Naval War College Review

Volume 54
Number 2 *Spring*

Article 9

2001

"Good Games"—Challenges for the War-Gaming Community

Stuart H. Starr

"GOOD GAMES"

Challenges for the War-Gaming Community

Stuart H. Starr

In 1999, the North Atlantic Treaty Organization issued a technical report laying out a "Code of Best Practice" for command and control assessment.[1] Although specifically aimed at command and control, this document offers a framework for thinking about the changing nature of war gaming. In the opinion of numerous practitioners and observers, war gaming has reached a turning point: the changing basis of international security at the dawn of the twenty-first century makes gaming an especially valuable tool, but a fundamental reformation of gaming is required for it to achieve its potential.

The Code of Best Practice, as a unifying and overarching framework, allows us to take stock of the present state of war gaming, to highlight the primary challenges that the war-gaming community faces, and to propose steps to improve every aspect of war gaming. It makes four central points.

First, as shown in figure 1, the framework of a good war game should be broadly based on the principles of sound operational analysis. Thus the cornerstone of any game must be a clear and unambiguous formulation of the problem to be addressed—the reason the game is to be played. A game's sponsors need to articulate very clearly the real issues of interest so that designers may develop (for the sponsors' approval) a conceptual framework within which these issues can be suitably analyzed.

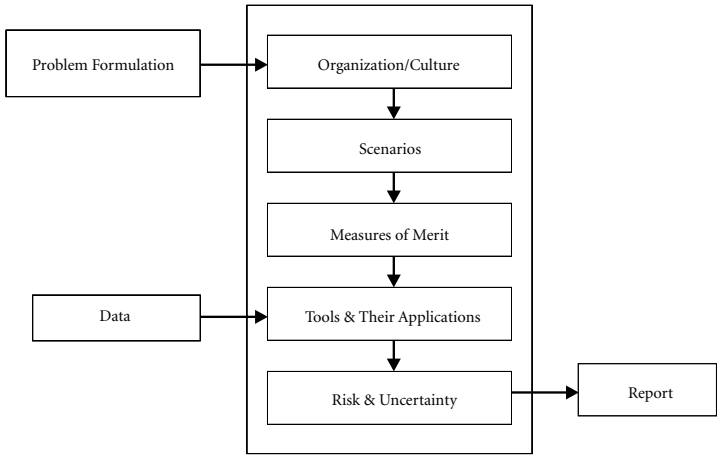
Second, as the Nato document instructs, the game's designers should identify and address organizational and cultural issues that emerge from the conceptual framework. What assumptions are to be accepted, for example, about the values, behavior, and decision processes of the various players?

Dr. Starr is Director of Plans at the MITRE Corporation in McLean, Virginia. *An earlier version of this paper was prepared for delivery at a war-gaming conference at the Naval War College in March 2000.*

Naval War College Review, Spring 2001, Vol. LIV, No. 2

Third, the war game must devise relevant scenarios. As is widely understood, no single scenario is adequate for the full range of issues found in a major game. Sponsors should expect to be presented with "families" of scenarios; a systematic and efficient mechanism to generate appropriate alternative scenarios allows a game to focus on the most interesting aspects of the problem being studied.

FIGURE 1
WAR-GAMING FRAMEWORK



Fourth, Nato's Code of Best Practice envisions the use of "measures of merit" to draw out insights about the game's results; for contemporary scenarios, hierarchies of interrelated—and, increasingly, nontraditional—measures are necessary. These measures, in turn, require the collection of appropriate data and the application of suitable analytical tools to be useful. For instance, ancillary tools can be used to perform analyses before a game (perhaps to define fruitful parts of "scenario space"), during it (to assess "moves"), and after it, especially to relate outcomes to measures of merit. The Nato Code considers it vital to perform risk analyses to illuminate the uncertainties associated with the issues of interest to the sponsor; many a game participant (and sponsor) has drawn a misleading inference from the idiosyncratic outcome of a single game. Finally, the results of the assessments must be documented, so there can be both peer reviews and a foundation upon which future analyses can be built.

Although figure 1 does not formally specify it, the Code of Best Practice emphasizes that an extensive feedback arrangement is needed to share insights among individuals carrying out these successive processes as game planning progresses. Further, the overall team must be an interdisciplinary one—comprising operations analysts, war-game designers, experimental designers, computer scientists, social scientists, and so on—if it is to address all of the issues of concern to a sponsor. A "good game," then, blends clear problem formulation, technical virtuosity, accurate data, scenario creativity, appropriate decision rules, and credible evaluation procedures. The rest of this article treats some of these points in greater depth.

PROBLEM FORMULATION

There is no shortage of problems amenable to useful analysis by war-gaming techniques. In fact, policy makers are likely to find war gaming the most effective tool for clarifying many issues and sets of issues that can be expected to come to the foreground in the near and middle term. Let us here consider a few problems for which war gaming could be particularly appropriate.

Strategic Visions. An indication of the variety of problems to which gaming might be applied as an analytical tool is the set of three lists of strategic problems assembled recently by former secretary of defense William J. Perry and former assistant secretary of defense Ashton B. Carter, in their book *Preventive Defense.*[2] Their "A list" comprises potential (and possibly preventable) future matters that could threaten the survival, way of life, and position in the world of the United States (such as a resurgent and hostile Russia, uncontrolled proliferation of weapons of mass destruction, or catastrophic terrorism). The "B list" contains direct threats (deterrable through ready forces) to vital American interests (for instance, major theater wars). The "C list" cites problems (like Kosovo, Bosnia, and Somalia) that "indirectly affect U.S. security but do not directly threaten U.S. interests."[3] Many of these issues, particularly those on the "A list," have yet to be explored adequately in war games.

The overall team must be an interdisciplinary one—comprising operations analysts, war-game designers, experimental designers, computer scientists, social scientists, and so on.

Homeland Defense. In its recent report, "Seeking a National Strategy: A Concept for Preserving Security and Promoting Freedom," the congressionally established Hart-Rudman Commission emphasizes the need to enhance what it calls "homeland security" to deal with emerging world threats.[4] A third and final phase of that study will address a variety of associated questions: Are responsibilities, authorities, and accountabilities clear? Do integrating mechanisms exist? What capabilities will be needed? Is the overall capacity sufficient, and if so, will it continue to be?[5] A suitably designed set of war games would be a promising way to illuminate these issues.[6]

Operational Tempo. One of the driving issues in the 1997 Quadrennial Defense Review was the necessity to devise "architectures" and personnel policies to allow U.S. forces to respond to operational demands that were expected to be high enough to put pressure on unit training and maintenance, as well as morale and retention. The "Dynamic Commitment" war game was developed and played to address that issue; it is being revised to serve the same need for the Quadrennial Defense Review of 2001.[7] That game—which is to play a single scenario, drawn randomly from a list of sixty-one "vignettes"—is itself a case in point, showing

that contemporary gaming does not reflect sufficient understanding of risks and uncertainties. The consequences of a particular vignette being played out in a single game, and of drawing conclusions therefrom about levels of demands that can be placed on U.S. equipment and personnel, are worrisome.

Service Transformation. Each of the military services is in the midst of sweeping modernization designed to take advantage of opportunities offered by the information age. Specific initiatives include the Navy's network-centric warfare, the Army's "Future Combat System for Smaller Scale Contingencies," the air expeditionary forces of the Air Force, and "Operational Maneuver from the Sea" of the Marine Corps. War games have contributed to preliminary assessments of each of these concepts singly, but there has been no attempt to game the totality of their effects. Doing so would appear to be a high-priority matter.

How can scenarios be kept (in Albert Einstein's formulation) as simple as necessary—but no simpler?

Joint Vision 2020. The chairman of the Joint Chiefs of Staff, in the recent doctrinal white paper *Joint Vision 2020,* conceives "a joint force capable of full spectrum dominance, persuasive in peace, decisive in war, and preeminent in any form of conflict."[8] The document reaffirms as the prerequisite of full-spectrum dominance four operational concepts—dominant maneuver, precision engagement, focused logistics, and full-dimensional protection—identified in an earlier white paper, *Joint Vision 2010.* These four operational concepts in turn depend on three factors: interoperability (joint force, interagency, and multinational), innovation leading to transformation, and "decision superiority" (to allow commanders to "make better and faster decisions than their opponents"). All of these factors, as well as their relationship to the central operational concepts of *Joint Vision 2020,* are very attractive subjects for gaming.

ORGANIZATION AND CULTURE

In games played by coalition allies prior to Operation DESERT STORM, differences in cultures were sometimes recognized as a major factor. For instance, British analyses reflected a particular appreciation of Iraqi characteristics that profoundly affected the planning and operational concepts of the British forces in the theater.[9]

Cultural differences were again acknowledged as central strategic factors in 1999, during Nato's coercive air campaign to terminate internecine hostilities in Kosovo. The subsequent debate about what actually prompted Slobodan Milosevic's acquiescence to Nato's demands has produced at least one analysis of the cultural and political dynamics of the Serbian leadership.[10]

It has been widely appreciated that war games require a much better theoretical basis than is now available for treating these matters in future conflict situations. One potential source of some necessary insight may emerge from work being done in the Office of Naval Research on "Adaptive Architectures for Command and Control." It examines the command-and-control staffs of various nations for pertinent "cultural artifacts" and their potential influence on decision making. In addition, the war-gaming community would do well to draw on the efforts of sociologists and political scientists, who could analyze the underlying cultural forces at work in such recent operations as Somalia and Kosovo, and who can be consulted in planning games for prospective involvement in foreseeable crises.

SCENARIOS

Today, basic issues in the selection and development of scenarios are being examined. Can a baseline scenario be used for a series (or "cluster") of games? How can scenarios be kept (in Albert Einstein's formulation) as simple as necessary—but no simpler? Can "excursions" into important issues be accommodated, and if so, in what ways?

Clearly, no simple answers to these questions exist, but there is a fundamental principle that game designers today should acknowledge—that no single scenario can adequately illuminate risk and uncertainty. The challenge is to develop an efficient mechanism for finding and exploring regions of "scenario space" where key factors play in significant ways. The Nato Code of Best Practice offers one approach to the problem, a scenario framework that subsumes three major categories—external factors (the political, military, and cultural situation), the capabilities of actors (friendly and adversary forces, noncombatants), and the environment (geography, terrain, and weather).

As an illustration of how such a framework might be used to develop a baseline scenario (and possibly scenario excursions), consider a methodology that enumerates the factors applicable to a given game.[11] For each of those factors, a number of values (specific geographies, particular orders of battle, etc.) can be assigned, each making a scenario more or less challenging in some respect that is significant in terms of a game's objectives. Between the bounding ("easy," "very difficult") values for each factor lie the elements of a potentially interesting baseline scenario; alternative scenarios can be readily produced for sensitivity analyses by selecting different values for particular factors. In effect, this approach generates a very large experimental-design matrix, each cell of which corresponds to a specific scenario. In traditional scientific experimentation, a number of iterations would be run for selected matrix cells in order to achieve statistically meaningful results; statistical uncertainty would be a function of

the number of cells examined and the number of independent trials of each. In war games, of course, a "full factorial experiment" would be impossible; still, it would be prudent to play at least a sampling of variants—a "sparse, fractional factorial experiment."

Closely related to scenarios is consideration of risk and uncertainty. As the Code notes, a useful way to display and characterize areas of uncertainty in a game is to play variations of the scenario. In doing so, however, it is important to take account of, and offset, the effects of learning that occurs in the play of a game. For instance, the sequence of variations should anticipate and minimize the "carry forward" insights obtained in each variation; one way to do this is to make the new problem appear different to the participants but have it contain the same essential stimuli. It will almost certainly not be possible to run enough iterations to bound measures of merit as tightly as a physical scientist would wish; nevertheless, to some extent well designed pre- and postgame analyses can refine those estimates.

Today's state-of-practice technology simply collects stand-alone collaboration tools—however, the state of the art has advanced to the point of integrating those capabilities into "virtual buildings."

A more basic issue is the estimation of risk. Risk analysis as a discipline is well developed in a number of fields, such as the insurance industry and stock brokerages, but in the context of national security there is little agreement even about the definition of risk itself. This is becoming a pressing issue, because the congressional mandate of the Quadrennial Defense Review specifically requires "a comprehensive discussion of [the] national defense strategy of the United States and the force structure best suited to implement that strategy at a low to moderate level of risk."[12] To meet this requirement the national security community will need to agree on definitions of risk, definitions that are amenable to evaluation in future war games.

MEASURES OF MERIT

For decades gamers have employed the familiar operations-analysis device of "measures of effectiveness" to structure game outcomes and relate them to sponsors' concerns. In recent years, however, the concept of measures of effectiveness has been broadened, resulting in the idea of "measures of merit."[13] As discussed in the Nato document, this conception not only embraces the conventional measures of effectiveness but allows a linked hierarchy of increasingly specific metrics to be considered as well. For example, the evaluation measures of a game might employ measures each of which "nests" within the next to provide both broad and detailed attention as appropriate. An example follows:

- *Measures of policy effectiveness,* assessing the extent to which the participants in an operation are able to achieve national or international security objectives;
- *Measures of force effectiveness,* examining the purely military effectiveness of a force in terms of its primary task (such as the time required to halt an attack);
- *Measures of mission effectiveness,* appraising the ability of the military force to perform key subordinate or subsidiary missions;
- *Measures of functional performance,* evaluating the success of a particular weapon system or command-and-control organization in important tasks, such as target engagement;
- *Dimensional parameters,* the properties or characteristics (such as bandwidth and resistance to jamming) of a specific system, such as a communications network.

Game designers might usefully devise measures for each level of this hierarchy, and analysts might explore their relationships during the course of the game. At the lower end of the hierarchy, extensive analyses have been performed for traditional warfare; that literature is being expanded upon to embrace information superiority.[14] It would be necessary, however, to formulate meaningful measures of merit for the top of the hierarchy. In one promising effort in this direction, economic measures were used to reflect the societal impact of military operations.[15] Participants were asked to estimate the effect that postulated crises might have on such indicators as the Dow Jones Industrial Average, the price of a barrel of crude oil, or the exchange rate between the dollar and the deutsche mark.

As the Nato Code of Best Practice concludes, games are not suitable for every analytical question. Indeed, no single assessment technique is likely to be sufficient (see table 1). Since games are increasingly likely to address such concepts as information superiority and information dominance, assessment tools must account for both friendly and adversary information processes. In addition, discipline is necessary; formal experimental-design matrices may be advisable, or multiple iterations of increasingly fine-grained analytical routines may have to be done (for instance, in successive attempts before a game to identify fruitful aspects of the scenario environment, clarify assumptions, assign values for key parameters, and model details).

Newly developed sophisticated collaboration tools may revolutionize war games by allowing geographically dispersed individuals to participate fully in deliberations and decisions. Today's state-of-practice technology simply

collects stand-alone collaboration tools—like video teleconferencing, shared whiteboards, and Internet chat rooms. However, the state of the art has advanced to the point of integrating those capabilities into "virtual buildings" in which participants interact in real time. Efforts are under way to improve "scalability" (usefulness for various numbers of players and complexities of scenario) and to deal with security issues regarding the transmission of game data.

TABLE 1
SPECTRUM OF CANDIDATE ASSESSMENT TECHNIQUES

Techniques	Typical Application	Systems	People	Ops/ Mission	Resources	Lead Time		Credibility
						Create	Use	
Analysis	Closed Form; Statistical	Analytical	Assumed or Simulated	Simulated	Relatively Modest	Weeks to Months	Weeks to Months	Fair to Moderate
Constructive	Force on Force Models; Communication Systems	Simulated	Assumed or Simulated	Simulated	Moderate	Months to Years	Weeks to Months	Moderate
Virtual	War games; Human in the Loop	Simulated	Real	Simulated	High	Months to Years	Weeks to Months	Moderate to High
Live	CPX* FTX**	Real	Real	Real or Simulated	Very High	Years	Weeks to Months	High
Actual Ops	After Action Reports; Lessons Learned	Real	Real	Real	Extremely High	N/A	N/A	Very High

*CPX-Command Post Exercise
**FTX-Field Training Exercise

One of the major advantages that these emerging collaborative gaming tools offer is the possibility that principals—commanders, heads of agencies, senior executives—will be able to participate personally. The demands on the time of such individuals normally make it difficult for them to get involved in war games, especially if travel is involved; typically they must delegate such matters to subordinates. Distributed, collaborative war-gaming technologies will make it possible for actual decision makers to play, increasing both the fidelity of the games and the real value of the entire activity by educating the principals directly about the intricacies and nuances of the problems being considered.

NOTES

1. North Atlantic Treaty Organization, *Code of Best Practice (COBP) on the Assessment of C2,* RTO Technical Report 9, AC/323(SAS)TP/4 (Hull, Que.: Communication Group, Inc., March 1999).
2. Ashton B. Carter and William J. Perry, *Preventive Defense: A New Security Strategy for America* (Washington, D.C.: Brookings Institution Press, 1999).
3. Ibid., p. 11.

4. Gary Hart and Warren Rudman, *Seeking a National Strategy: A Concept for Preserving Security and Promoting Freedom,* Phase II Report (Washington, D.C.: Commission on National Security/21st Century, 2000), available on the World Wide Web: http://www.nssg.gov (15 April 2000).
5. "Hart-Rudman Panel to Focus on Homeland Security in Final Phase," *Inside the Pentagon,* 25 May 2000, pp. 1, 19–20.
6. The Hart-Rudman Commission's final report, "Road Map for National Security: Imperative for Change," was released on 15 February 2001 (available on the World Wide Web: http://www.nssg.gov/PhaseIIIFR.pdf). It concluded that protecting the homeland against twenty-first-century threats "should be the primary national security mission of the United States." The report proposes answers for the institutional and organizational issues cited above, but it does not obviate the need for war games to assess critically the viability and desirability of the proposed solutions.
7. "Services Meet Snags in Readying Dynamic Commitment War Games," *Inside the Pentagon,* 27 April 2000, pp. 1, 17–8. In March 2001 the Pentagon announced that Dynamic Commitment had been postponed indefinitely.
8. Henry H. Shelton [Gen., USA], *Joint Vision 2020* (Washington, D.C.: Joint Chiefs of Staff; U.S. Govt. Print. Off., June 2000).
9. Discussions with Jim Moffat of the Defence Evaluation and Research Agency, U.K. Ministry of Defence, 4 April 2000.
10. Barry R. Posen, "The War for Kosovo, Serbia's Political-Military Strategy," *International Security,* Spring 2000, pp. 39–84.
11. See the author's "Developing Scenarios to Support C3I Analysis," in *Analytic Approaches to the Study of Future Conflict,* ed. Alexander Woodcock and David Davis (Clementsport, N.S.: Canadian Peacekeeping Press of the Lester B. Pearson Canadian International Peacekeeping Training Centre, 1996), pp. 291–300.
12. Public Law 106-65-Oct. 5, 1999, *Statutes at Large* 113: 716.
13. For a discussion, see North Atlantic Treaty Organization, *Code of Best Practice.*
14. IS Metrics Workshops sponsored by David Alberts (Office of the Assistant Secretary of Defense for Command, Control, Communications, and Intelligence [OASD (C3I)]) and by John Garstka (Office of the Director for Command, Control, Communications, and Computers [J6], Joint Staff).
15. Jeffrey Sands, *The Critical Link: Financial Implications of Threats to National Security* (a report on the Economic Security Exercise, cosponsored by the U.S. Naval War College and Cantor Fitzgerald, Newport, R.I., December 1997).