

2

AD-A255 880



A RAND NOTE

The Base of Sand Problem: A White Paper
on the State of Military Combat Modeling

Paul K. Davis, Donald Blumenthal

DTIC
ELECTE
OCT 6 1992
S C

Approved for public release
Distribution Unlimited

92 11 034

92-26446



RAND

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST QUALITY PRACTICABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

A RAND NOTE

N-3148-OSD/DARPA

**The Base of Sand Problem: A White Paper
on the State of Military Combat Modeling**

Paul K. Davis, Donald Blumenthal

**Prepared for the
Office of the Secretary of Defense
Defense Advanced Research Projects Agency**

APPROVAL FOR

SEARCHED	INDEXED
SERIALIZED	FILED

A-1

RAND

PREFACE

This Note in the form of a White Paper improves upon a draft circulated in early 1990 to senior Department of Defense (DoD) officials and military officers. The original draft came about because the authors had separately concluded that the DoD's approach in developing and using combat models, including simulations and war games, is fatally flawed—so flawed that it cannot be corrected with anything less than structural changes in management and concept.

This white paper was originally written in March, 1990. We updated it in mid-February, 1991, but did not attempt to reflect either changes in the world situation or new modeling and simulation initiatives in DARPA and OSD.

The draft built upon discussions at a small ad hoc workshop convened at RAND for several days in December, 1989. The working group consisted of senior scientists and analysts from RAND, the Lawrence Livermore National Laboratory, and the Jet Propulsion Laboratory. All present had conceived, built, and used military models, including sophisticated war gaming systems. Their experience extended from developing and validating physics-level algorithms to analyzing alternative theater and global military strategies and arms control agreements. Further, participants had worked at the technological frontiers of simulation, including knowledge-based simulation. They had had many successes and had learned from partial failures. The participants, then, were not gadflies, professional or organizational critics of the DoD, academic purists, or anti-model Luddites. Rather, they were insiders of the national security model-development and analysis communities. They were attempting to sensitize the DoD to a serious national problem and the need for drastic remedies. The draft in 1990 appears to have contributed to that goal, based on the numerous unsolicited responses we received. We hope this final version will be useful to a broader community.

This Note was sponsored by the Defense Advanced Research Projects Agency (DARPA). It was developed in RAND's National Defense Research Institute (NDRI), a federally funded research and development center sponsored by the Office of the Secretary of Defense and the Joint Chiefs of Staff. The original workshop and draft were accomplished using NDRI's research support funds. Comments are welcome and should be addressed to Dr. Paul K. Davis, RAND's Corporate Research Manager for Defense Planning and Analysis. The Internet address for electronic mail is Paul_Davis@rand.org. Mr. Blumenthal is a consultant to the Lawrence Livermore National Laboratory.

SUMMARY

The DoD is becoming critically dependent on combat models (including simulations and war games)—even more dependent than in the past. There is considerable activity to improve model interoperability and capabilities for distributed war gaming. In contrast to this interest in model-related *technology*, there has been far too little interest in the *substance* of the models and the validity of the lessons learned from using them. In our view, the DoD does not appreciate that in many cases the models are built on a base of sand. Nor does it appreciate that while replacing the sand with a more nearly solid foundation is feasible, it will be extremely difficult in scientific, intellectual, and managerial terms.

Fig. S.1 describes the base of sand as chaos in the modeling community. Many models and tools are simply inadequate. Further, they are seldom verified or evaluated well, there is confusion about what the models assume and do, and there have been many wasted efforts and lost opportunities. The principal contributors to this chaos, which are themselves problems, are (moving clockwise from the right in Fig. S.1): (a) inadequate theories, methods, standards, and practices for modeling and evaluation; (b) related computer hardware and software problems, such as achieving interoperability and software quality; (c) dissonance and lack of discussion across communities and organizations (e.g., operations planners vs defense planners); and, importantly, (d) the lack of a vigorous military science.

Upon looking into each of these contributors in more detail, we conclude that an *overarching problem is a variant of "No one is in charge."* There currently exists no national office with the responsibility of encouraging, nurturing, and sponsoring activities necessary to relieve the chaos described in Fig. S.1.

In considering remedies, it is important to recognize the need to invigorate the study of military science—shifting the balance between art and science farther to the science end of the spectrum, but doing so with full recognition that military science is a social science, beset with uncertainties and variables that cannot accurately be measured or anticipated. Such a science requires theories and associated models, coupled with activities to test and inform them (Fig. S.2). We recommend explicitly distinguishing between two overlapping categories of models: *research models*, which collectively embody and communicate our knowledge (including alternative theories and both objective and subjective data), and *application models*, which can be seen as tools allowing us to solve particular problems or inform decisions. By and large, policymakers and analysts have conceived models as mere ad hoc tools, to the detriment of continuity and coherence of military science. At the same time,

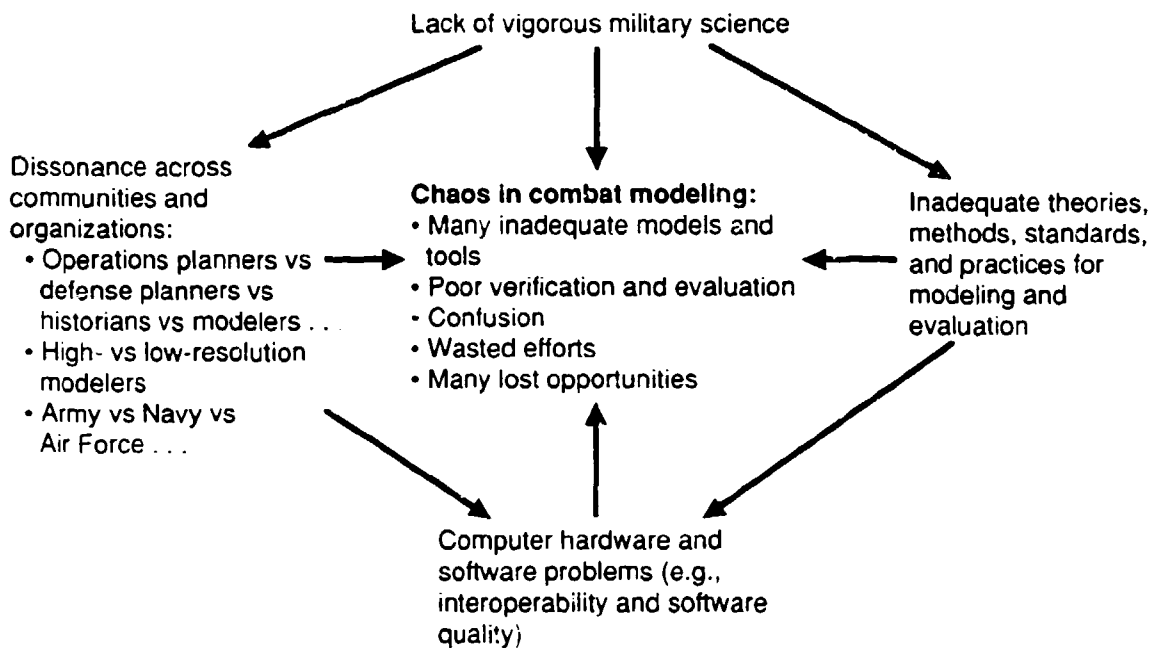


Fig. S.1—Contributors to the chaos in current combat modeling

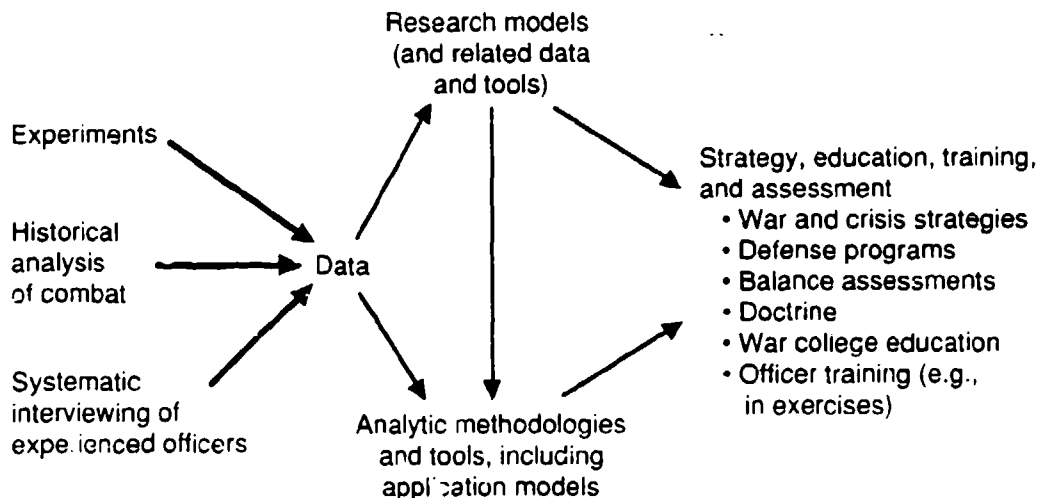


Fig. S.2—Elements of a comprehensive approach

those building more comprehensive models have typically not appreciated their limitations for analysis (they are often too complex for policy analysis), nor designed them as research models. The DoD needs to understand the different roles of research and application models and should sponsor and nurture both within a framework of an evolving military science.

For both research and applications, combat models should be viewed less as answer machines than as frameworks for summarizing and communicating objective and subjective knowledge (including knowledge of uncertainties), and as mechanisms for *exploration*. This view, which is especially important in designing complex research models, establishes stringent requirements for model transparency, comprehensibility, and flexibility. Increasingly, it is also important that models be designed so that they or their modules can be directly compared with, used in, or used with other models. This is crucial for scientific reproducibility and peer review, for efforts to calibrate aggregated models using higher resolution models, for analysis of empirical data, and for distributed war gaming. The theories and methods for designing models with these purposes in mind are not currently well understood.

In this paper, we recommend that the Secretary of Defense establish an Office of Military Science (OMS) to plan and administer the process of creating the national environment necessary for a vigorous military science. The OMS would not conduct research itself, but would instead encourage, nurture, and to some extent sponsor it—although relying primarily on the Services and other agencies for the vast majority of research and analysis. The OMS would be concerned primarily, but not exclusively, with research and research models rather than direct applications.

The OMS would (1) sponsor conferences, journals, and the development of textbooks; (2) sponsor development and iteration of experimental technical standards and data dictionaries to make model interoperability and model comparisons feasible and attractive; (3) encourage, coordinate, and sometimes sponsor historical and other empirical research to inform model building and evaluation; (4) disseminate optional standard model modules and data bases (to be thought of as baselines rather than something “blessed”); (5) sponsor cross-cutting military-science research that would not otherwise be accomplished, including model comparisons and “countermodeling” efforts; (6) support development of methods for the use of models and decision aids in distributed and conventional war gaming; and (7) facilitate the exchange of both research and application models and data bases.

We do not emphasize a czar-like OMS role for verification, validation, or accreditation, because verification and validation (V&V) are difficult to define and must often be accomplished in the context of a particular study. Further, because of the uncertainties

inherent in military science, even the term "validation" is arguably inappropriate because it may connote a sense of precision and certainty. In our view, the offices sponsoring particular studies must be the ones to judge the suitability of models. Attempts to centralize accreditation would be contrary to the scientific approach and would lead to bureaucratic mischief. The OMS would, however, sponsor and disseminate methods, tools, and data for V&V (or, better, model "evaluation"). As mentioned above, it would also encourage model comparisons and tests and would publish results, including disagreements of interpretation. Further, it could help agencies evaluate the suitability of particular models for particular applications, and in that important sense help provide a kind of limited "accreditation."

As a general proposition, the OMS would be more concerned with improving the quality and efficiency of the military-science marketplace, in part by *encouraging* parallel research, than with minimizing cost or "stamping out redundancy." Our image of an OMS, then, is very different from that of a tough cost-cutting central manager. What is needed is something combining the features of DARPA-style research with an interest in helping agencies work out current problems. DARPA's reentry-physics program several decades ago was in some respects an excellent model. It established a base of knowledge that made possible the subsequent military and civilian space programs.

In our view, then, major organizational changes and a long-term perspective are essential if DoD is to replace the base of sand with something more substantial. The present approach to the development and use of models is fatally flawed.¹ However, because of the widespread interest and growing importance of models, including distributed war gaming, the opportunity now exists to greatly improve the situation.

¹This white paper was originally written in March, 1990. We updated it in mid-February, 1991, but did not attempt to reflect the new modeling and simulation initiatives in DARPA and OSD that have been partly motivated by congressional interest. The initiatives had not yet "taken hold" as of the time we completed our work; nor did they go far enough. Further, we thought it best to complete our independent assessment. However, if the initiatives are sustained, they will address some of the issues we describe. The relevant DoD offices had the earlier version of our work.

ACKNOWLEDGMENTS

The authors appreciate workshop discussions with Dr. Richard Hillestad and Dr. John Friel of RAND, Dr. Ralph Toms of Livermore National Laboratory, and Dr. Joseph Feary and Dr. Hugh Henry of the Jet Propulsion Laboratory. RAND colleague Bruce Goeller provided a detailed and thoughtful review. Many other people also provided comments and suggestions. Finally, we wish to acknowledge the many colleagues in the analytic community with whom we have talked about the issues in this Note over the last few years—colleagues in FFRDCs, national laboratories, civilian and military agencies, and commercial companies. We hope that our critical assessments will be recognized as being directed at processes and mindsets, not individuals or organizations, and that none of the criticisms we provide will in any way diminish the many and substantial advances in modeling that have occurred in the last decade. There are no villains in the story we tell here.

CONTENTS

PREFACE	iii
SUMMARY	v
ACKNOWLEDGMENTS	ix
FIGURES	xiii
Section	
I. INTRODUCTION TO THE PROBLEM	1
Preliminary Comments on Terminology	1
The Increasingly Critical Role of Models	3
The Problem: "A Base of Sand"	6
II. A SYSTEM-LEVEL DIAGNOSIS	12
Overview	12
The Lack of a Vigorous Military Science	12
Dissonance Across Communities and Organizations	15
Inadequate Theories, Methods, Standards, and Procedures	16
Computer Hardware and Software Problems	17
The Issue of Military Science	19
Rethinking the Role of Models	22
III. A VISION OF HOW TO IMPROVE THE SITUATION	26
Issues of Scope, Resolution, and Perspective	26
A Comprehensive Approach	29
Model Attributes Needed	30
The Environment	33
Quality People	34
IV. A PLAN OF ACTION	36
Organizational Issues	36
Enlisting the FFRDCs and National Laboratories	39
Establishing Links with the Warrior Preparation Center and Joint Warfare Center	39
An Agenda for Immediate and Near-Term Actions	39
Conclusion	41
BIBLIOGRAPHY	43

FIGURES

S.1	Contributors to the chaos in current combat modeling	vi
S.2	Elements of a comprehensive approach	vi
1.1	Examples of U.S. combat models	2
2.1	Factors influencing chaos in combat modeling	13
2.2	Factors influencing the lack of military science	13
2.3	Factors influencing dissonance	16
2.4	Factors influencing inadequacy of theories, methods, standards, and practices for modeling and testing	17
2.5	Factors influencing computer hardware and software problems	19
3.1	Different levels of resolution and perspective	26
3.2	An illustrative, very simple model (of concentration and counterconcentration)	28
3.3	Components of a comprehensive approach	29

I. INTRODUCTION TO THE PROBLEM

PRELIMINARY COMMENTS ON TERMINOLOGY

This Note refers frequently to "models," "simulations," and "war games."

Unfortunately, these terms have multiple and contradictory meanings in common English. For example, "models" include the ship in a bottle, theories of expansion of the universe, and political scientists' conceptual frameworks. "Simulations" include the playing through of procedures by a team of workers, various types of models and computer programs, and electric-train networks enjoyed by hobbyists. There are some relatively general definitions, but none of them can encompass all the usages, because the usages are contradictory. In this paper, therefore, we adopt a particular set of definitions as follows (not because our definitions are better than others, but because they will simplify discussion here):¹

- a. A *model* is a mathematical or otherwise logically rigorous representation of a system or a system's behavior. It may or may not be computerized. It may or may not be structured as a game. It may or may not attempt to represent the internal functioning of the real system. It may be abstract only, or it may be implemented as a computer program, a nomogram, pencil-and-paper procedures, or in a variety of other ways. For brevity, we will also refer to the implementation (e.g., a computer program) as a model, even though the distinction (e.g., program vs model) is sometimes important.
- b. A *simulation* is a special kind of model that represents at least some key internal elements of a system and describes how those elements interact over time. Most combat simulations are implemented as computer programs. The principal exceptions are manual war games, discussed below. Computer simulations may be *closed*, in which case the user "pushes the button" and the computer generates a complete simulation. They may instead be optionally *interruptible*, in which case the user may intervene during the simulation and change assumptions in midstream. *Interactive* simulations may be optionally interruptible or may instead demand that users provide information and decisions during the simulation. Some simulations (closed, interruptible, and interactive) are structured as games—i.e., there may be explicit model entities representing opponents and allies.

¹See Hughes (1984), Rothenberg (1989), and Anderson, Cushman, Gropman, and Roske (1989) for careful discussions and model taxonomies. Other definitions have been proposed by "the Gorman Panel" in its work to establish a draft charter for a DoD oversight office for modeling and simulation.

- c. *War games* are either manual games or computer simulations with human players making some or all of the key decisions. War games are themselves models in that they attempt to represent a system (e.g., the nations participating in a war). However, they also require the use of specialized submodels; modern war games typically employ interruptible or highly interactive simulations, with the opposed players making periodic decisions about how to deploy and employ forces. These decisions are entered into the computer and the simulation is resumed. A few war gaming models (e.g., RSAS and CONMOD) can be used interactively in games and can also be used without player intervention, as closed simulations, by substituting decision models.

Fig. 1.1 shows how models, simulations, and games relate to each other and gives examples, drawing on U.S. models. It follows that in this paper when we use the term "model," we include simulations and war games. We deal in this specific paper only with combat models for general-purpose forces fighting at the battalion through theater levels. We do not, for example, discuss engineering-level models, human-factors models, or logistics models.

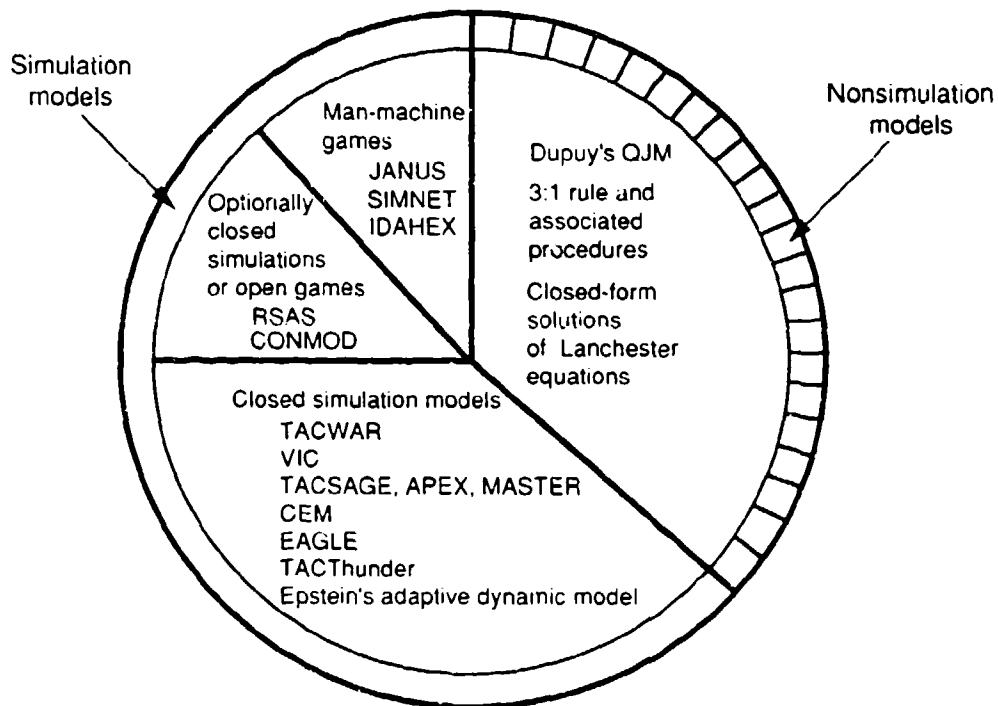


Fig. 1.1—Examples of U.S. combat models

THE INCREASINGLY CRITICAL ROLE OF MODELS

Factors at Work

We assume that readers already recognize generally the importance of models and that there are long-standing problems of verification and validation. They may be less aware, however, that DoD dependence on combat models, and particularly on simulations, is increasing rapidly. There are several factors at work:

- Scenarios, weapon systems, operational concepts, and forces are changing *qualitatively*.
- There is no experience base to guide much of the related planning, although the war with Iraq is mitigating this problem in some respects.
- Computer simulations, including war games, will be replacing many field exercises.
- Distributed war gaming in particular will be a major, and perhaps the principal, mechanism for joint and combined theater-level coordination and training of commanders.
- Computerized decision aids are becoming increasingly critical in command-control processes.

If these represent demand factors for modeling, then supply factors matter also; the supply changes taking place in computer science, information systems, simulation, and distributed war gaming are revolutionary, not incremental.

Scenarios, Forces, and Concepts

To elaborate on these themes, consider first that the changes taking place in Europe and the Soviet Union are transforming the strategic landscape and that the DoD's future planning cases will be highly diverse with respect to location, allies and antagonists, scale of combat, nature of combat (e.g., maneuver vs positional warfare), and many other factors. The old standby scenarios of Central Region defense at the intra-German border and Southwest Asian defense against the Soviet invasion of Iran are obsolete. They will be replaced by a broad range of scenarios that will include defense of the unified Germany, defense of post-war Kuwait and Saudi Arabia, contingency actions, such as assisting Poland in crisis (detering Soviet reentry into Eastern Europe), and other specific and generic

contingencies worldwide. Many of these will bear little or no relationship to big-war contingencies and will pose markedly different requirements.²

Second, revolutionary weapon systems will be entering the force throughout the 1990s, and operational concepts will change as a result. The war with Iraq (the *Desert Storm* operation) is, in a sense, a mere preview. Maneuvering of fires will be a basic element of operations; tactical aircraft will have long-range precision-guided standoff weapons; reconnaissance, intelligence, surveillance, targeting, and acquisition (RSTA) will be critical processes; low-observable aircraft will have unique and critical capabilities and will in time encounter reactive defenses; air forces will be an even more integral part of planning for ground-force operations; and, at the item level, the competition will continue between offense and defense at the level of tanks and anti-tank guided munitions (ATGMs), but in a complex environment that may include anti-personnel lasers, area munitions, highly lethal indirect fire, or chemical weapons.

High-tempo operational maneuver will be increasingly emphasized, even for many Central European scenarios, because negotiated and unilateral force reductions will reduce force-to-space ratios and place a higher premium on sound early decisions for concentration and counterconcentration. Also, long-distance movements will be necessary from peacetime positions to defense lines along the eastern border of Germany or, in "deter-reentry scenarios," in Eastern Europe. And, lastly, the lethality of modern weapons will increasingly force antagonists to avoid concentration except during critical periods.

The Increasing Need for Models

The war with Iraq will probably increase greatly the number of serving officers with major-combat experience, especially if a ground offensive proves necessary. It is already yielding a wealth of data on mobilization, deployment, and air operations. Even so, the new experience base will be of limited value when contemplating the next-generation battlefield. Soviet military theorists, who have not generally exaggerated such matters over the decades, consider that a revolution in military affairs is taking place. Many thoughtful Western observers are making comparisons between the 1990s and the 1930s, which was a period of innovation, experimentation, and change—in weapon systems, doctrine, and higher level concepts of operations. In the 1930s, there were wars to guide much of the assessment, and the *Desert Storm* operation will be similarly invaluable. However, the significance of the empirical data will have to be interpreted by models disentangling the many variables

²For discussion of future planning scenarios, see Davis and Howe (1990a), Kugler (forthcoming), and Winnefeld and Shlapak (1990).

affecting battle outcomes.³ Further, many assessments will involve combat phenomena not arising in the war with Iraq (e.g., operations when we do *not* have absolute air supremacy), and these will be driven by models and very limited field tests.

This increased dependence on models (including, remember, war games) has already been occurring with the emergence of SIMNET at the military-technical level and as distributed war gaming has become a reality under the leadership of NATO's Supreme Allied Commander, General Galvin. The technological opportunities now exist to use these tools far better than before. The results could be in some ways far superior to those achieved with traditional exercises, which are procedural, scripted, narrow in scope, and complicated by masses of people and activities that obscure key issues. At the same time, there is really no choice but to use models more extensively, because economic and political limitations will greatly constrain the number, scope, and realism of field exercises in the future.

Decision aids are also becoming increasingly critical. Already, commanders and senior staff depend heavily on computerized information displays from the weapon-system level up. Most of these displays reflect implicit models, since it is models that dictate what information one wants to see. At the weapon-system level this is often well understood, but decision aids going beyond relatively straightforward presentation of data and oriented specifically to higher-level field commanders are still relatively primitive (e.g., aids to help evaluate different courses of action or to find courses of action to achieve specified objectives). The same is true of models for assessing alternative defense programs.

Technological Opportunities

We are now well into the information era, and increasingly it is recognized that information dominance is becoming critical in warfare. Fortunately for the United States, we have advantages in the most relevant technologies, which include not only communications and computing power, but also model-related technology, such as new programming languages; analyst workstations; relational data bases; man-machine simulation systems, such as SIMNET; and knowledge-based modeling concepts. Unfortunately, however, there is a problem that has already become a limiting factor in what can be accomplished, one that is not yet widely recognized. We call this the base of sand.

³To illustrate how critical the use of combat models is in analyzing empirical data, consider that battle outcomes have historically borne no relationship to the raw force ratio. By contrast, when the outcome data is passed through models sensitive to situational factors such as terrain, preparations, asymmetries in fighting effectiveness due to better organization and training, and so forth, one finds that the data actually makes sense and that what matters is a ratio of *effective* forces. Unfortunately, the values of some of the key variables may not be known in advance. As a result, the models are sometimes more useful for after-the-fact description than for reliable prediction.

THE PROBLEM: "A BASE OF SAND"

Too Many Gadgets and Too Little Thinking

As background here, the increasing role of models has been observed for several years (e.g., in symposia held on the subject between 1987 and 1989).⁴ These symposia typically dealt, however, with advanced simulation *technology* and the technology of distributed war gaming. We and other participants emerged from these activities concluding that

- The DoD and the vast majority of the model community are relatively too enthusiastic about the advances in computer performance, communications, and human interfaces, and much too little interested in the *substance* of the models and the validity of the lessons that will be learned from them.
- The DoD does not seem to appreciate that, as discussed below, *the models are in many cases built on a base of sand*. Further, the compensating factors that have long mitigated this problem will not be adequate in the emerging environment of large-scale distributed war gaming and simulation,⁵ especially when distributed war gaming goes beyond the stage of merely distributing terminals, to stages in which the various participants wish to use and coordinate their separate models. At this stage, technical complexity will skyrocket (Bankes, forthcoming).
- The DoD has not understood that correcting the situation is a Herculean task that will require high-level priority, coordination, time, and money. *Replacing the sand with a more nearly solid foundation is feasible, but it will be extremely hard in scientific, intellectual, and managerial terms.*

Examples of How the Base Is Sandy

All those involved for some time in military analysis have examples of allegedly scandalous problems in models or the use of models.⁶ We mention a few of our own examples here, with only a minimal effort at structuring. We will then provide a more organized diagnosis.⁷

⁴See Bankes (forthcoming), Defense Science Board (1988), MITRE (1989) and MORS (1988).

⁵The bedrock tenet of the best users of models has been that results depend ultimately on the quality of the analysts and the analysts project rather than the model alone. Small teams can be aware of model limitations and can work around them. They can also impose quality control on studies. By contrast, as models are increasingly used as "black boxes," by people who had nothing to do with developing them and have little knowledge of what is inside them, the potential for serious errors increases enormously.

⁶See, for example, Comptroller General (1980), Stockfish (1975), and Brewer and Shubik (1979). The GAO report includes a more extensive bibliography to the pre-1980 critical literature. See also the exchange between T.N. Dupuy and W. Hollis in the August and October 1987 issues of *Armed Forces Journal*, although some of Dupuy's criticisms of DoD models were incorrect and some of Hollis' rejoinders too sanguine.

⁷Here we may be accused of polemical assertion: "Yes," it may be said, "there are problems, but there are also examples of people working effectively to alleviate them. Why not mention the good news?" Some commenting

Attitudes

- *Minimal Empiricism.* Little money is spent collecting empirical information to inform higher-level models. Much is spent in developing orders of battle, but little on improving the underlying assumptions about combat or other operational processes. In particular, there is much too little *systematic* effort to collect, structure, or exploit historical data.⁸
- *Parochialism and Ignorance.* There is little discussion and analysis across levels of resolution, across different perspectives of conflict, or even across models allegedly describing much the same thing.⁹ Instead, there are distinct ingrown communities, each with its own biases.¹⁰
- *Dubious Acceptance Criteria.* The criteria for model acceptance by military organizations often have little to do with empirical or historical information. Instead, they have to do with not disrupting operations (e.g., the new model should get the same answers as the old one, but faster), representing the organization's particular interests (e.g., the high potential effectiveness of a particular weapon or force), agreeing with the subjective impressions of the most senior relevant officer, and being compatible with existing computer data bases. In such an environment, the search for "truth" may not even be considered relevant.¹¹

on our first draft even mentioned kindly some of our own work, as well as their own. We agree that there are many successes for the analytic community to be proud of, but we believe our overview depiction here is not overstated.

⁸An exception here over the years has been the work of Trevor Dupuy, who with small and sporadic funding over the years has been the leader in using historical information to inform highly aggregated combat modeling. The Army's Concepts Analysis Agency has also supported useful historical work and continues to do so (e.g., McQuie, 1988, and Fain, et al., 1988; see especially Helmbold, 1990, for abstracts and a bibliography). However, these have been largely independent efforts without much guidance from or input to simulation modelers. There is usually a mismatch between the form of the historical work's output and the information needed by simulation modelers and analysts. For example, Dupuy's and McQuie's work organizes historical information in a way most suitable to *static* models. Nonetheless, it has been exceptionally valuable. For other all-too-rare examples of valuable empirical research, see Hughes (1986), Rowland (1986), and Molnar and Colyer (1988). These deal, respectively, with naval tactics, the effects of suppression, and the effects of interdiction.

⁹The principal exceptions occur *within* organizations. For example, both U.S. Air Force Studies and Analysis and the FRG's Industrielagen-Betriebsgesellschaft GmbH (IABG) have used hierarchies of models with some success, and within RAND and the Institute for Defense Analyses there are some recent examples of work using together models of greatly different character and resolution (e.g., JANUS and both corps- and theater-level models). The Army's Model Improvement Program has attempted such work across Army agencies. Doing such things well, however, is difficult, time-consuming, expensive, and socially complex. Further, the *theory* for doing so is lacking.

¹⁰By contrast, in the traditional scientific and engineering disciplines, it is recognized that one must be able to move from one level of resolution to another. For example, statistical mechanics is used to explain and enhance the laws of thermodynamics, but the laws and measurements of thermodynamics are the peg points that the more detailed theories must agree with to be valid.

¹¹By contrast with that of the U.S., the Soviet military regards it as a sacred duty to pay attention to historical information—a failure to do so may mean the lives of many soldiers in a future war.

Content of Models

- *Phenomena Omitted or Buried.* Typically, ground-combat simulations focus on complex calculations of attrition while treating command-control processes, tactics, and strategy in terms of stereotypes embedded in the data bases. This ignores the evidence of history that such matters (and other "soft factors") are first-order determinants of both deterrence and war outcomes, and should therefore be highlighted.¹²
- *Teaching Wrong Lessons.* Theater-level war games used in major exercises are often built around overaggregated "hexes," which characterize all of a region by certain average features of terrain for that region. As a result, officers "learn" that certain regions are impenetrable when they are not. Similarly, in any number of theater-level war games, officers have "learned" that offensive operations fail without a big numerical advantage (because the models do not give adequate credit for tactical surprise, high-tempo maneuver, flanking operations, and counterattacks during the period of time in which the original attacker's forces are extended and exhausted but not yet aligned effectively for defense).
- *Mirror-Imaging the Opponent.* Theater-level war games used in major exercises typically limit the attack speed of Soviet units to figures acceptable in the Western military community based on its experience, despite more ambitious Soviet norms and a different Soviet philosophy of operations.¹³
- *Inconsistent Assumptions.* "Data" regarding the force-generation capabilities of Soviet and NATO forces are developed with different methodologies and assumptions, and are clearly in conflict with one another. In our view, NATO exaggerated this aspect of the Soviet threat for years by overestimating how quickly Soviet reserves could be effectively employed in assault operations. Interestingly, a significant factor in this was the failure of models to distinguish between effectiveness for offensive vs defensive operations as a function of training time (see Davis, 1988b).
- *Models Implicitly Tied to Obsolete Settings.* Current linear-defense models developed for the old Central Region are ill-suited to the study of large-scale

¹²For discussion of how force employment (operational strategy) has dominated battle, see Davis and Howe (1990b). See also discussions by retired Air Force generals Hosmer and Goodson in MORS (forthcoming). Hosmer's paper is summarized briefly in *Phalanx*, 21, No. 3, September, 1988.

¹³See Simpkin (1984) for an analysis of Soviet operational concepts by a distinguished British general officer specializing in armored warfare. While dismissing many of the intelligence community's claims about how different Soviet maneuver concepts are from Western concepts, he also concludes that we should take seriously the high attack rates emphasized in Soviet doctrine. To our knowledge, such matters are seldom analyzed in simulations. Instead, they are buried in "data."

maneuver combat in the new Central Europe or elsewhere. The rules of thumb underlying the aggregated models are invalid, and the current detailed models are not typically useful for policy analysis. Many of the detailed models also assume stereotyped tactics, limiting their value for maneuver studies.

The Modeling and Analysis Process

- *Failure to Converge.* Studies often reach contradictory conclusions without addressing those contradictions and without the sponsors insisting that they do so—especially if the sponsors are in different organizations. Although we (the authors) like to believe that we, our organizations, and numerous colleagues in the analytic community do somewhat better in this regard, we believe nonetheless that the search for and convergence on truth is not a strong, much less dominating, ethic in the analytic community or most of the sponsoring offices.¹⁴
- *Pro-Detail Bias.* Most funding organizations favor detail in models, whether or not the detail is justified for the studies anticipated. For example, weapon-on-weapon calculations are often preferred over force-on-force calculations. The effect is to proliferate uncertain parameters. The parameter values, although sometimes systematically generated from more detailed models that are organizationally “accepted,” often produce results that are demonstrably incredible (e.g., predicted attrition rates far higher than has ever been observed historically). Thus, the pro-detail bias complicates and obfuscates, but does not necessarily improve analysis. It *might*, if there were sufficient effort to establish a scientifically valid base of parameter values under a distribution of battlefield conditions and stochastic factors, but that is not what typically occurs.¹⁵
- *The Myopia of Aggregated Analysis.* At the other extreme, methods and models using force scores (e.g., WEI/WUVs, ADEs, EDs, or DEFs) are sometimes applied precisely where they are most inappropriate—in making cost-effectiveness comparisons of alternative divisional structures or total-force structures. Light infantry divisions always fare poorly in such work—even though in more detailed analysis accounting for realities of terrain, they can appear highly cost-effective for

¹⁴One observer not in an FFRDC has commented: “Many individuals are engaged in combat modeling, but little cross-fertilization occurs. Peer review, independent verification and refereed publication do not exist. In general, the professional environment is one of beauty contests and political competition, closely held knowledge, and a fear of exposition. Because all simulation systems in use contain serious flaws, and because of the fierce competition for turf, money, and the ear of senior officers and officials, to expose one’s work for professional examination and critique is to commit suicide.”

¹⁵For an unflattering but perceptive parable on the Services’ approach to models and analysis, see Builder (1989, pp. 107ff).

certain missions, thereby suggesting a mixed-force approach. One might reasonably ask why the score-based methodologies have not long ago been extended to be situationally dependent.¹⁶ One might also search in vain for any current documented basis for the ADE (or DEF) values now used. No organization is responsible for scientifically developing and documenting the rationale for these scores—or for replacing them with a more theoretically valid concept—even though policy-level work *requires* simplified methodologies at approximately this level of resolution.

- *Bureaucratic Implications of Verification and Validation.* There is a tendency to confuse “approval” with “validation” and a tendency by some organizations to use concern about verification and validation to improve their own positions (e.g., by attempting to veto or disparage efforts or models they do not control) and eliminate competition.

Examples of Questions That Can't Currently Be Answered

It may be appropriate to end this section with examples of recent policy-level problems for which the analytic community has been ill prepared. First, consider that in contemplating possible deep force reductions in 1989, NATO was concerned about force-to-space issues and the possible existence of an “operational minimum” below which reductions should not be made even if one could be confident of force parity at the theater level. Understanding and evaluating this operational minimum was therefore an important analytic problem. When policymakers asked about it in early 1989, however, estimates of the operational minimum varied from 15 to 45 divisions and those presenting the estimates did not initially discuss uncertainties or even acknowledge the inconsistencies among analyses! Further, analysis depended on anecdotal, ambiguous, and misunderstood rules of thumb about divisional coverage capabilities that were and are nowhere justified with modern empirical experiments or credible and detailed simulations.¹⁷

A second example involves the relative value of different tactical aircraft. Some military officers and analysts believe that A-10s should be the first aircraft to be eliminated as part of force reductions. Others believe they should be the last. This long-standing disagreement is not resolved by merely running theater-level simulations, because most or

¹⁶RAND has developed a procedure for making situational modification of scores in the RSAS, but the methodology, developed primarily by Patrick Allen, has not yet been calibrated or tested. Plans call for this to be accomplished in Spring, 1991.

¹⁷See Davis (1990b) and Davis, Howe, Kugler, and Wild (1989) for discussion of these matters, including the basis of controversy.

all of the models are structurally biased—omitting known but “soft” influences of air forces, although in different ways. More detailed physics-level simulations are probably needed, as well as new historical analyses to better characterize—by situation—the observed effects of tactical air forces on maneuver and logistics.¹⁸ We also need specialized field tests and innovative types of gaming to provide at least bounding information. To our knowledge, at least, there has been little systematic and “scientific” cross-model, cross-discipline effort to illuminate the issue, although recent work on close-support led by Bruce Don and Fred Frostic at RAND suggests that much can be done with modern simulation and visualization techniques; also, recent parametric work on the Persian Gulf demonstrates how critical countermaneuver and counterlogistics effects probably are (Shlapak and Davis, forthcoming). The war with Iraq may greatly improve our empirical base.

As a final example of how the analytic community has failed policymakers and general officers planning operations, consider that in November, 1990, when the United States was contemplating the need for *offensive* ground operations against Iraqi forces in Kuwait, the combat models being used throughout most of the analytic community focused almost exclusively on *defensive* operations by the United States and its allies. Neither the models nor many analysts were prepared to explain what would be required for a successful allied offense (e.g., to investigate in detail how U.S. tactical air forces and B-52 strategic bombers could be best employed to support an offensive, taking into account qualitative Iraqi weaknesses).¹⁹

¹⁸One interesting compilation of historical information on this is Molnar and Colyer (1988), done at the Warrior Preparation Center.

¹⁹Presumably, the Joint Chiefs of Staff and the Central Command were exceptions here, conducting realistic human war games and analysis in support of deliberations about the offensive option, but we can only speculate on the matter.

II. A SYSTEM-LEVEL DIAGNOSIS

OVERVIEW

That serious problems exist in combat modeling is by no means a new observation,¹ but we believe that it has not previously been recognized that the problem is national and *structural*, rather than something that can be cured by one or more of the military services merely spending a bit more money and trying harder. Fig. 2.1 depicts our view of the problem as a whole. It is an "influence diagram" in which, if an arrow connects two items, an increase in the first item (at the tail of the arrow) tends to cause an increase in the second (at the head of the arrow).

The claim of Fig. 2.1 is that from a national perspective the state of combat modeling is characterized by chaos, a chaos that has produced numerous inadequate models and tools, inadequate verification and evaluation, general confusion, many wasted efforts, and many lost opportunities. The principal causes, we argue, are four: (a) inadequate theories, methods, standards, and practices for both modeling and evaluating models;² (b) certain computer hardware and software problems (e.g., of interoperability and software quality); (c) dissonance across communities and organizations; and, importantly, (d) the lack of military science (or, at least, a vigorous one).³ The lack of a military science is also a factor in the dissonance across communities and the inadequacy of current theories, methods, standards, and practices. These "causes" are themselves problems; in Figs. 2.2-2.5, we discuss their causes. We also begin seeing, implicitly, what remedies might be considered (the subjects of Sections III and IV).

THE LACK OF A VIGOROUS MILITARY SCIENCE

We discuss the issue of military science in more detail below, but let us first finish describing our diagnosis. If the lack of a vigorous military science is one of the causes of chaos (Fig. 2.1), then Fig. 2.2 indicates the principal reasons we see for this lack. One reason (lower right of figure) involves the complacency stemming from not having had a big war for decades (the nature of combat in Vietnam was quite unlike that envisioned for the big war with the Soviet Union and was not something combat models were expected to describe

¹ See for example Hardison (1979), Comptroller General (1980), Low (1981), Callahan (1982), and Defense Science Board (1988).

² Following suggestions of RAND colleagues Steve Bankes and James Hodges, we use the term "evaluating" here, in preference to the commonly used "validating," because the latter connotes a certainty and precision of evaluation that is generally inappropriate in combat modeling. More on this later.

³ To what extent the United States has a true military science is discussed later in the paper. We have in mind a good deal more than is covered in existing departments of military science in the war colleges, for example.

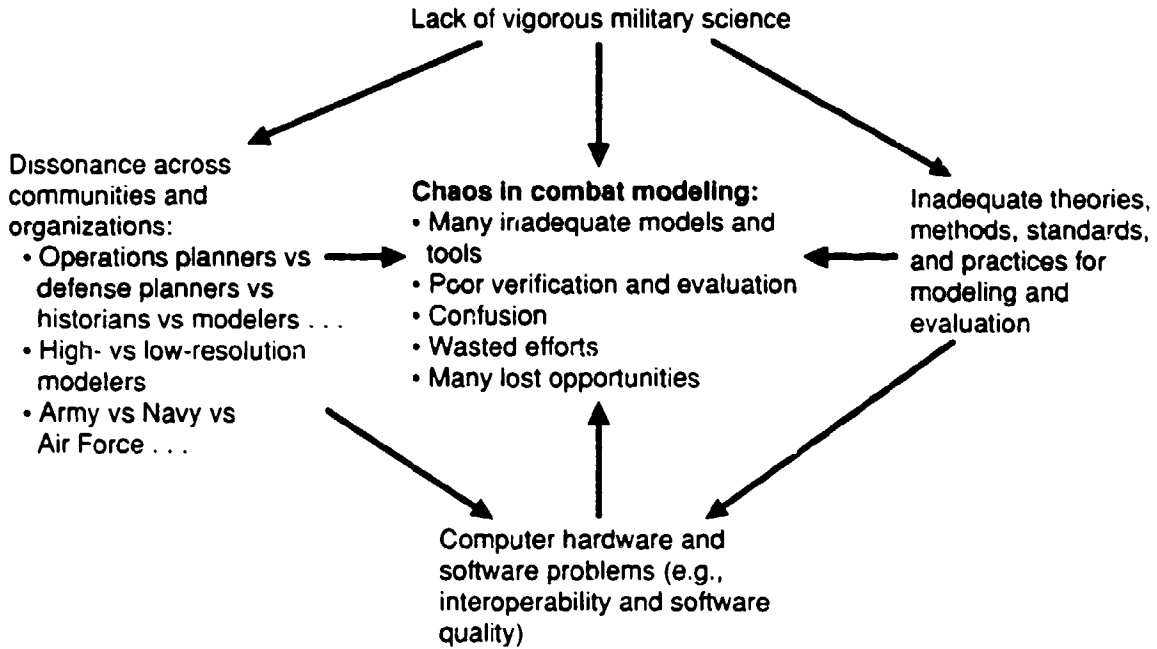


Fig. 2.1—Factors influencing chaos in combat modeling

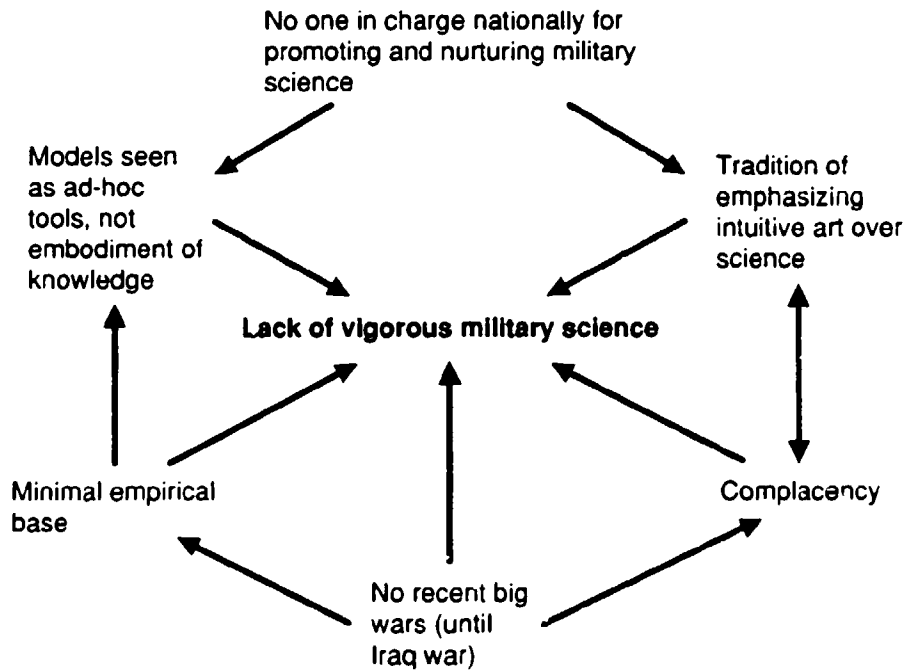


Fig. 2.2—Factors influencing the lack of military science

accurately), the tradition of emphasizing military "art," and—importantly—the simple fact that no one has been in charge nationally of promoting a military science.

Another important reason, in our view, is that to the extent there are theories of combat, they are represented in models, but models have typically been seen (top left of figure), at least by policymakers, agency directors, and analysts, as mere "tools," not embodiments of "knowledge" (including alternative theories and hypotheses and different types of objective and subjective data).⁴ This has been particularly the case in higher-level analysis, such as that conducted in and for OSD and the Joint Staff. Indeed, the classic texts on systems analysis developed by RAND in the 1950s and 1960s emphasize this view,⁵ in which the problem being worked is paramount, one develops models adequate for addressing that problem, and one has no particular aspirations for the model beyond that. Indeed, the reductionism of systems analysis includes the ethic of eliminating from models anything that is not necessary for the problem at hand. A corollary of this perspective is that one may very well discard the model altogether after completing a particular study. Why not, if it is merely an ad hoc tool, not something of value in itself?⁶

Government organizations using detailed models often have a different attitude that also acts against relating models to a military science: They see models as machines for producing what their organizations are required to produce, which is a bureaucratically acceptable product, not necessarily something well based in empirical reality. We have heard countless complaints about this from frustrated analysts working in such organizations. The analysts would like to spend more time on research and true analysis, but the pressure to meet deadlines, coupled with the demands on time necessary merely to care for and feed big models, results in "turning the crank" on the existing model. This problem can also be found in the studies and analysis organizations, such as the FFRDCs, but much less commonly. Interestingly, the events in Europe and emergence of new challenges and new scenarios have stirred the pot, with previously "acceptable product" being recognized as not appropriately acceptable. The Joint Staff's J-8, for example, has called for a period of innovation in the use of human gaming and a range of new special-purpose

⁴Many of the modelers and analysts, however, have implicitly or explicitly identified with an approach closer to that we recommend. There has been, over the years, a great deal of solid research on combat models by FFRDCs, national laboratories, government agencies, and a few commercial contractors, such as Vector Research. Its impact would have been much greater, however, if there had been an organized military-science community. We had considered listing some of the many excellent studies or documentation reports, but have chosen not to, for fear of offending some of our community colleagues by unintentionally omitting their contributions.

⁵See, for example, Quade and Boucher (1968).

⁶As part of a PhD dissertation (Thompson, 1987), Michael Thompson observed several years ago how this attitude had been a factor in RAND's not having polished, verified, validated, and documented many of its own combat models. As he noted, however, the models were *not* discarded, but were in fact used repeatedly over a period of years. Thus, they *should* have been approached as something very different from a temporary and ad hoc tool.

models. Other organizations have done similarly. The atmosphere, then, is better than in previous years, because people are looking afresh at problems.

DISSONANCE ACROSS COMMUNITIES AND ORGANIZATIONS

Another factor contributing to chaos is, according to Fig. 2.1, the dissonance across important boundaries.⁷ For example:

- Operational commanders often have difficulty communicating with combat modelers and civilian analysts concerned with building and explaining the defense program, and vice versa.
- Combat modelers working with high-resolution models, such as JANUS or VIC, typically interact seldom and poorly with modelers and analysts working at lower resolution (e.g., analysts seeking to characterize military balances, assess alternative operational strategies, or assess the potential value of different force structures). It is easier to do better when the individuals are part of the same organization, but even then, the difficulties are substantial.
- Modelers and analysts in or working for the various military services often have difficulty communicating well, because they see the issues so differently, and each is in an organization that is competing for scarce resources.

Fig. 2.3 indicates some of the influences that cause this dissonance. The lack of military science is primary, because it is from a science that one draws theories and methods to accomplish cross-cutting work. Again we cite the role of complacency and the absence of a national authority to encourage, nurture (and, frankly, enforce) cooperation. Parochialism and competition are major factors here, although the Goldwater-Nichols act and the resulting reorganization may be influencing the situation favorably by strengthening the coherence of the Joint Staff and the roles of the Chairman of the Joint Chiefs of Staff and CINCs. There has been considerable evidence of high-quality joint and combined planning in the war with Iraq and, before that, in the invasion of Panama. Similarly, there has been increased attention to such matters in the war-college curricula and studies conducted within the DoD and in some of the FFRDCs.

⁷This has troubled one of us for some years and influenced design of the RAND Strategy Assessment System (RSAS) (see Davis, 1986).

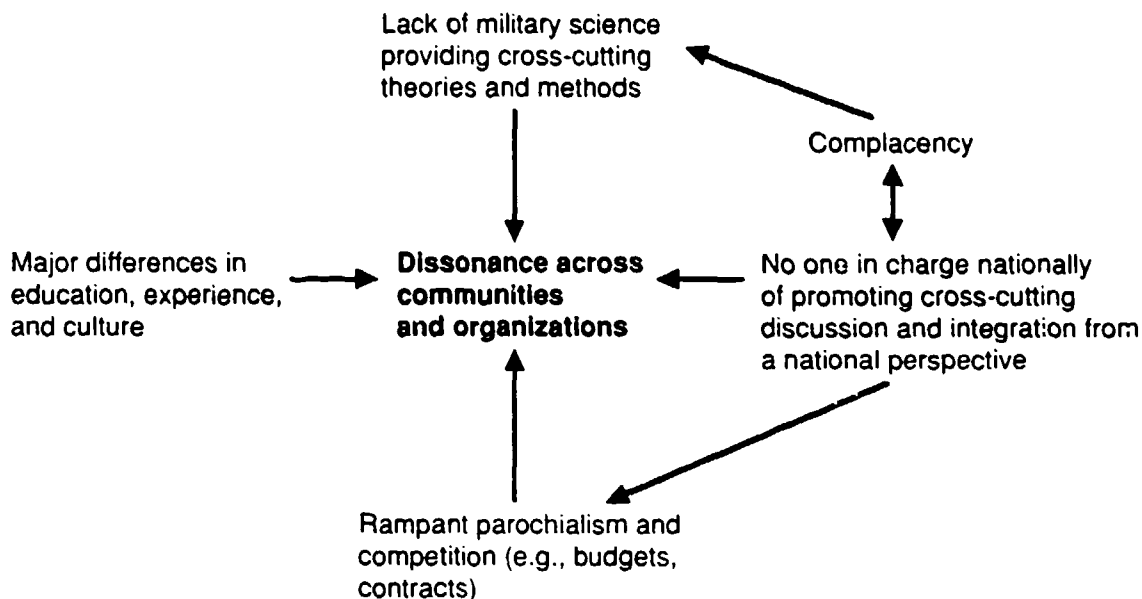


Fig. 2.3—Factors influencing dissonance

INADEQUATE THEORIES, METHODS, STANDARDS, AND PROCEDURES

Fig. 2.4 elaborates on factors influencing the inadequate theories, methods, standards, and procedures for modeling and evaluating models. These inadequacies are especially serious (left side of figure) when one is dealing with subjects on which there is a great deal of uncertainty—about both analytic relationships and the values of parameters in those relationships. Once again, the lack of a vigorous military science is prominent, because it is science-style work that produces theories, methods, and the like. Further, it is in science-style work that one makes a point not only of testing in the narrow (e.g., verification testing and testing to see that an organization's new model gets the same answers as its old one), but of outside comparison: One exposes one's work to the outside world and one tries to understand and reproduce the work of others—not just superficially, but in depth.

Another problem highlighted in Fig. 2.4 (bottom left) is that few of those who do combat modeling have been adequately exposed, in depth, to some of the disciplines most useful to good modeling, notably computer science, simulation theory, and software engineering.⁸ And, to make things worse, there has been relatively little experience tying complex and squishy models, such as combat models, together in a rigorous way, especially

⁸Here we may seem to be confusing "modeling" and "programming," but we are not. In practice, there is a strong relationship between the quality of model design and specification and the quality of the program built to implement the model. A good understanding of software engineering is, in practice, very helpful in the modeling itself. For discussion of some of these issues, see Allen, Bennett, Carillo, Goeller, and Walker (1991).

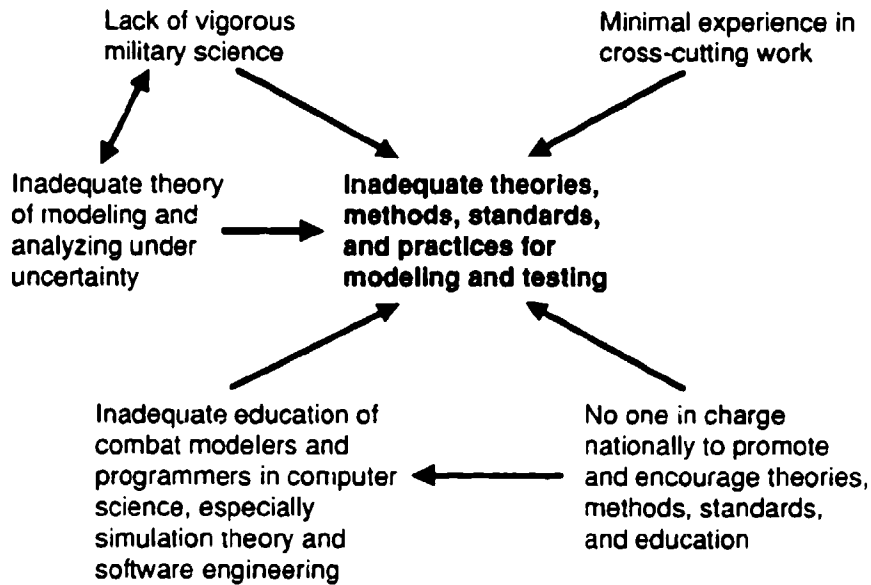


Fig. 2.4—Factors influencing inadequacy of theories, methods, standards, and practices for modeling and testing

across levels of resolution or across submodels built by different organizations. Various organizations, such as the Warrior Preparation Center, Air Force Studies and Analysis, RAND, and IDA, have at least some experience with such matters, but we do not believe any of them would claim a high degree of rigor. To the contrary, the lashups have typically been accomplished with the software equivalent of bailing wire. Germany's IABG seems to have done better over the years.

As with the other problem areas identified in Fig. 1.1, an overarching problem here is that no one is in charge (Fig. 2.4, bottom right)—in this case, to promote and encourage the needed theories, methods, standards, education, and so on.

COMPUTER HARDWARE AND SOFTWARE PROBLEMS

The remaining major contributor to chaos is, according to Fig. 2.1, various computer-related problems such as the noninteroperability of models, computers, and graphics programs, and the uneven quality of model software (i.e., the software programs implementing the models).⁹ Fig. 2.5 identifies what we see as major factors in this problem

⁹Here we distinguish between the quality of models and implementing software programs, having in mind a distinction between substantive content and such software characteristics as structure, performance, and maintainability. As noted earlier, however, well-designed models provide the specifications for good model software.

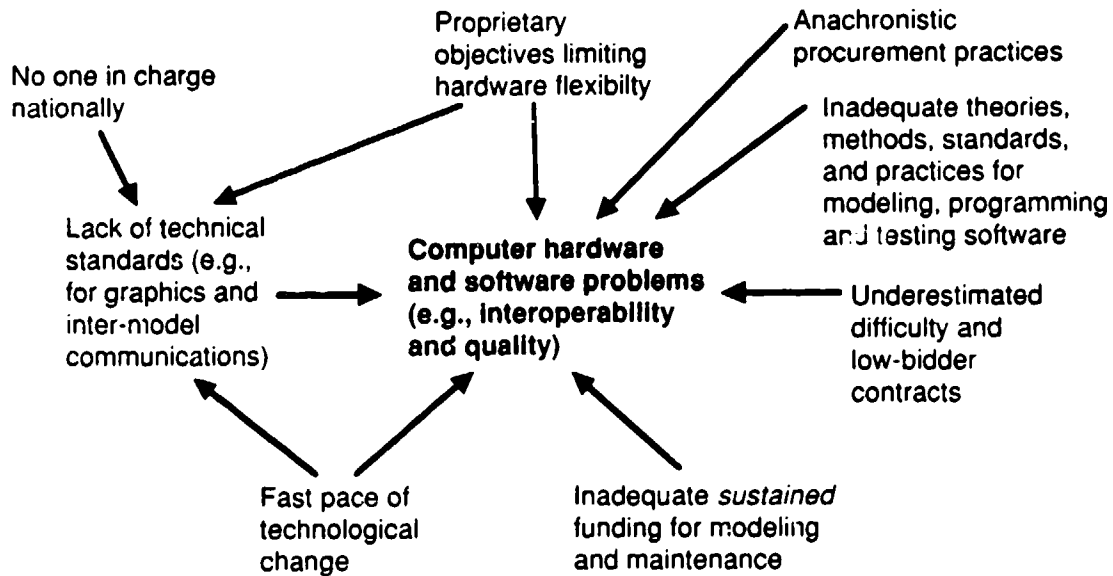


Fig. 2.5—Factors influencing computer hardware and software problems

class. We do not elaborate on most of these issues in this paper (e.g., we don't discuss the procurement practices that leave the DoD with obsolete computers and software, the chronic tendency of agencies to underestimate how difficult it is to develop high-quality models, the associated chronic tendency to underfund development efforts and select lowest-cost bidders who must pad their staff with poorly trained junior workers to win the competition, and the inadequate attention given to sustained product improvement over a period of years after an expensive model development effort¹⁰). Nor need we elaborate on the pace of technical change, which is evident. It is worth elaborating, however, on the lack of technical standards (left side of figure) and the decidedly mixed quality of model and interface software that has been produced (right side of figure). The principal points to make are probably these:

- Without a national organization managing the process, it is probably impossible to achieve interoperability of models. The tendency of each modeling group is to develop its own data structures and protocols unless there is a recognized standard and incentives to use it (or to be able to meet the standard upon occasion).¹¹
- The uneven quality of computer software is attributable to chronic underfunding, unstable funding (it takes a number of years for complex models and other

¹⁰The need for an evolutionary approach to product improvement is described by RAND colleague Bruce Bennett in his paper in MORS (1989).

¹¹Similar points were made in Defense Science Board (1988) and, in work accomplished in 1988, in Bankes (forthcoming).

software to mature and reach potential), and the lack of both mechanisms and incentives for verification, evaluation, and straightforward comparisons.

- In addition, the quality problem is due significantly to the lack of a well-developed theory of simulation modeling *for complex systems*. Ultimately, model software is limited by the quality of the model specifications, and these are often developed by people with no formal training in modeling and, often, little education in advanced mathematics or science.¹²
- As noted in the influential book, *The Mythical Man Month*, top-notch programmers are a factor of ten more productive than less capable ones (Brooks, 1982, p. 30). Qualitatively, at least, the same is true for modelers and analysts. The moral here is that one cannot successfully do serious combat modeling on the cheap.

Having provided a system-level diagnosis of the problem, let us now discuss in somewhat more detail the issue of military science. Does it already exist? Can it exist? Should it exist?

THE ISSUE OF MILITARY SCIENCE

Conflicting Counterassertions

Our assertion that the United States (and West) lacks a military science (or at least a vigorous one) is sometimes challenged for both of two opposite reasons. Some claim that there *is* a military science: It is what professional military officers learn over decades of experience and the study of history, doctrine, and problem-solving procedures worked out in countless field units and headquarters. Others claim, by contrast, that there can be no real "military science," because the phenomena at issue are too complex and too muddled by human factors, such as fear, genius, and fatigue, and by random factors, such as weather. In this view, one should talk of the "art" of war, as did Sun Tzu. Let us consider each view briefly.

Is There Really No Military Science? Art vs Science

Is it valid to claim that there is no military science in the United States, or at least no vigorous one? We believe the answer is "yes," although there have been important informal

¹²Another problem here is that simulation modeling is not a large, well-developed, or prestigious discipline. Few texts and few departments specialize in it, especially in simulation modeling of truly complex systems. Jay Forrester's work in *Systems Dynamics* at MIT was an exception. So also is the work of Bernard Zeigler at Arizona State (see, e.g., Zeigler, 1964 and 1990). RAND's RSAS represents a major effort to model military complexity from a top-down perspective (Davis and Hall, 1988).

efforts to develop such a science,¹³ and there are portions of an implicit science that can be detected if one looks hard enough. Further, there are many war-college departments of military science that teach important parts of what would go into what we would regard as a more vigorous military science. Perhaps it would be better to say that in the United States the formal study of military affairs (e.g., by military officers in the war colleges) and the practice of most military officers, combat modelers, and analysts is not structured around a concept of military science that is analogous to other sciences. Instead, it emphasizes more the dimension of "art."

But what do we mean by military science, as distinct from military art? In fact, the distinctions are unclear. Consider first the dictionary definitions (*Webster's New Collegiate Dictionary*, Merriam Webster, Springfield, Mass., 9th edition, 1989):

Art: skill acquired by experience, study, or observation

Science: knowledge covering general truths or the operation of general laws, especially as obtained and tested through scientific method.

The scientific method: principles and procedures for the systematic pursuit of knowledge involving the recognition and formulation of a problem, the collection of data through observation and experiment, and the formulation and testing of hypotheses.¹⁴

From this, it is not immediately evident how art differs from science. We believe, however, that most people associate "art" with a combination of experienced-based skill and *intuition*, as distinct from knowledge and skill based on a relatively deep and general theory of relationships, and as distinct from skills that depend more on "calculations" and rigorous logic than intuition. Associated also with "art" is developing methods to solve problems as they arise rather than developing general methods.

Consistent with this sense of the distinction, the knowledge associated with "art" or "skill" is often embodied in procedural knowledge, such as doctrine manuals and standard operating procedures, coupled with concepts and beliefs communicated through anecdote and a variety of other informal or descriptive mechanisms (such as studying the great Captains'

¹³Notable here are the efforts of the Military Conflict Institute and Trevor Dupuy's HERO group. The issue of military science has also been discussed in conferences over the decades. See, e.g., Callahan (1982) and Low (1981).

¹⁴Unfortunately, these dictionary definitions fail to mention that, in both science and the study of combat phenomenology, the "testing" of hypotheses is accomplished not only by comparing with objective data, but also by applying such criteria as theoretical simplicity (Occam's razor) and intuitive credibility. Thus, even physical science is much less cut-and-dried than is often realized, and the art vs science distinction breaks down if one believes that science is what scientists do. In the military domain, and in parts of the social sciences, it is flatly impossible to validate in detail many of the considerations essential in any reasonable model.

military campaigns, as recommended by Napoleon). The knowledge may not extend explicitly to how one should deal with unusual situations and why.

The enduring *Principles of War* are an example of the military art. They are fuzzily stated and to some extent mutually contradictory. They are unquestionably of fundamental importance, but the way they are stated and the fact that there is relatively little structured guidance on how to resolve the contradictions is part of what we mean by there being no military science. As another illustration of military art, consider how much of the "real military knowledge" is embodied only in organizational practices (e.g., at the level of an Air Force wing) rather than in textbooks, formal doctrine manuals, or journals laying out the various problems and situations methodically.

Finally, we assert that anyone attempting to learn about military operations so that he can build a good combat model needs no convincing that the U.S. military approaches the subject of war as an art rather than a science: He can't find anything remotely like the textbooks he used to learn other disciplines. His view is reinforced when he sees how little effort goes into empirical studies or attempts to understand and reproduce the work of others as part of learning and converging on "truth."

Can There Be a Science Amidst the Friction of War?

The other view is that there can be no real military science, because of the manifold uncertainties in war. Proponents of this line sometimes wax eloquent on the subject, and usually quote Clausewitz (although, in our view, distorting his message). Quoting from a recent article warning against the folly of slipping into a body-count mentality, which in turn is quoting Clausewitz:

They [those who overquantify] aim at fixed values, but in war everything is uncertain, and calculations have to be made with variable quantities. They direct the inquiry exclusively toward physical qualities, whereas all military action is intertwined with psychological forces and effects. They consider only unilateral action, whereas war consists of a continuous interaction of opposites. Military activity is never directed against material force alone. It is always aimed simultaneously at the moral forces which give it life.¹⁵

One *might* conclude from these comments of Von Clausewitz that numbers and calculations are irrelevant and that one should fall back on intuition and military genius. However, such a conclusion would be a mere excuse for sloppy thinking. Furthermore, concluding from such arguments as those quoted that there can be no military science reflects a serious misconception of science. Science (and engineering) abounds with problems

¹⁵Harry Summers, "Body Count Proved to Be a False Prophet," *Los Angeles Times*, February 9, 1991.

characterized by large uncertainties. Social science, in particular, must deal with at least as many fuzzy variables and effects of human behavior as does the study of war. To be sure, the science needed for coping with massive uncertainties is different from that needed to catalog the dimensions of planets or to categorize species and is exceptionally difficult, but there can be science nonetheless.¹⁶ It is interesting that some of us, in reading Von Clausewitz on these matters, conclude not that models and calculations are irrelevant, but rather that to be useful, models and modelers must include a wide range of qualitative factors and must also be humble with respect to random factors and horseshoe nails.¹⁷ Further, models become tools of *exploration*, not answer machines.

RETHINKING THE ROLE OF MODELS

Earlier, we argued that a factor in the lack of military science has been the widespread view that models should be seen as ad hoc study-specific tools. An alternative view, which one of us (Davis) has been promoting for some time, is that combat modeling should be regarded as to some extent part of a science (albeit a social science designed to improve decisionmaking under uncertainty, rather than a physical science). Another way to state this is that *we should regard many of our models as embodiments of knowledge—including knowledge of alternative hypotheses, uncertainties, and a broad range of objective and subjective data*. It is not merely knowledge of "facts." Consider that if one stops for a moment to think about how one would "write down" knowledge of (and bounded uncertainties regarding) combat phenomena, if one had that knowledge, the answer is that the natural vehicle is models. One can't really do military science without using models to represent the knowledge, including theory.

Research Models and Application Models

It does not follow that *all* combat models must be or even should be viewed in this manner. To the contrary, it seems useful to distinguish crudely between two classes of combat model, what we call *research models* and *application models*. The reason for the distinction is that there is a fundamental tension between building a model or set of models

¹⁶One of the most influential expositors of some of these issues is Herbert Simon, who won the Nobel Prize in economics for developing concepts of dealing with *bounded rationality*. These include the use of heuristic reasoning and play a central role in modern computer science, including artificial intelligence.

¹⁷Trevor Dupuy considers his QJM methodology to be based in the lessons of Von Clausewitz (Dupuy, 1987). Similarly, the Soviet military sees no contradiction between those lessons and its emphasis on quantification and military science (or, as suggested to us by John Battlella and Judith Grange, what can be understood as an engineering approach in which one seeks to study the range of possibilities so as to be able to build in an appropriate error margin). Within RAND, development of the RAND Strategy Assessment System (RSAS) has emphasized a wide range of qualitative variables as part of basic methodology. See, e.g., Bennett, Bullock, Jones, and Davis (1988) and Davis (1989).

that embody all one's knowledge in a particular subject (including knowledge about alternative theories and hypotheses and various types of objective and subjective data) and building a model to help inform decisions on a particular question, such as what weapon system to buy, what level of forces to maintain, what operational strategy to employ, or what doctrine to adopt. The reality is that when models are used for decision support, they must be relatively simple, by which in practical terms we mean they must be comprehensible to those seeking to use them. This implies that models for decision support must have fewer variables, fewer relationships, and fewer processes than models attempting to represent "what's going on" in detail.¹⁸ These application models also embody knowledge, but more selectively—so much so that in some cases they are ad hoc one-time constructs tailored specifically to the question at hand

Research Models

What form should we expect our research models to take if they are to be a repository and embodiment of knowledge? The answer, of course, is that we need a variety of different model types. In some instances, one is interested in characterizing relationships, steady-state conditions, endpoints, or optimal strategies. There are roles in such cases for such diverse types of models as static models, closed-form analytic models (e.g., solutions of Lanchester equations), game-theoretic optimization models, and artificial intelligence models using heuristics to find good strategies.

This said, however, we argue that since combat is inherently a dynamic process involving the interaction of many entities, and since a considerable part of military science deals with understanding those interactions over time, it follows that simulation models—including rather complex computer simulations—are especially important.¹⁹ There seldom is a good alternative if one wants to capture and understand dynamic cause-effect relationships in complex systems. Unfortunately, current large models are not usually designed with enough attention on transparency and comprehensibility for them to fulfill the role of representing and communicating knowledge adequately. That can be changed with time,

¹⁸To illustrate what we mean here, a recent RAND study dealing with post-war requirements for defense in Kuwait and Saudi Arabia depended primarily on a parametric analysis using a highly aggregated simulation model implemented as a spreadsheet program on the Macintosh computer. This proved very useful in discussions with both civilian officials and officers in CENTCOM and the Joint Staff—discussions focused on identifying key variables and key issues rather than precise answers. The work was complemented, however, and to a significant extent calibrated with, more detailed simulations and war games using the RSAS (See Shlapak and Davis, forthcoming).

¹⁹Another implication here is that the modeling must seek structural validity in the sense of representing correctly, at whatever level of aggregation, the systems' natural entities and cause-effect relationships. It is not enough to find statistical relationships from historical data. Indeed, it is better to use physical insight and experience-based judgments about cause-effect relationships than to omit what are believed to be the key variables of the problem.

effort, and a sense of priorities, but developing transparent and comprehensible models is extremely difficult and requires top-flight talent. Only recently have the requisite concepts and techniques begun to emerge.²⁰

Application Models

Although some application models may be research models and vice versa (i.e., the distinction is not entirely clear-cut),²¹ focusing on a particular application allows one to narrow and specialize (application) models—emphasizing clarity, flexibility for policy-relevant sensitivity analysis, reproducibility, and explicit dependence on the particular variables policymakers are able to control. That is, application models often and perhaps usually reflect a particular perspective. One should therefore evaluate application models more narrowly than one would research models: For example, do they help inform judgments, decisions, or even descriptions? More precisely, do they help, *on balance*, when due consideration is paid to their omissions, approximations, communication effectiveness, and so on?

Striking a Balance

In summary, we believe that a balance should be struck between the classical view that models are ad hoc devices developed to solve particular problems and the view that models should be seen as the mechanisms by which we express and communicate our knowledge and theories. The former view has been too dominant in the policy-analysis and decision-support communities.²²

It follows that the DoD as a whole, and not just the constituent parts, should be concerned about the development of sound research models and application models—including models of uncertainty. Further, it follows that models (both research models and application models) are too important to be treated as though they were mere tools to be discarded at the end of a study or to be developed and tuned for effective use in adversarial aspects of the budget process. And it follows that the approach to be taken *must* be more akin to science than to art. Again, however, the science at issue is a social science. As a

²⁰Some examples of this are the top-down hierarchical modeling, high-level RAND-ABEL programming language, and graphics used in the RSAS (see, e.g., Davis and Hall, 1988, and Davis, 1990a). The graphics represented in JANUS and CONMOD are another example. Object-oriented modeling and programming are yet another example (e.g., Zeigler, 1990), with related work going on at Los Alamos, TRADOC, RAND, Livermore, and elsewhere.

²¹By analogy, the simple gas law $PV = NkT$ is an important part of research and theory, as well as a simple and practical application model for low-density gases.

²²A separate problem is that the view of high-impact models as production machines has been far too dominant, since they are typically not good enough to be used for good analysis amidst uncertainty.

result, there are many "squishy" aspects and the need for forthright use of subjective judgments as well as quantitative data.

III. A VISION OF HOW TO IMPROVE THE SITUATION

In this section, we sketch out what we believe is needed, first discussing the desired environment in which work would take place and, second, describing what we see as the critical features of next-generation modeling and simulation.

ISSUES OF SCOPE, RESOLUTION, AND PERSPECTIVE

A basic problem in military analysis, as with the physical sciences, is that there exist many different levels of resolution. In physics we might think of thermodynamics at one end and quantum statistical mechanics at the other (ignoring particle physics). In military analysis the natural breakdown for many purposes is strategic, operational, tactical, and military-technical levels.¹ At each level, there are also alternative representations to reflect different *facets* of the problem or to serve different purposes (see Fig. 3.1). For example, much policy analysis relies upon static or quasi-dynamic analysis in the form of Red vs Blue

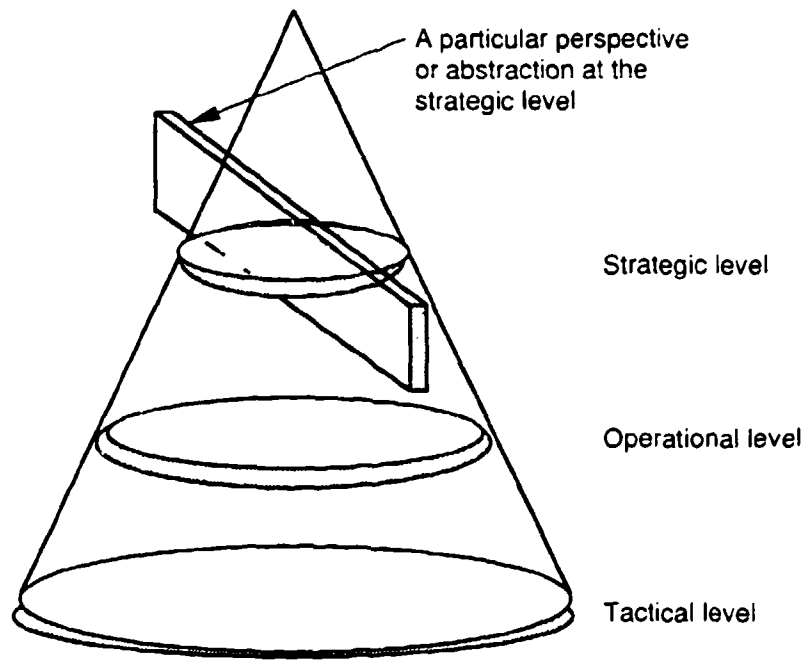


Fig. 3.1—Different levels of resolution and perspective

¹These "levels" are themselves difficult to define adequately and do not correspond neatly to the size of units. In a small theater, activities that would be "tactical" in larger conflicts may be operational or strategic in function and significance.

"equivalent divisions" (EDs) vs time. Such analysis can cut through the morass of complexity and focus attention on critical issues (e.g., the importance of prompt response to strategic warning, strategic mobility systems, and allies). Such analysis can be embellished for specific problems and made to represent rather more complex issues (e.g., the difference in anticipated operational-level force ratios in different corps sectors, perhaps to explain the importance of achieving the time goals of a POMCUS program for Europe or Saudi Arabia).

In contrast, other strategic- or theater-level analysis may employ closed theater-level simulation and gaming with hundreds of entities and an explicit representation of terrain. So, for example, the SACEUR or the Commander of AFCENT would not be satisfied with war planning or analysis expressed in the simplified terms of the policy analysis. Such commanders should, however, be interested in *other* simplified treatments abstracting the principal elements of specific problems on which they are working. For example, theater commanders are often comfortable with the concept of "equivalent divisions," however flawed the calculation of equivalent division scores may be in detail. They are also comfortable—for *some* purposes—with simplified depictions of concentration, such as that shown in Fig. 3.2, which represents *one* perspective of a situation that also involves, for example, terrain, lines of communications, and force quality, none of which are shown. For other purposes, those same commanders need detailed depictions of one or more aspects of the problem. The rule of thumb here is that a commander at one level will sometimes need to see depictions of the situation that are two to three levels more highly resolved than his natural level. For example, a corps commander will sometimes need to see the deployments and missions of individual battalions (which are determined by brigade-commanders subordinate to division commanders, who are subordinate to the corps commander).

The points to be made here are:

- There is a fundamental need for *variable resolution models* (or families of models) in which there is true consistency across levels² and for concepts and methods making it easier to do cross-resolution work, including work with models not originally designed to be compatible.³

²Such consistency is often difficult to achieve, but what we are talking about is the same consistency one has in going from lower to higher scale maps within a family. As those familiar with theoretical work relating microscopic and macroscopic worlds in physics recognize, one cannot expect to produce accurate aggregated depictions by merely integrating high-resolution depictions; to the contrary, because of the propagation of uncertainties in high-resolution "data," the results of such integration are typically inaccurate (e.g., implausibly high attrition rates). Consequently, a sound theory must allow for inserting empirical and judgmental data at several levels in the overall system, including at the top level (e.g., with historical division-level attrition rates), and then iterating to achieve a multilevel stability and consistency.

³RAND has begun studying such issues in detail under DARPA sponsorship.

- There is also a fundamental need, at every level, for *variable perspective models* and for *special-purpose abstractions*, including simplified models and even rules of thumb.

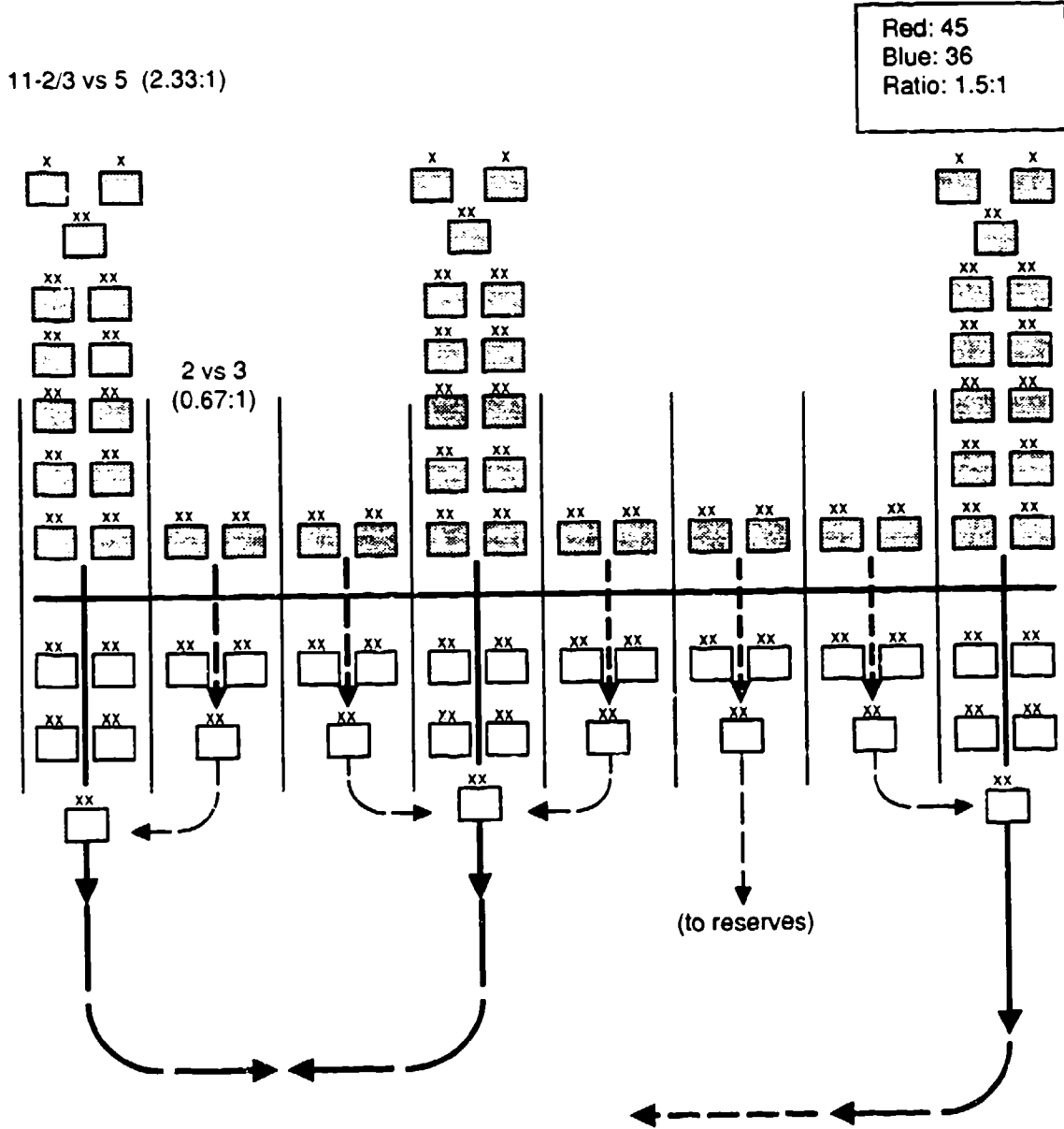


Fig. 3.2—An illustrative, very simple model (of concentration and counterconcentration)

We emphasize these matters because chronic misunderstandings exist about issues of resolution. Individuals and organizations working with high-resolution models commonly look down upon aggregated models and aggregated rules of thumb (e.g., the “3 to 1 rule”), and those working with low-resolution policy-oriented models similarly look down upon the more detailed models. It is uncommon for modelers and analysts to move back and forth among levels of resolution and perspectives.⁴ In years past, doing so was not feasible computationally, but the limiting factors now are conceptual and organizational.

A COMPREHENSIVE APPROACH

In addition to considering alternative levels of resolution and perspective in the modeling itself, a comprehensive approach needs to attend seriously to all the items in Fig. 3.3, which highlights the fact that at least three different classes of activity can add to our body of data: experiments (including exercises), historical analysis, and systematic interviewing of officers with relevant experience. Data, in turn, should inform theory, especially in the form of research models of various types. Theory leads to simplifications and specializations in the form of application models, analytic methodologies, and tools such as graphical aids. The theories, methodologies, and tools all contribute to strategy,

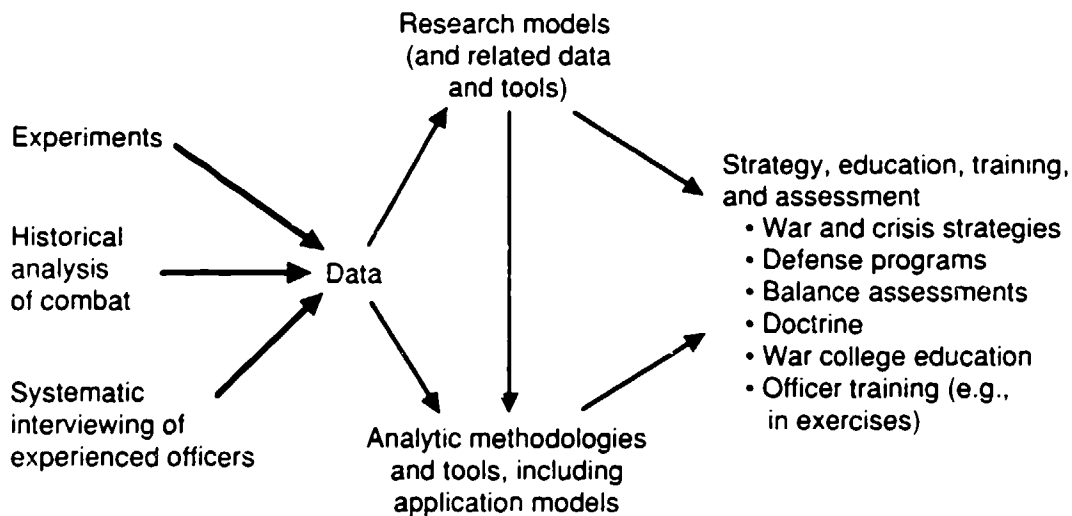


Fig. 3.3—Components of a comprehensive approach

⁴One reason that the participation of general officers is important in the development of sound and useful models and methods is that they have worked at different levels and from different perspectives and are better able to appreciate the need for all of them. Whereas junior staff almost invariably seek more detail (in areas in which they are knowledgeable), general officers and some senior staff often appreciate or seek abstraction.

education, and training in all their dimensions, some of which are noted as war and contingency strategies, defense programs, military balance assessments, doctrine, war college education, and officer training. Not shown, for simplicity, are the many feedback loops. That is, in a comprehensive approach, theory informs what experiments are conducted, what historical analysis is commissioned, and so on.

Today, the weight of effort is on building and using models as tools, developing orders-of-battle data bases, and conducting analysis. Minimal attention is paid to "theory" per se, and little support or technical guidance is given to experiments, historical analysis, or the collection of anecdotal and subjective information. While it is true that combat models will continue to be driven heavily by postulated relationships that cannot readily be proven or disproven in detail, much could nonetheless be gained from carefully designed experiments, historical studies, and interviews. It is not an all-or-nothing proposition,⁵ as can be seen from experience with National Training Center data and SIMNET experiments.

MODEL ATTRIBUTES NEEDED

Because models play such an important role in our scheme, we will first describe what we see as desirable model attributes. We will then turn to such issues as how to create the general environment in which the approach of Fig. 3.3 could best be conducted.

Whenever meetings are held to discuss desirable attributes of models that should be emphasized in next-generation work, certain suggestions almost always emerge. Some ideas date back years,⁶ but much has been learned in the interim, and technological advances make some things possible that were impossible in earlier days. Also, in some cases, we have attempted to be more realistic than the cliches of a decade ago (e.g., in our comments on modularity). Most of the items apply to research models, but some also apply to applications models. Throughout the list, one should interpret "Models should" as meaning "We need at least some models that . . ."

The desired attributes of models and modeling are as follows:

- Models should be consistent with and reinforce the principles of war.

⁵As an example here, physics-level simulations should be able to place credible *bounds* on defensible division frontages (e.g., bounds determined by ability to service targets rapidly enough to avoid being overrun, as a function of terrain, weather, and other factors). Appropriately stochastic models might do much to illuminate the risks and potential benefits of alternative tactics.

⁶For a description of the objectives that went into RSAS development, see Davis and Winnefeld (1983). For a highly technical description of the results, see Davis and Hall (1988). See Davis and Howe (1990b) and Bennett, Bullock, Jones, and Davis (1988) for a mix of philosophical discussion and model descriptions. Most detailed RSAS documentation is restricted to government agencies. For a description of Lawrence Livermore's CONMOD, see Chiu et al. (1987).

- **Models (or families of models) should be developed hierarchically to correspond to real-world command levels.**
- **Different *kinds* of models should exist to serve different needs, but there should be strong professional and governmental pressures to establish relationships among them (e.g., comparisons, calibrations, or merely clear descriptions of similarities and differences), as well as to validate them on their own terms, which may be quite different from class to class of models.⁷**
- **Overall modeling activity should not be dominated by any one approach (i.e., neither the "bottom up" nor the "top down" approaches are "right").**
- **Models should support multiscenario analysis in many dimensions (e.g., Davis, 1988a). This requires unusual flexibility and so many degrees of freedom as to make those seeking cut-and-dried answers uncomfortable.**
- **Variable-resolution features should be designed in from the outset, as should features allowing for iterative cross-level calibration, tuning, and consistency checking. As noted earlier, calibrations should use low-resolution data, as well as high-resolution data. The bottom-up paradigm is wrong.**
- **Models should reflect stochastic effects, at least optionally. Further, deterministic operations should be based on results of stochastic modeling. This may require distinguishing among cases rather than settling on expected-value calculations, because outcome distributions are so often skewed or multimodal.**
- **Models should be sensitive to command-control choices at all levels; they should include optional submodels optimizing or satisficing-decisions as a function of information and criteria; they should also include optional submodels describing *likely* (and probably very nonoptimal) decision processes, accounting for human limitations well known in behavioral psychology (e.g., the inability to understand "sunk-cost" issues). Decision models should operate on the basis of perceived information, not perfect information, except for limiting-case analysis.**
- **Adaptive strategies and tactics should be emphasized; scripted strategies and tactics should be avoided, since they teach stereotypes and encourage complacency.**
- **Models should be sensitive to Reconnaissance, Surveillance, and Target Acquisition (RSTA) factors.**

⁷Some models are calibrated roughly to historical data; others use test-range weapon data as inputs, although they usually include scaling factors to bring their results into rough accord with expectations. Neither approach is "correct."

- New models should be modular to permit plug-in plug-out operations within an organization and should be designed so that suitable models can be readily constructed⁸ for cross-organization review and use (e.g., use in distributed war gaming).⁹ However, this is not likely to be feasible without a *far* greater degree of consistency in data structures than typically exists. That, in turn, is infeasible without standards, such as could be implemented through a common data dictionary approach.¹⁰
- Far greater emphasis should be placed on comprehensibility and appropriate documentation. This has major implications for the choice of computer languages, data base management systems, and on-line "helps" and graphics, as well as more traditional issues, such as simply writing down the algorithms. Much of what must be comprehensible is inherently qualitative or judgmental, not quantitative and empirical (e.g., key elements of operational strategies and decision rules within them).
- Technical *standards* must be developed if it is to be possible to exchange and review models across organizations, or to fully exploit the potential of distributed war gaming. These standards would make it possible to develop appropriate data dictionaries and practical documentation using a diversity of techniques ranging from diagrams to equations and discussion.¹¹ Standards are also needed on the subset of specific data items (variables) to be communicated. Importantly, however, developing the standards is a *research* problem that will take time, experimentation, and iteration. Premature or inflexible application of standards would be counterproductive.
- The traditional distinction between gaming and closed simulation is obsolete. Models for analysis often need to be interruptible and flexible. Also, a gaming style of analysis is often quite valuable.

⁸Our phrasing here is important, nontrivial, and in contradiction to what is often claimed as virtuous by those who have not yet built the model at issue. It is not feasible to require that all models have the same modules, not even models of the same general class. Further, different modularity is needed for different studies. Thus, the design challenge should be to anticipate the need to be able to construct exportable modules using standard variables (if only a standard Data Dictionary existed).

⁹Modularity is an ambiguous concept, however, and we are by no means endorsing a particular programming language or a particular computing environment.

¹⁰Modularity should be natural, not artificially imposed. If two processes are tightly coupled in the real world, they should not be depicted as loosely coupled modules. Tactical air operations and ground operations are a good example of where the usually assumed coupling has been overly simplistic, causing serious problems.

¹¹The Macintosh program HyperCard illustrates one type of flexibility needed.

THE ENVIRONMENT

If military simulation is to reach its full potential, it should be conducted in an intellectual environment with the following features:

The Research Environment

The first requirement is to have the concept of research and development (R&D) be as strong with respect to combat modeling as it is in other aspects of DoD work. Such R&D funding needs to be steady and focused on the midterm and long term. We cannot overemphasize how difficult a problem area this is.¹² The funding and management environment should also include

- Long-term funding of military combat modeling as applied science, not mere tool building.
- Substantial redundancy and overlap of research—precisely the opposite of what is frequently recommended by panels—but with strong incentives and requirements for open publishing and comparison efforts.
- Protection of research from such common problems as organizational shortsightedness (e.g., an emphasis on doing the current job faster rather than better), corruption (e.g., an emphasis on supporting the organization's natural positions in the adversary process), and parochialism (e.g., ignoring the roles of political factors, allied military forces, other services, or other theaters).
- Exposure to sunlight through peer-review processes and publications.
- Standards and practices to improve model verification and validation (by which, in practice, we may mean evaluation), and to make model transfer and interoperability feasible.¹³
- The development of academic-quality textbooks.

¹²As data points, models such as JANUS, JESS, and the RSAS are developed over periods of many years and cost tens of millions of dollars, ultimately. Success requires dependence on top-quality senior people. Our development experience also redemonstrated the principles cited in Brooks' *The Mythical Man Month* in that a large fraction of the work was accomplished by a small number of key people with exceptional talent. Such trailblazing efforts can no more be accomplished by people of average talent or education than analogously important efforts can in engineering or physics.

¹³We emphasize again, however, that complex combat models cannot be "validated" in the sense of certifying their correctness: There are simply too many uncertainties and unknowns. Further, a centralized effort to "certify" only certain models and data bases could be *seriously* counterproductive. "Blessed models" are not the same as "valid" models, and a great deal of mischief can be accomplished under the rubric of centralized quality control and certification.

Military-Science Society

It is difficult to imagine medical research without *The New England Journal of Medicine* or physics research without *Physical Review*. Further, it is difficult to imagine a quantum leap in the quality of military research without numerous seminars and conferences presenting scholarly papers for peer review and the associated intellectual rewards. The closest thing that now exists is the Military Operations Research Society (MORS), which is highly valuable but which has not attempted to take on the functions at issue here. Regular MORS conferences have a low signal-to-noise ratio, papers are not reviewed or published, and only a moderate percentage of participants are professional military analysts or scientists.¹⁴

A major problem is the lack of military journals. Here, the closest thing we know of is the classified *Journal of Defense Research*, which is not widely read—in part, probably, because it is indeed classified. Thinking must often be done in researchers' homes—during the evening or weekends. This is especially so when it is technically difficult work. The academic journals *International Security* and *Defence Analysis* are also of some value to military science.

QUALITY PEOPLE

As all managers know, the *sine qua non* for solid research and analysis is top-quality people with appropriate backgrounds. It is not enough to have bright people or people with technical degrees. One must have *very* bright and exciting people with appropriate backgrounds of advanced graduate work, often in "hard subjects," such as the physical sciences, engineering, operations research, or computer science, although some top-notch people are also needed in softer disciplines, such as history and behavioral science. Such people gravitate toward interesting fields and interesting institutions. Thus, our vision of the future includes improving the attractiveness of military research and modeling. It also includes the development of university programs to attract and prepare appropriate students. Currently, the universities are of little help in developing first-rate analysts or modelers, whether for military applications or others. Further, many departments purporting to do so emphasize special tools (e.g., linear programming) rather than the broader problem-solving skills and attitudes we seek. We believe work environment issues are less of a problem, because the federally funded research and development centers

¹⁴Recently, however, MORS has had some minisymposia on selected topics (advanced simulation, human factors, and operational art) with smaller attendance and more attention to the quality of papers. These have been stimulating and productive.

(FFRDCs) and national laboratories have a long track record as being attractive places to work.

It is worth noting here that defense work will probably be much less attractive for graduating students in the future than in the recent past as a result of the Cold War ending. The war with Iraq will mitigate this, but only temporarily, even if the war goes well. Any actions that enhance the *technical and scientific* content of defense work will tend to counter these trends and permit a continuing supply of talent.

IV. A PLAN OF ACTION

ORGANIZATIONAL ISSUES

The most important problem underlying the current morass (as indicated in Figs. 2.1 through 2.5) is that no one is in charge—with both responsibility and authority. The need here is less for a permanent czar (since czars are seldom good for science) than for a principled organization with the charter, resources, and clout to make military science flourish—not by doing the research itself, but rather by sponsoring and facilitating it.¹ This organization, an Office of Military Science (OMS), would be in addition to the many organizations that currently exist within the Services and government agencies. The OMS would need to be protected from the parochial and adversarial interests of the military services and defense agencies, and yet be responsive to their needs. It would need to exude the ethics of scientific inquiry rather than the attitudes that have sometimes characterized model building and military analysis in the past. It would need to be prestigious, stable, and effective. It should focus exclusively on R&D. In particular, it should not sponsor potentially controversial analysis, although one would expect those conducting the research to be engaged also in analysis for other sponsors. Also, it might be appropriate for the Secretary of Defense to use the Office to convene special scientific review panels, perhaps in connection with Defense Science Board activity.

The OMS should focus its efforts, initially at least, on phenomena above the levels of engineering analysis and weapons testing. It should be concerned primarily with joint and combined effectiveness at the theater level and below, perhaps down to the level of battalions. It should be concerned with the full range of items in Fig. 3.3—i.e., development of analytic methods and tools, as well as empirical and theoretical work on the phenomenology of war and military operations more generally. It should help disseminate some of the fruits of research by the military science community (e.g., model modules, models, and data, along with commentary and information about how they have been tested and what views exist about their realm of validity).² Overall, then, the OMS would be

¹Since the draft of this Note was first issued, OSD has begun an initiative that may lead to the creation of a somewhat comparable office. As of February, 1991, there was an interim office for Modeling and Simulation located in OSD (FM&P). That office had sponsored a study to develop a draft charter and policy (the Gorman study). No final actions had yet been taken, however, and we have therefore chosen not to comment here on the interim results and recommendations.

²We are leery of using phrases such as "the OMS would 'certify' . . ." because we believe that a zealous effort to certify some models as valid and others as invalid would prove mischievous. There is a natural tendency for military organizations to seek a discipline over modeling analogous to the discipline that constitutes doctrine, but that approach is antithetical to science and might result in "blessed" but invalid models and mediocre analysts (good analysts would be uninterested in merely exercising blessed models). Further, government organizations might use

relatively more concerned with research, research models, and relatively generic methodologies and tools than with specific applications, although it should also be interested in *assisting* agencies with their near-term problems, which would buy good will and establish important contacts for information and insights.

In assessing effectiveness of models for describing phenomenology or supporting classes of analysis, the OMS should consider political-military factors. Further, it should strongly encourage research and analysis emphasizing issues of operational strategy, tactics, and command-control, not just weapon-system effectiveness. It might also sponsor cutting-edge work on decision aids for commanders and staff at all levels, but here as elsewhere it should insist on relevant empirical studies and not be satisfied with theories and flashy computer systems. Most application models, including decision support systems, should be developed by user agencies, but "research models" should include models experimenting with decision support.

In all of its work the OMS should propagate the spirit of scientific inquiry, including the associated requirements for documentation, free exchange of ideas, peer review, and development of theories and texts. It should, for example, sponsor and otherwise encourage both "comparative modeling" and "countermodeling" activities, such as have proven very useful in other domains of policy-relevant simulation (see, e.g., Greenberger, et al., 1976).

In thinking how to implement the concepts introduced here, we recognize that a variety of options are plausible. We believe, however, that a new organization is desirable because of the need to change deeply seated attitudes. We suggest, as the starting point for discussion, creating an OMS that would report to the Secretary-of-Defense through a Military Science Board (MSB) consisting of

- A three-star general officer appointed by the Chairman of the Joint Chiefs of Staff
- An assistant-director-level representative of the Director of DARPA
- A DASD-level representative of the ASD (PA&E)
- A comparably ranked representative of the Director, Defense Research and Engineering
- A comparably ranked representative of the Assistant Secretary for Force Management and Personnel (specifically, from an office charged with training responsibilities)
- Senior representatives from the Army, Air Force, Marines, and Navy.

the certification process to gain power over other government organizations and to limit competition with outside organizations.

The director of the OMS would report to MSB. The MSB would have a staff-level working group with representatives from the above organizations and other defense and intelligence agencies. A similar approach has been used by OSD's Office of Net Assessment to guide development and dissemination of the RAND Strategy Assessment System (RSAS), and the process has worked well—creating a cross-cutting community within the government with a good working relationship with the developers. In creating an OMS, it is important to establish early-on approval by the Secretary of Defense and not allow the initiative to be treated as a subject on which the separate agencies have vetoes. Further, the OMS should not be constrained in its actions by requirements for consensus. In particular, while interested participation by the Services is critical—especially since they sponsor the vast percentage of relevant work and would be its primary eventual users—models can be weapons in the resource-allocation process, and decisions about the OMS and its activities ultimately need to be made by OSD and the Joint Staff.

For a variety of reasons related to the need for a scientific approach and a mid- to long-term perspective, we suggest a co-chairmanship by the representatives of DARPA or DR&E, and the Joint Staff.

The OMS would require a significant budget, for which long-lead-time arrangements should be made immediately (e.g., by setting aside portions of the DARPA and OSD budgets for use in late 1991 and 1992). Assuming that previous modeling efforts continued—i.e., that the OMS would be funded with “new money” rather than reprogramming from other modeling projects—we estimate that after a startup year at much lower levels (e.g., \$10M), a budget of \$30M per year might be adequate, but by no means lush. This would *not* include the costs of any specialized field tests, but we assume that such tests could be requested or directed by the Chairman of the Joint Chiefs of Staff or the Secretary of Defense upon the recommendation of the oversight board. Nor would this budget cover the substantial cost of distributed war gaming conducted for training purposes, since such expenses should and will be borne by other organizations. This budget level *would* permit the DoD to support a diversity of research efforts and to create small but meaningful links to academic programs that would in time generate a supply of appropriately trained new scientists. A much larger budget would, of course, be desirable, at least if it did not come at the expense of ongoing modeling efforts by the separate agencies.

ENLISTING THE FFRDCs AND NATIONAL LABORATORIES

The DoD already has institutions that are well suited for this type of research, some of the FFRDCs and national laboratories.³ Further, some of them already conduct a great deal of related research. None of them, however, has anything so broad and systematic on their research agendas. Further, there is little cross talk among these institutions, which has reduced the quality and impact of their work from what might have been possible. We recommend that the DoD initiate immediately a limited program of cooperative research involving several of these institutions. RAND and Livermore have shown interest in the recent past and we believe there would be interest from some of the others as well.⁴

ESTABLISHING LINKS WITH THE WARRIOR PREPARATION CENTER AND JOINT WARFARE CENTER

We have been impressed by the Warrior Preparation Center's (WPC's) ability to work with commanders and respond to their requests in short periods of time. It is likely that the WPC and the Joint Warfare Center (JWC) will demonstrate a high degree of competence, innovation, and a "can do" spirit; it is also likely they will continue to have warm ties to operational and training commands. We therefore believe it is desirable to bring them into the research program suggested in this White Paper by giving them a charter for early testing of prototype models and decision aids.⁵

AN AGENDA FOR IMMEDIATE AND NEAR-TERM ACTIONS

With this background, then, we suggest the following actions:

- DARPA or the Secretary of Defense should sponsor a study to define a detailed management plan and research agenda for an OMS. This could involve a Blue-Ribbon committee, a research project, or both. There should be participants from outside the defense community (e.g., academic figures interested in defense and knowledgeable about the real-world mechanisms of scientific inquiry, especially in fields with inadequate empirical data). The research agenda should address all

³In addition, of course, the DoD has in-house offices and agencies, such as Air Force Studies and Analysis and the Army's Concepts Analysis Agency. These do some original research and regular contract research with commercial contractors in addition to their extensive analysis functions. The DoD has consistently concluded over the years, however, that it was essential to go outside the government for a large part of its advanced research and some of its analysis.

⁴We also recommend cooperative studies with non-U.S. military analysis groups in the FRG, UK, and elsewhere.

⁵This idea has been proposed and discussed at a number of workshops in the last several years, starting with the DARPA-sponsored workshop at RAND on distributed war gaming in 1988 (Bankes, forthcoming). At the same time, we believe that neither the WPC nor the JWC are appropriate places for basing scientific research, even of the applied variety of interest here, because the organizations' natural dispositions are to support here-and-now needs, cutting corners as necessary.

aspects of Fig. 3.3, including empirical tests, historical analysis, systematic interviewing, and so on. Research should draw upon primary materials of other nations, and not just the experience base of the United States.

- DARPA or the Secretary of Defense should sponsor a two-year prototype effort to create scientifically valid interfaces between high-, medium-, and low-resolution models of warfare. This would be an intensive experiment in variable resolution modeling and the sound use of families of models for mutual calibration and exploration.⁶ As part of this effort, the institutions would organize workshops and conferences with broader participation (especially within the FFRDC/national-laboratory community) to discuss selected issues in technical detail. At the technology level, the model building would place extraordinary emphasis on clarity, comprehensibility, and flexibility—even if this meant substantial penalties in performance. A major objective here would be to explore the concept of using simulation models as the repository of knowledge and a principal mechanism for communicating and debating about that knowledge, including uncertainties. Another objective would be thinking out the appropriate relationships between research and applications models.
- The OSD or Joint Staff should sponsor a related study to use physics-level (e.g., JANUS/CONMOD) models to bound or estimate key parameters used in the RSAS and other theater-level models, preferably in the context of a policy-relevant issue such as evaluating proposals for deep cuts in Central Europe or defending in post-crisis Kuwait with minimal ground forces.—This effort would have two objectives: the immediate objective of serving the needs of policymakers, and the technical objective of learning by experience how to do a sound analysis of this type, which requires cross-organizational transfer and comparison of models.⁷
- DARPA or the Secretary of Defense should initiate a journal of military research. The journal should be unclassified, although classified special editions should not be precluded.
- Internally, the Secretary of Defense should appoint a committee to define organizational mechanisms for instituting a vigorous long-term effort in military

⁶We have in mind high-resolution models, such as JANUS, CONMOD, and COSMOS (a German model at the University of the Bundeswehr; see Huber, 1990), and low- and medium-resolution models, such as the RSAS.

⁷Many studies have been conducted in which one model was allegedly calibrated from the results of a higher-resolution model. By and large, however, those "calibrations" are at best crude and sometimes downright cynical. They can also be highly misleading, because the high-resolution model is itself unreliable. Cross-model calibration and tuning is an intellectually difficult and time-consuming effort, and there is not much of a theoretical base to start from.

science. This might be stimulated by a joint memo from, for example, the Director of DARPA, the Director of Defense Research and Engineering, the Director of OSD's Net Assessment, and the Director of J-7 or J-8.

- Long-lead-time plans should be made to fence budgets for FY 1991 and thereafter. A funding stream might be \$10M in FY 1992 and a constant level thereafter at \$30M in 1991 dollars. A substantially larger budget might be necessary, depending on the scope and charter of the OMS that we postulate.

CONCLUSION

In this white paper, then, we have sought to sensitize policymakers and senior military officers to the need for a true military science within the United States and to suggest both substantive and procedural action items. We have done this without the benefit of a detailed study, drawing instead on our personal experience and the ideas that we could pull together quickly in discussion with our colleagues. Our principal recommendation is that the DoD should recognize that it is not enough to build models, or even to build and "manage" them. Nor is it feasible to achieve good results by merely imposing a military-style discipline on the undisciplined community of modelers and analysts. Nor is it feasible to "validate" models by creating a program to do so.⁸ Instead, the DoD should nurture development of a vigorous military science. Accepting this as a prime objective will shape fundamentally the DoD's approach to the building, testing, comparing, and using of models in all their varieties from closed-form analytic models through computerized simulations and war games. Currently (February, 1991), the DoD seems likely to pursue an effort to provide oversight on modeling and simulation, but it has only begun to discuss seriously the issue of military science. We hope this white paper will contribute to that discussion.

⁸As discussed by Clayton Thomas in his chapter in Hughes (1984), verification and validation are very hard even to define, much less accomplish. In our view, decisions about model acceptability will have to be made by the organizations using them, and not by higher level centralized organizations working without the benefit of context.

BIBLIOGRAPHY

- Allen, Patrick, Bart Bennett, Manuel Carillo, Bruce Goeller, and Warren Walker, *Quality in Policy Modeling*, RAND, P-7640-1, forthcoming.
- Anderson, Lowell B., John H. Cushman, Alan L. Gropman, and Vincent P. Roske, *Simtax: A Taxonomy for Warfare Simulation*, a draft paper developed under the auspices of the Joint Staff's J-8 division, 1989.
- Bankes, Steven, *Issues in Developing the Potential of Distributed Warfare Simulation*, RAND, R-3746, forthcoming.
- Battilega, John and Judith K. Grange (eds.), *The Military Applications of Modeling*, Air Force Institute of Technology, 1978.
- Bennett, Bruce, Arthur Bullock, Carl Jones, and Paul K. Davis, *Main Theater Warfare Modeling in the RAND Strategy Assessment System (3.0)*, RAND, N-2743-NA, 1988.
- Brewer, Gary and Martin Shubik, *The War Game: A Critique of Military Problem Solving*, Harvard University Press, 1979.
- Brooks, Frederick P., *The Mythical Man Month: Essays on Software Engineering*, Addison-Wesley, Reading, Massachusetts, 1982.
- Builder, Carl, *Masks of War*, The Johns Hopkins University Press, Baltimore, Maryland, 1989.
- Callahan, Leslie G. (ed.), *Modeling and Simulation of Land Combat*, proceedings of a workshop held at Callaway Gardens, Georgia, 28-31 March, 1982, Georgia Institute of Technology, Atlanta, Georgia, 1982.
- Chiu, Yu-Hsing, et al., *The Conflict Model (CONMOD) Interim Technical Overview*, Lawrence Livermore National Laboratories, UCID 21234, 1987.
- Comptroller General of the U.S. General Accounting Office (1980), *Models, Data, and War: A Critique of the Foundation for Defense Analyses*.
- Davis, Paul K., *Game-Structured Analysis as a Framework for Defense Planning*, RAND N-2751-RC, 1988. Reprinted from Reiner K. Huber (ed.), *Modeling and Analysis of Conventional Defense in Europe*, Plenum Publishing Corp., 1986.
- Davis, Paul K., *The Role of Uncertainty in Assessing the NATO-Pact Central Region Balance*, RAND, N-2839-RC, 1988a.
- Davis, Paul K., *Toward a Conceptual Framework for Operational Arms Control in Europe's Central Region*, RAND, R-3704-USDP, 1988b.
- Davis, Paul K., *Modeling of Soft Factors in the RAND Strategy Assessment System (RSAS)*, RAND, P-7538, 1989.

- Davis, Paul K., *An Analyst's Primer to the RAND-ABEL Programming Language*, RAND, N-3042-NA, 1990a.
- Davis, Paul K., *Central Region Stability in a Deep-Cuts Regime*, RAND, P-7610, 1989. Published also as a chapter in Reiner Huber (ed.), *Conventional Stability in Europe: Prerequisites and Analytic Requirements*, Nomos Verlagsgesellschaft, Baden-Baden, 1990b.
- Davis, Paul K. and H. Edward Hall, *Overview of System Software in the RAND Strategy Assessment System*, RAND, N-2755-NA, 1988.
- Davis, Paul K. and Robert D. Howe, *Planning for Long-Term Security in Central Europe: Implications of the New Strategic Environment*, R-3977-USDP, 1990a.
- Davis, Paul K. and Robert D. Howe, *Representing Operational Strategies and Command-Control in the RAND Strategy Assessment System (RSAS)*, RAND, N-3232-OSD, 1990b.
- Davis, Paul K., Robert Howe, Richard Kugler, and William Wild, *Variables Affecting Central-Region Stability: The 'Operational Minimum' and other Issues at Low Force Levels*, RAND, N-2976-USDP, 1989.
- Davis, Paul K. and James A. Winnefeld, *The RAND Strategy Assessment Center*, RAND, R-3535, 1983.
- Defense Science Board, *Computer Applications to Training and Wargaming*, 1988.
- Dupuy, Trevor N., *Understanding War: History and Theory of Combat*, Paragon House Publishers, New York, 1987.
- Fain, Janice B., Richard C. Anderson, Trevor N. Dupuy, Gay M. Hammerman, and Charles F. Hawkins, *Forced Changes of Combat Posture*, a HERO report by Data Memory Systems, Inc., September 30, 1988. Prepared for U.S. Army Concepts Analysis Agency.
- Greenberger, Martin, Matthew A. Crenson, and Brian L. Crissey, *Models in the Policy Process*, Russell Sage Foundation, New York, 1976.
- Hardison, David C., et al., *Review of Army Analysis*, Vol. I., Headquarters, Department of the Army, April 1979.
- Helmbold, Robert L., *Point Paper: CAA History Activities, 1980-1990*, U.S. Army Concepts Analysis Agency, Bethesda, Md., 1990.
- Huber, Reiner (ed.) *Modeling and Analysis of Conventional Defense in Europe: Assessment of Improvement Options*, Plenum Press, New York, 1986.
- Huber, Reiner (ed.), *Military Stability: Prerequisites and Analysis Requirements for Conventional Stability in Europe*, Nomos Verlagsgesellschaft, Baden-Baden, 1990.
- Hughes, Wayne P. (ed.), *Military Modeling*, Military Operations Research Society, Alexandria, Va., 1984 (revised, 1989).

- Hughes, Wayne P., *Fleet Tactics: Theory and Practice*, U.S. Naval Institute, Annapolis, Maryland, 1986.
- Kugler, Richard, *NATO's Future Conventional Defense Strategy in Central Europe: Theater Employment Doctrine for the Post-Cold-War Era*, forthcoming.
- Low, Larry, *Theater Level Gaming and Analysis Workshop for Force Planning*, SRI International, Vols. I and II, 1977 and 1981, respectively.
- McQuie, Robert, *Historical Characteristics of Combat for Wargames (Benchmarks)*, U.S. Army Concepts Analysis Agency, Research Paper CAA-RP-87-2, July, 1988.
- Military Operations Society (MORS), *Human Behavior and Performance as Essential Ingredients in Realistic Modeling of Combat—MORIMOC II, Vols. 1-2*, proceedings of a minisymposium held 22-24 February, 1989, Military Operations Research Society, 1989. Steve Murtaugh, ed.
- Military Operations Society (MORS), MORS Mini-symposium at National Defense University, "Operational Art and Analysis (OPARTAN)," held March 6-9, 1990, forthcoming.
- Military Operations Society (MORS), *MORS Workshop on Simulation Technology 1997 (SIMTECH 97)*, May (Session I) and October (Session II), Edward Brady, Chairman, 1988.
- Military Operations Society (MORS), *MORS Workshop on Simulation Technology 1997 (Simtech 1997), Draft proceedings for Session III*, June, Edward Brady, Chairman, 1989.
- MITRE, *Advanced Technologies Workshop, August 1989 (Draft)*, Stuart Starr, Chairman, 1989.
- Molnar, Alexander and Jerry Colyer, *Historical Notes on Interdiction Operations*, The Warrior Preparation Center, December 1988.
- Quade, E.S., and W.I. Boucher (eds.), *Systems Analysis and Policy Planning: Applications in Defense*, American Elsevier Publishing Co., New York, 1968.
- Rothenberg, Jeff, *The Nature of Modeling*, RAND, N-3027-DARPA, 1989.
- Rowland, David, "Assessments of Combat Degradation," *RUSI*, pp. 33-43, June, 1986.
- Shlapak, David, and Paul K. Davis, *How Little Is Enough: Estimating Post-Crisis Force Requirements in the Persian Gulf*, RAND, N-3314, forthcoming.
- Simpkin, Richard, *Red Armour*, Brassey's, London, 1984.
- Stockfish, J.A., *Models, Data, and War: A Critique of the Study of Conventional Forces*, RAND, R-1526-PR, 1975.
- Thompson, Michael E., *Political and Military Components of Air Force Doctrine in the Federal Republic of Germany and Their Implications for NATO Defense Policy*

Analysis, RAND, P-7557-RGS, a PhD Dissertation in the RAND Graduate School, 1987.

Winnefeld, James and David A. Shlapak, *The Challenges of Future Nonstandard Contingencies: Implications for Strategy Planning, Crisis Management, and Forces*, RAND, N-3098-DAG, 1990.

Zeigler, Bernard P., *Multifaceted Modelling and Discrete Event Simulation*, Academic Press, Orlando, Florida, 1984.

Zeigler, Bernard P., *Object-Oriented Simulation with Hierarchical, Modular Models*, Academic Press, New York, 1990.