gestalt of my action is unmistakable. If you and I together work out the attack so that its pattern is unmistakable, we need a third person to test it for us. We assume that the theme will come through the jazz variations, that the face will be seen in the picture. But we know what we're doing because we designed it; we designed it because it looked "obvious" to us that there was a clear message contained in the pattern; and by simply talking about it we reinforced our sensitivity to it. We are a little like the people who confront the kids with an intelligence test asking them to add one more number to the sequence of numbers, 2, 4, 8, expecting 16 and getting 4 instead because some kid thought it was log-cyclical and insists that that is the "obvious" pattern of numbers.

I remember a game in which a team responded to an air strike on its ground troops with a quick air strike on an enemy airfield chosen at random. (It insisted that the response be quick so that its connection with the preceding enemy air attack would be unmistakable.) The other team devoted 45 minutes to hypothesis about why that particular airfield was chosen, and ended up with a beautiful hypothesis (though in fact it had been drawn at random) and in the process completely missed the association in time with its own attack on enemy ground troops. What was unmistakable "obvious" to seven people who had jointly designed the plan was utterly unnoticeable to seven very similar Americans in the next room who had not designed the plan but only saw it carried out. This kind of perceptual phenomenon can only arise and be observed in a process that is organized as a game.

I may have belabored this point more than I should have in relation to Bob's paper. But in arguments about the treasures or dangers that one may stumble on in games it is significant that there is at least something that games can do or generate that cannot be done or generated in any other way. One cuts himself off from an entire set of phenomena by analytical processes or planning processes that do not involve the interaction of two or more decision centers.

Just to prove that there is almost nothing in Bob's paper that I agree with, except the final paragraph, let me make a few terminal

comments. Bob is worried about the dynamics of plausibility: any story I give him about the Soviets taking over northern Finland becomes plausible just by the telling. This is a terribly important phenomenon. It means, among other things, that not critizing such scenarios may leave us with a monstrously distorted conception of probable events for sheer lack of competing possibilities to judge them by. Two or three scenarios may do the harm Bob refers to. Twenty or thirty may do a lot of good, especially if there is open competition in scenarios. Most important, in my experience, the events that gain plausibility by being put in an imaginative scenario usually are more plausible than they were credited with; it is usually the truth that hurts.

Bob's remarks about Pearl Harbor seem to me irrelevant. For the price of Pearl Habor we could have afforded a lot of games, and one might have made the Pearl Habor attack look plausible, and Bob would be wrong; but I wouldn't make the case for games rest on that possibility. Games won't play music or cook fish, cure a man of stuttering or improve my children's French, just as they may not predict Pearl Harbor, but unless Bob can show that games would have accentuated the tendency to ignore Pearl Harbor his remark is immaterial. Games might have taught us something else useful.

Bob is worried about the unreality of scenarios pitched far in the future. They are indeed a nuisance. A game pitched five years from now needs an almost encyclopedic scenario to cover all the relevant information: did DeGaulle die, who got to the moon first, are the Chinese in the U.N., did Cyprus ever get cleared up, whatever happened to Civil Defense, who's the opposition party in Britain and what do they feel about the M.L.F., did we phase out the rest of the B-47's, has Cuba been quiet, did anybody denounce the test ban or cheat on it, have nuclear weapons been used in anger, what are the troop dispositions in NATO and the fleet dispositions in the Far East, is there still a trade embargo with China, whatever happened in the Congo, what are Nasser's relations with the Shah of Iran or have both of them died? And so forth. The advantage of a scenario that occurs immediately or in the near future is that one can specify

a few significant events and then say that everything else is as known to the participants, and they have in their minds all the background information they need. (Even if nobody quite knows what Nasser's attitude is toward the Shah of Iran, people at least know what the uncertainties are and they can be told to abide by their own expectations.) For that reason it is an enormous convenience to stick to the present or the near future. One cannot do that, of course, if he wants to run a crisis that results from competitive lunar landings, from the second stage of disarmament, or from the acquisition of nuclear weapons by ten or fifteen countries or one that takes place in the context of fully deployed ballistic missile defenses. And the reason is not that games, like everybody else, are poor at predicting the future; it is just that it is a terrible nuisance in a game not to be able to rely on common knowledge and to have to specify an enormous amount of background information that people must learn for the occasion by reading the scenario.

Bob's remarks about the Berlin crises seem to me a little beside the point. He says the Berlin crisis games of 1961 had the advantage of recent history. I do not really think that the wall, or any other immediate preceding history, made all that difference. They had the advantage of a scenario that was pitched in the present or the near future. This was a great convenience and contributed much to the game, but it did not mean that the results were therefore fresh rather than out of date. The 1961 Berlin-crisis games are about as relevant to 1964 as they were to 1961. If I wanted to study a Berlin crisis of 1967 -- not to focus on some striking difference between now and 1967 that I wanted to introduce but just to have a game that would yield results that would last three years -- I would pitch it in the immediate future, not in 1967. Most of what comes out of a game is not invalidated by the passage of time (just as a lot of it is not invalidated by going to another hemisphere). A great deal of what came out of the 1961 games is just as valid now as it was then. It may not be as fresh; it may have faded in memory; and there might be an argument for doing it again. But that is not entirely because the situation has changed and the results are out of date. It is

because the main beneficiaries of the game are the participants, and if a new bunch of people are working on Berlin maybe they deserve a new game. Actually, a game sited in Yugoslavia or Spain, Spitsbergen or Finland, would have a lot of relevance for Berlin today, or even in 1967. Not nearly as much as the Berlin game, but enough to refute Bob's point. It is a happy coincidence with a game "predicts" an event or identifies some specific tactic or situation that occurs shortly after. But that is not the main result of the games. I would expect the games of a few years ago that were sited in various parts of the world to be about as relevant to Vietnam today as they were for the areas they were concerned with at the time they took place.

Bob raises the question on page 9 why we should have a red team at all, rather than just a malevolent control. There is a good answer. You lose most of the benefits of the game. I have tried it. All of the awareness of a conscious adversary who is somewhat in the same boat -- all of the problems of collaboration and communication, of accommodation and intimidation, of designing a pattern of actions to convey something to an adversary and evaluating the situation the adversary is in, all the sense of risk and danger, the apprehension of over-response and under-response, and the opportunity to exploit an adversaries apprehensions -- disappear when a team knows that it is just playing against a control team that has no stake involved and that has complete access to the team's thinking. According to the definition I used above, it ceases to be a "game." And I would add that the positions of the blue and red team are not as different as Bob makes out. There is of course a much greater sense of involvement on the blue team because we are in fact all blue. And this is why it is important to have the red team properly motivated. Blue gets nothing out of the game unless it knows that red is sweating it out also. If red is "irresponsible," the essential "game" element is missing altogether. Blue is just solving problems unilaterally, and may as well go do it by itself.

Bob's extensive worries about stimulating the presidency relate ~~only to the kind of game in which people are supposed to simulate~~

the presidency. I much prefer to have the blue team be the presidency, in a corporate sense, letting control stop them when they do something that is totally out of American character or unacceptable for reasons of domestic politics. Subject to constraints imposed by control, the blue team is the U.S. government and is supposed to do the best it can (in the games I've been associated with). I know that there are games organized differently and some of Bob's criticisms may apply to them. But they need not apply, and do not, if the game is organized so that this mimicry of a president or a premier is waived aside by the instructions to the participants.

I do not understand why Bob pays so much attention to hypotheses about "blue attitudes" on pages 12 and 13! There are, I believe, some phenomena involving attitudes that a game may get at. One is the way attitudes are formed in a group. Another is the way attitudes change over time. A third is the way attitudes are changed by one's perception of what the adversary is doing. But I did not know that games were being widely used to elicit basic attitudes, in the sense in which Bob describes them, so I do not take his criticisms seriously there.

I have enjoyed disagreeing with Bob Levine for a change. My disagreement is suspect, of course, because I bought a lot of stock some years ago in the enterprise whose product Bob is warning the consumer to be wary of (and whose recruitment claims Bob is disputing). Still, I vividly remember discussions at RAND in 1959 about the perils of games, and more in 1961 when I was getting involved. I believe I heard just about every argument against games then that has been voiced in the English-speaking world. I remember being worried. My worries have declined, almost disappeared. With respect to the crisis games I have witnessed I do not recall serious complaints about exaggerated or mistaken policy conclusions, about the high cost or the waste of time, about exaggerated emphasis or misplaced prediction, about actions taken in consequence of the game that were seriously mistaken, or about alternative modes of research being preempted by the game. I have heard criticisms in principle like Bob Levine's; I have heard games deprecated for the harm they might do. But I do

not recall that the dangerous possibilities occurred or that the serious warnings were borne out.

Surely I am biased; and surely people who thought some of the games a menace or a nuisance would have abstained from telling me so. Still, I am becoming relaxed, and confident that the bad things that games might do they usually don't do.

As I said at the outset, Bob is expressing reasonable conjectures about what games might do or fail to do. Everything he says is plausible. But there has been quite a bit of experience now; and, while I cannot speak for all of it, and although I can think of some games that seem pretty silly, I believe the record shows that certain kinds of "crisis games for adults" have been persistently fruitful, have not had disastrous side effects, and have covered their costs very nicely by a wide variety of side benefits.

There may be games of which that is not so. But then let's improve them. I feel about games somewhat the way Alcoa is reported to have felt about aluminum when it was first introduced. They are terribly good for some things, not very good for others, and there is a danger that enthusiasts will discredit them by cooking some things that turn black in the pot. Some people run games who don't know their limitations; and, while I think they do remarkably little harm, they also do no good. I'd be happy to write a sequel about situations and problems for which game organization is inappropriate, superfluous, and even in danger of discrediting the technique.

But, as I have said, the only part of Bob Levine's paper that I can agree with is the final paragraph on page 21, which somewhat modified, says one of the most important things that can be said about these games. Another thing, equally important, is that the main beneficiaries of these games are very often participants, not people who read about other people's games.

## CRISIS GAMES FOR ADULTS AND OTHERS

William M. Jones

I have three reasons for adding my comments on crisis gaming to those recently made by Bob Levine and Tom Schelling. First: Crisis games are run in various organizational environments for various purposes. To me it seems to follow that they should be discussed, praised, or criticized only on the basis of how efficiently and how successfully they meet the needs of the environment. Secondly: I wish to heartily endorse Tom's use of "the hidden face in the picture" analogy in refutation of the criticism that games tend to give profound insights into the obvious. Last, I cannot resist an opportunity to protest Bob's implication that gaming is the fallen woman of research procedures because it is both seductive and expensive.

## I. THE INFLUENCE OF ORGANIZATIONAL ENVIRONMENT ON THE DESIGN OF CRISIS GAMES

In my opinion, the two most salient features of crisis games (or any other sort of team game) are: (1) The rapidity with which the bits of information on the problem area, known at the outset to individual participants, enters the common information base of the whole group; (2) The rapidity and accuracy with which individual members of a game team achieve an understanding of (although, not necessarily, in agreement with) the feelings of their team mates about the situation/problem being simulated. These two features are, I believe, the main bases for the phenomena remarked on by Tom: The noticeable post game improvement in the ability of participants to communicate meaningfully.*

Two different crisis game environments are mentioned by Bob in his paper and still another is, I feel, being discussed by Tom. I am under the impression that Tom is basing his comments on the crisis games of the Washington environment. I gather that Bob is basing his comments on the reports of crisis games conducted at RAND and the manual simulations conducted by Harold Guetzkow. While I can pretend to only general familiarity with the work of Guetzkow, I can speak with some authority on the other two.

The crisis games, played in the Washington environment, have certainly been productive in the ways described by Tom. It was this desired effect that motivated the Joint War Games Agency of the Joint Staff to participate in the initiation of these games.**

The games recently conducted at RAND are (at least to me they are) directed toward the expeditious transferral of information and

---

*This is not to say that people who have shared participation in a game are forever after in complete accord.

**Remember that in 1961 the communications barrier between the military and the new set of civilian authorities was at its highest. The military services, faced with a new group with a new jargon (cost/effectiveness -- program packages -- counter force -- et. al) had similar feelings to those of the high school freshman whose mother has just said, "Johnny, say something in algebra."

views among a group of researchers and analysts of rather disparate backgrounds, skills and interests. In discussing the results of these experiments (and they probably should still be viewed as being experimental) it would seem appropriate to compare and contrast them with other procedures, used by multi-disciplinary research organizations to achieve similar results. Perhaps, although not certainly, the lesser quantity of detailed data available in a small research organization, compared to the voluminous and varied detailed data available at the seat of government, makes the diffusion of the available data more important at RAND than in Washington. In either case, it seems important.

Knowing very little about the Guetzkow simulations, their environment and objectives, I find it easy to be critical. Their use of analysts as observers of the simulation rather than as participants seems to me to be a weakness -- analogous to kibitzing a poker game being played for matches -- a rather far remove from real life. On the other hand, I note that the simulations have been in process for a considerable period of time so, presumably, the sponsors find them useful.

## II. ON THE DEVELOPMENT OF A PROFOUND INSIGHT INTO THE OBVIOUS

Little need be added by me to Tom's discussion of the utility of this feature of gaming. The "hidden face in the picture" analogy precisely typifies and illustrates a gaming function that I have frequently wished, but found difficulty, to describe. Perhaps there is some small display of normal human pride involved in the assertion that what is obvious to one today was obvious yesterday. If so, I can only recommend, exhibit and ask for some degree of charity.

Under the same heading I will comment on Bob's charge that games seldom present anything novel which might be called a contribution of the game to the universe of analytical knowledge. This seems to be an excessively high standard against which to measure any study technique. It might be wise to agree to "bury this one" for fear that our employers hear of it and adopt it as a standard for measuring our output -- with rather unfortunate results.

## III. A FEW, VARIED COMMENTS

(1) On Scenarios

Developing scenarios of the future, plausible or probable,* is a rather difficult process of creation, much more akin to art than to science. Having little claim to artistic talent in this (or any other) respect my personal bent is toward the writing of a scenario using a technique used (with considerable success) by many weather forecasters: persistence. Using this technique the weather forecaster carefully plots and analyzes the conditions over a wide area, determines the trends that are developing, and then forecasts that tomorrow's weather in the local area will be like today's. More seriously, perfectly plausible, useful and not improbable game scenarios of a situation three or four years in the future can be had by simply initiating a game using today's newspapers -- interjecting only tomorrow's equipment. Arab-Israeli controversy has existed for some time and it seems probable that it will be in existence in 1970. Turkey and Greece have not seen eye-to-eye on the Cyprus situation for a considerable period and, unfortunately, this controversy is likely to persist. German nationalism has been a strong, persistent factor in Western Europe for about a century and it seems rash to expect one generation of separation to eliminate it entirely. Notice here that I am not saying that games, played from scenarios which are extrapolations of today's problems, are perfect for determining tomorrow's force posture and weapons characteristics. I am saying that I can think of worse procedures (re-fighting the last war, for example) and I know of no other way in which I have even such a limited amount of confidence.

(2) On Role Playing

This, to me, is one of the more sensitive variables on crisis gaming. How one deals with this tends to be a matter of personal taste. I, for one, prefer games in which the precise role to be played

*I know a scenario of a probable future situation when I read one. Unfortunately I've never had much success at convincing others of its probability nor have others had much success at convincing me.

is neither directed nor announced. Similarly I think it rather pointless to direct at the outset of a game that the players simulate decisionmakers as they think they are or as they think they should be. (Although if I must make a choice it would have them acting as they think the decisionmakers should.) My reason for this preference is simply that few people are good simulators of others in either sense and it seems a bit pointless to direct them to try. In the course of a game people generally tend toward making those decisions that they think should be made. Incidentally, it is my opinion based on involvement in quite a number of such games that researchers are not noticeably worse (though different) simulators of civilian officials than civilian departmental subordinates; that retired officers and staff officers are not very good simulators of military commanders; and that, comparing game decisions to publicly announced and implemented decisions, actual decisionmakers in a game situation are not precise simulators of themselves in their official capacity.

If this lack of precision is taken to mean that games are imperfect predictors of the future, I can only agree. If this is taken to imply that the changed and enriched views of certain common features of crisis situations produced by games is valueless, I disagree. Of course, any results, conclusions, inferences drawn from games should be -- and sometimes are not -- tested for reasonableness against experience and good judgment -- but the same can be said of any study procedure.

(3) On the Credibility of Game-Generated Conclusions

Bob, in his paper, states, "If a game result appears in language that can be read as a confirmed policy conclusion, there is a substantial chance that it will be so read."* From my ten years of experience with games of almost all types I cannot recall a single instance where this has been the case -- except -- when some naive person has previously held that belief and is trying to "pile up additional evidence."

* Page 17.

Contrary to what Bob seems to fear, the number of policy makers who view policy conclusions (or any other kind of conclusions) as having added credibility, solely because they are game products, is negligible. My guess is that the proportion is about that that blindly accepts a policy conclusion solely because it is expounded in a RAND RM.

## IV. FINALLY

I join with Tom in concurring with Bob's closing assertion that "game evidence" is not to be cited as being authoritative -- and that, in fact, when one is presenting conclusions it might be well not to mention that one had gamed. On the other hand one need not be embarrassed to concede that one has gamed in the course of the study that led to the conclusion being presented.

# CRISIS GAMES: A REJOINDER TO TOM SCHELLING AND TO SOME EXTENT TO BILL JONES

Robert A. Levine

## I

The one thing that stands out most clearly from Tom Schelling's pitch for crisis games is that, whomever else crisis games stimulate, they certainly stimulate Tom. Since one of my criticisms of crisis gaming was based on doubts as to the real stimulation they provide, this may appear to disprove one basic point at the very outset. It does not. It is true that crisis games stimulate Tom, but <u>everything</u> stimulates Tom. As an example, I can point out that the machinations of Tom's four sons stimulate him to think of deterrence theory; the machinations of my three stimulate me to think of mayhem. I know Tom's boys and can state that their devilishness is certainly in a ratio of no less than 4:3 as compared to my own.

And the trouble is that I suspect that I am far more typical than Tom. Tom can cite chapter and verse as to where crisis games have stimulated him, but I look at two larger samples -- that of hypotheses derived by people in general from crisis games, and that of hypotheses derived by Tom Schelling from all sources of stimulation -- and conclude that the variable is the person and not the method. I happen to be of the school of thought that believes that progress in analysis and theory (in any field) is dependent on a very few outstanding individuals, and if crisis games provide additional stimulation for Tom Schelling they may be worthwhile on this account alone. But what I was writing about was a more generalized case.

## II

Beyond this, however, I think that Tom and I were talking about rather different things -- different game environments as Bill Jones points out, and also different functions of games. I tried to limit at the beginning of my piece the area I was covering, but I did not stay completely within my own bounds. To restate what I was discussing, then, it was: <u>Games as generators of research conclusions (and to a lesser extent, hypotheses) concerning the way in which nations act or should act in crises, and concerning the relationship of such actions to military forces</u>. And further, although it was perhaps only implicit in my paper, I was talking about research conclusions derived by games in research organizations

(e.g., RAND, Guetzkow's shop). I contended that, because of the suspect nature of phenomena such as ersatz history and ersatz people, such research conclusions must also be suspect to the extent they are based on game "evidence." My major conclusion was in a final paragraph, with which both Tom and Bill agreed, to the effect that if gaming is used, any research conclusions should stand and be defended entirely on the basis of non-game "real-world" evidence.

It seems to me that, in defending games, Tom was talking about two gaming functions which I was not attacking. He was talking about games for educating decision-makers; and he was talking about games as generators of general theory about crisis interactions. This latter is close to what I was criticizing, but it is not the same thing, as I shall try to show in Section IV, and the difference is important.

## III

It is clear from Tom's paper that he plays games with different people than I do. The games he is referring to are games played by officials of the Defense and State Departments, etc., and very high officials at that. I gather that it would be inaccurate to say that these games have been played at the journalistic "highest level," but they've been pretty damn close. It would certainly be correct to say that Tom's games are played by "decision makers." And most of what Tom is talking about is <u>educating</u> these decision makers as to what crises look like and how people act in crises. This education is of two types: classroom education, in which decision-makers are taught in a vivid and unforgettable way some crisis principles for which they had previously had no appreciation or only an intellectual appreciation; and laboratory training, in which decision-makers operate together as a team, discover some of the constraints which stem from team operations, and obtain a deeper realization of what their colleagues are like. Tom, then, is discussing gaming as enabling "classroom" instructors to communicate <u>with</u> a group, and gaming as establishing a mode of communication <u>within</u> a group which may have a need for quick intercommunication in a real crisis. This sort of education Tom calls a by-product of gaming, but it is also implicit in much of his

basic defense. As basic accomplishments of games, for example, Tom points out (p. 6) that they make possible "walk-through" dress rehearsals of contingency plans, and (p. 7) living through intense experience and thus making crisis meaningful (and making it possible to talk to others who have been through the same thing.)

These contributions of gaming at the decision-maker level should not be denigrated; they can be extremely important. The question is, however, what applicability does this education function have to games at places like RAND?

The payoff on training RAND types as decision-makers is questionable. We make our contribution of Hitches, Enthovens, and Rowens (and it should be noted that Hitch, Rowen and Enthoven were educated at RAND without particularly participating in crisis games), but this is a by-product of RAND. To become decision-makers, most RAND types need the crucible of Washington, and I think that there is general agreement that we do not possess the hardware here to reproduce the heat. The one example I recall of RAND people sitting in Santa Monica and trying to help Washington during a real crisis was, at best, kind of silly.

In this connection, it is interesting that in answer to my comments about positive and normative role-playing, Tom says that in his games, the question "does not seriously arise. Everybody is instructed to play normatively, not to mimic or predict decisions but to play the best they can." But in games at this level, the question need not seriously arise. I am perfectly willing to admit, as both Tom and Bill point out, that General X or Assistant Secretary Y in a game do not simulate exactly General X or Assistant Secretary Y in a real crisis, but for these gentlemen, trying to act as they should and trying to act as they would, come to the same thing. The gap in Washington is not between should and would, but between trying and acting.

This problem, which is not important in Washington, is at RAND, however. Here, should and would can more easily diverge. Harvey Averch and Marv Lavin, in an early draft of their RM-4202-PR on gaming,* state in the summary that "We believe manual gaming techniques to be uniquely useful for studying intricately constrained decision-making -- particularly when one's focus is on what U.S. decision-makers will do in a crisis

* See above.

with the military forces on hand, rather than what they could do." (underlining theirs.) The fact that they removed this statement in the final version is unimportant -- I'm not trying to pin Harvey and Marv on anything. What is important is the illustration that while it is easy for Tom's players to decide to play normatively, it is not so easy for us. It is by no means clear for us that, if we game, attempted positive simulation is incorrect. If we are to get anything out of it, we may have to make some attempts to constrain our own theories not only by force-composition realities, but by Washington realities.

This leads to the next question -- games to train RAND types not as decision makers but as RAND types. Here it is difficult to feel particularly negative. If games aid in establishing communication among members of a research team, fine. If they help all members of the team achieve a common body of knowledge, fine. But I allude to my final caveat, with which Tom and Bill agree, about dropping the fact of the game from the research presentation. One should no sooner report that the team started out to get a common basis by gaming than that the team started out to get a common basis by all reading pages 1-30 of Herman Kahn's On Thermonuclear War (Princeton University Press, 1960). Both may be very useful things to do, but they are not relevant to the report.

One example of the utility of games for group research, however, is discussed by Tom and endorsed by Bill, and I think needs some comment. This is the "hidden face in the picture" analogy, in which the researcher cannot see all the implications of what he has done, or sees some obvious implications which someone else cannot see at all. Again, I would not deny the phenomenon, nor would I deny the utility of games in exposing the "hidden face." But the key is that although games can provide a method of exposing one's thoughts to others, it is doubtful that this method is a unique* contribution of games, as Tom and Bill seem to imply. The first thing that will be done with this D, before it goes to the typist, will be to show it to both the Jones brothers, Bill and Norm -- and not for them to tell me how wonderful it is, but to look for hidden faces. And, judging by the number of things I am asked to read, I am hardly unique at RAND in getting others to look for hidden faces. It's

*Bill says that he claims that games are outstanding in this respect but not unique.

still really not too tough to get a good argument going here; it's a lot tougher to get an RM through all the levels of review.

As a matter of fact, Tom's specific example of the hidden face phenomenon is an odd one. He suggests that "I may think that if I attack PT boats in four harbors of North Vietnam a few hours after PT boats have attacked some of my own ships, the pattern or gestalt of my action is unmistakable. If you and I together work out the attack so that its pattern is unmistakable, we need a third person to test it for us." But the point here is that, yes, we need a third person to test it for us, but to say that submission to a third person is a "game," is to stretch the bounds of the definition a bit. The distinction here is not between testing by a game or testing by using a critic; the real advance in the state of the art would come by having a North Vietnamese criticize it for us.

In sum, then, I agree with Tom's statement that "the main beneficiaries of the game are the participants," and I agree that in the case of participants in decision-making positions, the benefits are likely to be high. In the case of participants at research organizations, however, I have my doubts, and would fall back on the caveat that when crisis games take place in these circumstances, it should be clear in all minds that it is for the participants and not as a tool of research as such.

## IV

Actually, of course, Tom is claiming more for games than mere training value. He is claiming that games are useful for deriving general behavioral principles applicable to crisis study. And if the claim is taken carefully, and understood as I think he means it, I agree.

For one thing, in Tom's participation in high-level games, he is deriving value from studying how decision-makers react to a crisis. This goes beyond the stimulation the games provide him; as he discusses it, it is true research into decision-making by decision-makers in crisis situations. But this is not in opposition to what I wrote. In my concluding "constructive" section, I suggested that the best way to do

fundamental research into decision-making in crises is to study real decision-makers in crises. Frankly, because of my RAND worm's-eye view of gaming, it had not occurred to me that a powerful tool for such study would be watching the decision-makers act in crisis games, or participating with them. Tom having suggested it, however, I don't think it unfair to claim that I can consider it a case included in what I advocated, rather than in what I criticized. Further, I gather that those participating in the games Tom has in mind included not only executive decision-makers of the "expert" type, but also legislators and executive decision-makers of the "political" type (to draw a very arbitrary distinction), which was also along the lines of what I was suggesting. I would still worry about the fact that the President as such makes a big difference, that the President or the Presidency is very difficult to understand, never mind simulate, and that attention ought to be paid to the ideas of those who may have thought of themselves as President. But now we are in the realm of differences from the real world which are narrow enough so that perhaps they can be systematically allowed for.

The direct study of decision-making in crisis, however, cannot be accomplished by games at RAND. Tom seems to see another kind of specific research value which might come from RAND games. This is the deriving of general and rather theoretical principles about behavior in crisis or other conflict situations. An example he gives in this D is the phenomenon that "teams typically overestimate the boldness of particular moves they make" -- that the average between hawks and doves looks more dove-like than hawk-like. He explains this by pointing out that dove-like actions are immediately communicated to the opponent, whereas hawk-like actions are contingent and frequently never executed.

This is an extremely interesting observation, a novel observation so far as I know, and (not surprisingly from Tom) a very perceptive observation. What is more, it is undoubtedly true that wherever Tom might have derived it, he did derive it from gaming. But, what is important for this case is the conflict between two teams of people, not the fact that the games in question were concerned with Berlin crises, North Flanks, South Flanks, or anything else with specific military-political

content. As Tom himself says, "the most striking observations that come out of a game usually have little to do with the scenario or even the countries involved or the part of the world in which it all takes place." Some of Tom's work in the last years has been with a rigorous bargaining game utilizing students in a context that has nothing directly to do with any real-world conflict. And the principles derived from this, by Tom and his collaborator Bernhard Lieberman, have been of the same important but general character as the hawks-doves one mentioned above.

I would argue that the real research value of games outside of Washington is similar to the above, but that so long as these games are imbedded in a pseudo-real context with exciting ersatz history and ersatz people dominating the terrain, attention will inevitably be drawn away from this possible real contribution. Nobody yet, not Tom or Bill, has claimed that the specific context of research games provides much (or anything) additional on what to do in a crisis situation. This, together with the seductive nature of this context* in making some people think it is the real world (I had some examples in my last D, and cannot refrain from one more in the next section of this one), implies that perhaps the prescription should be that games in RAND should turn from the concreteness of scenario and role-playing, toward the much more abstract.

V

Tom claims that I have "made a very plausible conjecture that just happens to be substantially wrong," and suggests that I amend my criticism "to apply only to bad games, not good ones." This, of course, is what I have done in this piece, although I have tried to classify "bad" and "good" according to what the games are used for, as well as to carelessness in their use. What is left, in the residue I continue to criticize, is the use of games at RAND for looking directly at specific crises phenomena such as, for example, the "fire-break," or allocation of forces

between NATO's Central Front and Flanks. The use of research gaming for such specific-context phenomena is really all I was shooting at in the other D, and it continues to be all I am shooting at.

What is more, if Tom wants to press his point, I will be willing to admit that maybe he can even distill something useful concerning such phenomena out of RAND games -- but I would still claim that this is a function of Tom and not of the games. Perhaps there are good examples of useful things along these lines which have come out of games at RAND or at similar places, but at this point the discussion must become classified and perhaps oral. On this basis it might be worth holding a seminar the next time Tom is in Santa Monica.

And to defend my conjecture about the dangers of games as being more than plausible -- as being repeatedly demonstrated -- I need some examples to put against the possible ones of what Tom can do or what classified research has done that I don't know about. In my last piece I cited as a horrible example the statements of Guetzkow, and as an example of the dangers of seduction, some statements, later changed, from a RAND draft. Let me add now (in the horrible category, I suspect) a document from SDC entitled Plans 1: A Vehicle for Studying National Policy Formation in a Less Armed World,* which has fallen into my hands through arcane means. (Someone sent it to me.)

> PLANS is an acronym for Planning and Negotiation Studies. It refers to a series of continuing simulation experiments addressed to problems of national policy planning and negotiation. Specifically, PLANS-1 is an experimental vehicle for studying these problems under conditions of a "less-armed world."

Plans-1 is a game in that it has both the ersatz history and ersatz people phenomena I have discussed. It is not two-sided -- it has no Red team -- but it turns out to be many-sided -- it has teams representing American Business, Labor, Civil Rights Groups, Military, Internationalists, and Nationalists. What it seems to be is a serious attempt at simulating the American political system -- no less.

And in defense of my avowal that games for such research purposes are seductive and dangerous, let me quote from the "Results" section, after the first runs of this "vehicle." And remember that SDC is a very

*System Development Corporation, Washington Division, TM-WD-119, by R. Boguslaw, R. H. Davis, and E. Glick.

respectable research organization, a child sometimes thought to be in the image of its parent.

> The experiments described herein should be considered as pilot rather than definitive studies. The goal of a pilot study is to help insure that relevant hypotheses, variables or ideas are not prematurely excluded from consideration.
>
> Our analysis of the results is, therefore, focused on generation of hypotheses and explication of ideas suggested by observations made during the preparation and conduct of experiments -- ideas and hypotheses which appear to merit more intensive examination. Some of these can be tested further with the PLANS experimental vehicle. The examination of other hypotheses and ideas may well require use of the so-called real world or experimental vehicles other than PLANS.

While it is nice of PLANS to give the so-called real world a possible role in testing hypotheses, it is obvious that the burden of testing and providing conclusions will be on PLANS and its friends. And perhaps even more interesting than this statement of future methods for drawing conclusions is that, after the standard disclaimer that these first runs are only for generating hypotheses, come thirteen pages of statistical and graphic statements of the numbers generated by these "hypothetical" first runs. I am willing to make a small bet that someone will be quoting these numbers as conclusions before long, if it is not being done already.

Now, let's face it, games are not alone in making vast claims about synthetic real worlds. Raytheon has done the same thing on a computer without the benefit of ersatz people or (I think) ersatz history. So I am not claiming that research games are uniquely dangerous, any more than I was willing to concede many of the unique values sometimes attributed to them. RAND has a well-deserved reputation for being careful on claims; nonetheless I would suggest that RAND games move toward the abstract. If they cannot do this, then perhaps they should do things like moving away from ersatz history by starting games with today's newspapers, as suggested by Bill Jones. This suggestion is close in spirit to my point in the last piece that the 1961 Berlin game was remarkably successful as a policy generator because it had real history to go on.

And, withal, if these suggestions are not taken up, as I suspect they will not be, I conclude by repeating the agreed caveat:

> Game if you will, But in presenting policy results, don't tell anyone that you gamed. Present it in essay, model, or other analytical form, without mentioning the game. If it is convincing in this form, then the game has been as good an instrument as any. If it is necessary to fall back on game "evidence," however, then the whole process is of very doubtful validity.

Perhaps this warning is obvious, as it may have been to Tom and Bill, and perhaps it is not needed at RAND. If so, good.

POSSIBLY THE LAST WORD ON GAMES

Robert A. Levine and Thomas Schelling

HARVARD UNIVERSITY

CENTER FOR INTERNATIONAL AFFAIRS

ROBERT R. BOWIE, *Director*
ALEX INKELES
HENRY A. KISSINGER
EDWARD S. MASON
THOMAS C. SCHELLING
RAYMOND VERNON

6 DIVINITY AVENUE
CAMBRIDGE 38
MASSACHUSETTS

RECEIVED
1964 OCT 20 AM 8:31
MAIL & RECORDS

October 16, 1964

Mr. Robert Levine
The RAND Corporation
1700 Main Street
Santa Monica, California

Dear Bob:

At last I have something to get really concerned about--the definition of "game." On the fifth page of your rejoinder of September 29 you acknowledge the value of having a third person test the message or Gestalt or pattern in some complex of events. You assert, though, that "to say that submission to the third person is a 'game' is to stretch the bounds of the definition a bit."

No, it is not. A game can involve one person for one minute, and one decision. When I submitted a questionnaire involving such questions as where to meet somebody if you forgot to make an arrangement, each such question was a "game." The reason why it is a game is that the respondent (the "third person" in your quoted passage) must not be told the solution--possibly must not be told even what to look for--but must come to it fresh, as a separate center of consciousness and decision, not contaminated by the reasoning or the images that went into the thing he is asked to react to. He is not to be shown the face and asked whether he shares our evaluation of whether he would have seen it had we not shown it to him; he must be told to look for something (possibly told that he must look for a face) and we must then see whether he finds it. It is a "game" the way a riddle is; it is a puzzle containing clues, and the answer must be not too easy and not too hard; in particular the "correct" answer to a riddle is the answer we had in mind. The purpose of a riddle is to see whether somebody can read a hidden message (and, while we may give some credit if he can find any message at all, his answer must at least be in the spirit in which the riddle was asked).

In other words, when you test it on the third person you must organize the test as a "game" or the test won't work. Or, to be a little less dogmatic, it is a better test the more you organize it like a "game."

ANS. BY Pat
LETTER NO. L-21520
DATE 10-27-64

RAND—Santa Monica
PRESIDENT
VICE PRESIDENT
OPERATIONS MGR
LIBRARY
MFS
PERSONNEL
KEYPUNCH
SECURITY
TRAVEL
FILE NO. 3735-8

Even having a North Vietnamese criticize it for us, as you prefer in the particular test you cite, may not be much help unless you organize it as a game. It may be some help, because you may be afraid that North Vietnamese can't even see the face in the picture when it is pointed out to them, because they draw pictures differently from the way we draw them in the West. Somebody in the Economics Department at RAND has a splotchy picture that is supposed to look like Christ, I believe, if you can find the hidden form. I couldn't find it even when I was told what to look for. I probably could have, if somebody had traced it out to me, or told me whether I was looking for a face, a crucifiction, something that comprised the whole field or a small image camouflaged in a larger one. Just how much communication or information to allow the second party, or to allow between the participants, is a matter that should be controlled. And with North Vietnamese you might organize the one-shot game differently.

We do not know in advance whether you will learn more by having a North Vietnamese colleague work with you in a non-game collaboration, or by having an American colleague who is not privy to your intentions respond to a game-organized inquiry. The latter would still be worth doing even if you have the North Vietnamese colleague design the operation with you.

Since my argument has absolutely no bearing on the merits of the games you have been arguing about, but is purely a definitional matter that I offer in order to isolate a crucial difference between the technique that I call "games" and the more straightforward analytical techniques, I recommend you concede the point, unless you have in mind accumulating enough material to start a journal devoted to Levine on games.

If you think it belongs in the record, I suggest you circulate this memo to your subscribers.

Best regards,

Thomas C. Schelling

TCS:ac

27 October 1964 L-21520

Mr. Thomas C. Schelling
Center for International Affairs
Harvard University
6 Divinity Avenue
Cambridge 38, Massachusetts

Dear Tom:

I concede.

Sincerely,

Robert A. Levine
Logistics Department

RAL:dh

RAND/P-7719

# CRISIS GAMES 27 YEARS LATER: PLUS C'EST DEJA VV

Levine
Schelling
Jones