Special Issue Article

JDMS

Journal of Defense Modeling and Simulation: Applications, Methodology, Technology
1–16

DOI: 10.1177/15485129211073126
journals.sagepub.com/home/dms
SAGE

Artificial intelligence for wargaming and modeling

Paul K Davis[1] and Paul Bracken[2]

Abstract

In this paper, we discuss how artificial intelligence (AI) could be used in political-military modeling, simulation, and wargaming of conflicts with nations having weapons of mass destruction and other high-end capabilities involving space, cyberspace, and long-range precision weapons. AI should help participants in wargames, and agents in simulations, to understand possible perspectives, perceptions, and calculations of adversaries who are operating with uncertainties and misimpressions. The content of AI should recognize the risks of escalation leading to catastrophe with no winner but also the possibility of outcomes with meaningful winners and losers. We discuss implications for the design and development of *families* of models, simulations, and wargames using several types of AI functionality. We also discuss decision aids for wargaming, with and without AI, informed by theory and exploratory work using simulation, history, and earlier wargaming.

Keywords

Artificial intelligence, wargaming, modeling and simulation, cognitive modeling, decision-making, decision-making under deep uncertainty, massive scenario generation, exploratory analysis and modeling

1. Introduction

In this paper, we argue that (1) modeling, simulation, and wargaming (MSG) are related methods of inquiry that should be used together, that (2) artificial intelligence (AI) can contribute to each, and that (3) AI for wargaming should be informed by modeling and simulation (M&S) while AI for M&S should be informed by wargaming. We sketch an approach, focusing for brevity on political-military MSG involving states with weapons of mass destruction (WMD) and other high-end weapons. Section 2 provides our view of how MSG and analysis can relate to each other. Section 3 shows that this is feasible by discussing a 1980s system. Section 4 notes today's challenges and opportunities. Section 5 sketches aspects of architecture. Section 6 highlights the selected challenges in developing AI models and decision aids. Section 7 draws conclusions. Throughout the paper, we use ''model'' to cover the range from simple math formulas or logic tables to complicated computational models; we use ''wargame'' to include everything from small seminar exercises (e.g. Day-After exercises[1]) to large multi-day, multi-team wargames.

2. An integrated view of modeling, simulation, gaming, and analysis

MSG can be used for a wide range of functions, such as those in Table 1. *Each* function can be addressed by *each* MSG element, although relatively simple human activities such as seminar wargames and *Day-After Exercises* have proven uniquely valuable for the last two items.

The usual forms of M&S and wargaming have different strengths and weaknesses[2–6] as stereotyped in the first three columns of Table 2. M&S is seen as quantitative, rigorous, and ''authoritative,'' but severely limited for failure to reflect *human* considerations. Critics of M&S go farther, arguing that the ''rigor'' of M&S translates into generating

[1]RAND Corporation and Pardee RAND Graduate School, USA
[2]Yale University, USA

Corresponding author:
Paul K Davis, RAND Corporation and Pardee RAND Graduate School, 1776 Main Street, Santa Monica, CA 90407-21391, USA.
Email: pdavis@rand.org

Table 1. A few of the many purposes of MSG.

Functions of gaming and modeling	Comments
Familiarization	Learn about the "chess board," actors, and processes.
Education for adaptive planning	Learn about "the system," linkages, and cause–effect relationship so as to be better prepared for adaptations in the real world when events occur.
Testing	Test coherence of and identify vulnerabilities in a pre-existing plan.
Exploratory analysis	Confront deep uncertainty. Explore possible scenarios to open minds, identify possible problems, note opportunities, and design robust strategies.
Concept generation	Unleash creative thinking about how to employ a new technology; or given an operational challenge, how to address it.
Tightening and rehearsal	Given a plan, tighten details, and rehearse participants for implementation.
Sensitization	Make issues intuitive, compelling, and demanding of action.
Communication and socialization	Communicate rationale and purposes of strategy; develop social networks important to implementation or cooperation in crisis.

Table 2. Currently perceived and future differences among modeling and gaming.

Attribute	As often perceived		Potential
	Modeling	Gaming	Analytic wargaming
Quantitative versus qualitative	Quantitative	Qualitative	Qualitative and quantitative
Rigorous and reproducible	Yes, but perhaps precisely wrong	No	As appropriate
Authoritative	Yes, with blessed models and data (which may be quite wrong)	No	As appropriate, for specific purposes
Scope	Kinetic	Up to full PMESII	Up to full PMESII
Character	Sometimes opaque	Explainable in discussion	Understandable with explanation capability
Creative, forward-looking	No	Yes	Yes
Adaptive	No	Yes	Yes
Able to address human issues and foibles	No	Yes	Yes
Interesting, compelling, good for team building	No	Yes	Yes
Man in loop	No	Yes	Yes (optional)
Representation of top decision-makers	No	Yes (optional)	Yes (optional)
Clear and persuasive in communication	No	Yes (experiential learning)	Yes (experiential learning)

PMESII: political, military, economic, social, information, and infrastructure.

results that may be precise, but wrong. Wargaming, in their view, corrects for the shortcomings of M&S. M&S advocates have a different view.

We assuredly recognize and have long criticized the shortcomings of normal modeling. We have also benefited greatly from wargaming, in part through long associations with Herman Kahn (P.B.), RAND, and Andrew Marshall, but the quality of wargames ranges from being a waste of time or even counterproductive to being a rich source of insights. Although such insights cannot be trusted without follow-up study, that is true also of insights from modeling.[7]

A thesis of our paper is that the stereotype need not be correct and that the aspiration (unabashedly lofty) should be for the last column of Table 2—''having it all'' with a semi-integrated mix of modeling, simulation, and gaming. Figure 1 shows a corresponding vision.

This idealized activity over time begins (Item 1) by assembling knowledge about a domain (e.g. international security issues for the India-Pacific region) from studies, war games, military and diplomatic experience, human history, anthropology, and so on. Metaphorically, this is characterizing the chessboard, actors, potential strategies, and rule book.

Two efforts proceed asynchronously. As in the upper half of Figure 1, wargaming proceeds, structured for some purpose. This may occur independently, whether or not

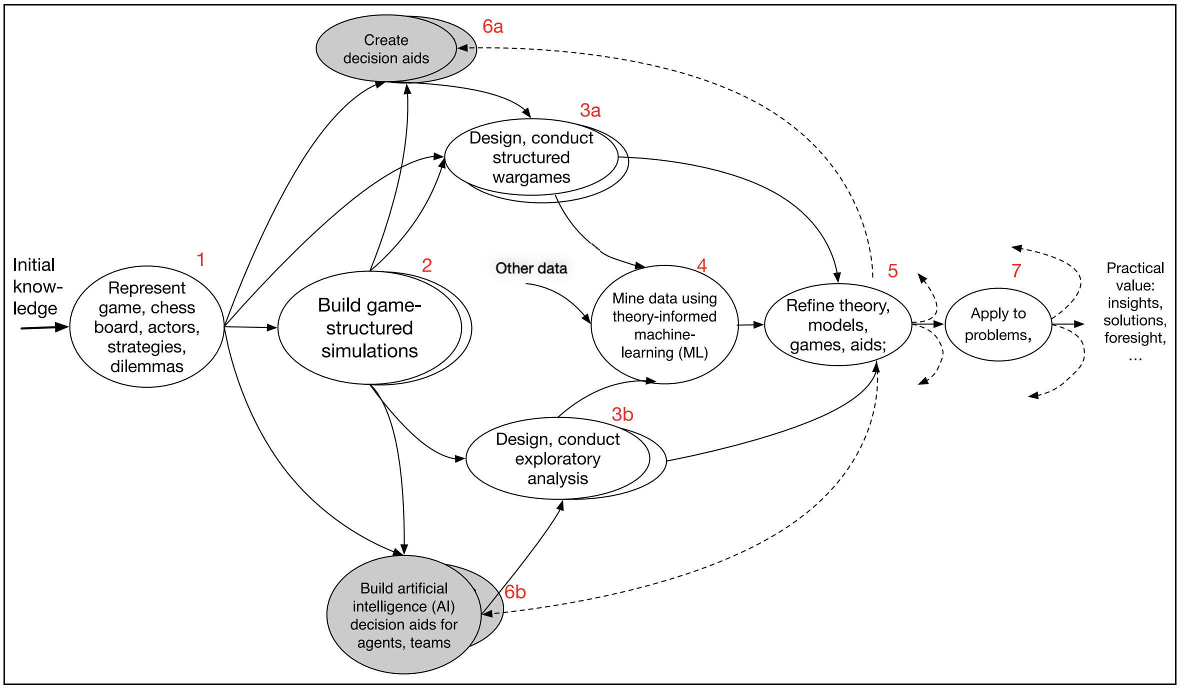

Figure 1. Connecting M&S, wargaming, and analysis.

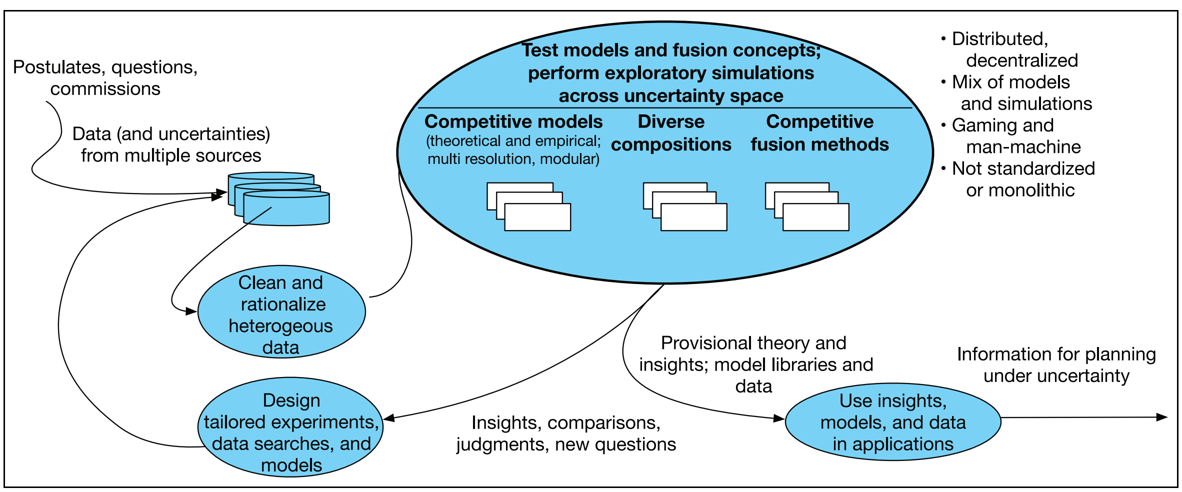

Figure 2. A virtual social-behavioral modeling laboratory (SBML) (taken from a recent study[8]).

the rest of the diagram is successfully executed. In parallel, M&S proceeds in the form of game-structured simulation. Over time, the lessons from M&S and wargaming are assimilated using AI to mine data from M&S experiments (Item 4), so as to refine theory and data for subsequent cycles (Item 5). At any given time, problem-tailored MSG addresses real-world problems (Item 7). As in the light-gray bubbles, decision aids for human teams (Item 6a) and heuristic rules for agents (Item 6b) are generated and updated. Some are constructed directly, but others distill knowledge from analytic experiments and wargaming. Some agents incorporate AI directly, some indirectly, and some not at all. Figure 1 encourages coordination among MSG activities, although the coordination may sometimes be informal and might occur only occasionally.

The intent of Figure 1 could be accomplished in a single organization (e.g. for sensitive in-government work) and/or a more open continuing program of efforts in think tanks, laboratories, private industry, academia, and government, as in Figure 2—in what a DARPA study called a social-behavioral modeling laboratory (SBML).[8] In either case, the approach would encourage diversity, debate, and competition. It would also encourage *composing* purpose-built MSG components using community modules. This contrasts with focusing on one or a few blessed, monolithic models. Bluntly, the vision is revolutionary.

3. Existence proof

An inspiration for the vision of Figure 1 was the RAND Strategy Assessment System (RSAS) of the 1980s (Appendix 1 points to documentation). In response to a DoD request for better use of wargaming for strategic analysis, an RAND team led by Carl Builder proposed *automated wargaming* that would exploit that era's AI, expert systems, but that would allow interchangeable AI models and human teams.[9] That led to a multi-year project, which one of us (P.K.D.) led after joining RAND in 1981.

The project began with an in-depth design, which retained the seminal idea of interchangeable teams and AI agents, but also included a flexible global military model; new AI-related concepts such as *alternative* Red and Blue agents, each with models of each other; a Green Agent representing other parties with simple parameterized rule-based submodels; the capacity for Red and Blue agents to do ''look-aheads'' before making decisions; and ''analytic war plans''—adaptive slotted-script AI models representing military commanders. The design also anticipated: multiscenario analysis, incorporating ''soft factors'' such as qualitative fighting effectiveness, and explanation capability for the AI models. Figure 3 sketches the high-level RSAS architecture.[10] Implementation proceeded throughout the 1980s. RAND used the RSAS for DoD studies, for example, of the conventional balance in Europe and proposals for conventional arms control,[11,12] and exported it to various government offices and war colleges. The Joint

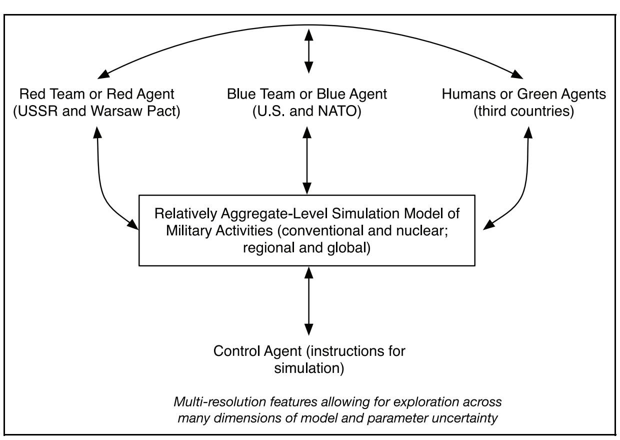

Figure 3. RSAS architecture.

Staff received the RSAS, but continuity proved impractical because as soon as appropriately talented officers learned to use it, they were promoted to other assignments.

Despite its technical successes, the RSAS was ahead of its time in some ways. On one hand, its innovative global combat model was widely embraced and used for both analysis and joint wargaming. It became the Joint Integrated Combat Model (JICM), which has evolved over the last 30 years and is still used. On the other hand, the AI portions of the RSAS were seldom used outside RAND except for demonstrations. Most government offices steering the RSAS work had no interest in political-level matters such as crisis-decision-making, paths to war, or escalation. A few were, which led to RAND studies,[13–15] but for the most part their needs could be addressed by relatively simple wargames, including the *Day-After Exercises* (Roger Molander, Peter Wilson).[16] Furthermore, the full RSAS was expensive, complicated, and demanding. More generally, DoD interest in wargaming plummeted with the collapse of the Soviet Union.

Fortunately, cream-skimming proved possible: *some* of the important insights from RSAS-like simulations with AI agents can be obtained with very simple models and games,[17–19] as illustrated recently in unpublished work gaming nuclear war with alternative images of the adversary.[20]

The RSAS incorporated to some degree most of the ideas of Table 2's last column, so it demonstrated feasibility. That is, it serves as an existence proof of sorts. That, however, was during the Cold War with 1980s technology. What could be done today?

4. Challenges and opportunities

4.1. National security challenges

Today's international security challenges go well beyond those of the Cold War.[21] They cry out for fresh wargaming[22] and fresh M&S. The new challenges include the following.

4.1.1. Multi-polarity and proliferation. The world now has multiple decision-making centers whose actions are interdependent. Conceptually, this places us in the world of *n*-person game theory. Unfortunately, it seems that the convoluted solution concepts for *n*-person games have not yet proven terribly useful,[23] although such phenomena as the tragedy of the commons and the diner's dilemma can be described in the language of *n*-person game theory, and mean-field theory can sometimes be useful as an approximation. For various reasons, such solutions have not been widely adopted. Business-school strategy courses rarely use these techniques and defense think tanks seldom incorporate them in their M&S. It may be that real-world multipolarity is simply too complicated to model, although a few efforts have been made relating to strategic stability.[24] As with even the three-body problem in physics, behavior in *n*-party systems can even be chaotic (as noted in Chinese science fiction by Liu[25]). We also note how small a role randomized mixed strategies usually play in *n*-person games. Again, there may be so much inherent complexity in calculating other players' moves that an added layer of uncertainty arising from randomization would contribute little to our understanding of future crisis dynamics.

More states have WMD than in the 1980s (i.e. India, Pakistan, North Korea) and even more have weapons of mass *disruption*. The addition of cyber as a *strategic* weapon further complicates matters. Here, AI might be helpful in understanding events. As an example, suppose a nuclear force is attacked by bringing down the electrical power system it uses for its electronic control (that may not be easy because of dispersal and defenses). A missile force can only perform for a short time on backup power systems. The major powers are surely aware of this vulnerability for themselves and their rivals. In the commercial electrical power world, AI is becoming important for rapid reallocation of power sources to demand nodes after disruptions, such as occurred in Texas with state-wide freezing temperatures in 2021.

4.1.2. Multi-dimensional warfare. Changes of weaponry have extended the dimensionality of high-end crisis and conflict, as with long-range precision strike and new forms of cyberwar, information warfare, and space warfare. This means that the 44-rung escalation ladder introduced long ago by Kahn[26,27] must now be replaced by something more complex, as discussed later in section 6.3.

4.1.3. Feasibility of limited strategic war. A corollary is under-appreciated, that *the world is now more ripe than earlier for limited high-end warfare* in which—despite assertions to the contrary by the more ardent enthusiasts for deterrence theory—there may be meaningful winners and losers. This becomes evident when considering possibilities such as a Russian invasion of Baltic states, a North Korean invasion of South Korea, or Chinese aggression against Taiwan. Some of the issues that arise include Russia's dalliance with a strategy of escalate-to-deescalate (a Russian version of NATO's Cold War strategy)[28] and prospects for cyberwar and attacks on space systems. So also, it is troublesome to observe more states deploying precision strike weapons with transoceanic range. Even protracted "limited" strategic war may now be possible, although escalation could easily occur as discussed in section 6.3.

4.1.4. Conflicting objectives among allies and partners. Today's US security partners have different vital interests and perceptions. The remarkable unity demonstrated by NATO throughout the Cold War might not be reproduced in a modern crisis or conflict. In the Asia-Pacific region, the ambivalent relationships among North and South Korea, China, Japan, Taiwan, India, and Pakistan constitute an omen of difficulties in crisis.[29] All of these nations have escalation options through use of space, cyberspace, or regional-range precision weapons.

The overall issue here is that alliances are still *very* important, but today's alliances are likely to be different than the taut blocks of the Cold War. We may be entering a phase of multipolarity akin to that of the early 20th century. One factor in the outbreak of World War I was Berlin's belief that London would not join France in a war to block Germany in Europe. This led to the belief that war would resemble the Franco-Prussian War of 1871—limited, short, and not particularly destructive. Even France was uncertain until August 1914 over whether Britain would join in. Such calculations as to what one's allies will do are critical to stability. Uncertainty here is truly a strategic problem of enormous importance.

4.2. Technological changes and opportunities

New technological opportunities abound when contemplating the prospect of modern analytic wargaming. The following sections list a few.

4.2.1. Agent-based modeling. Agent-based modeling (ABM) has progressed greatly and is especially important for *generative* modeling that provides a cause–effect understanding of how phenomena unfold. Such generative modeling is a revolutionary development of modern science.[30,31] Unlike the agents of earlier expert systems, today's agents are typically goal-seeking or position-improving in nature, which may allow them to be more adaptive.

4.2.2. AI. AI research more generally is, of course, far broader than ABM. It provides limitless possibilities as laid out in modern texts.[32] We do not discuss it much in this paper, but in contemplating the future of M&S, and of decision aids for wargaming, it would be desirable to have lengthy sections on each of the types of AI sometimes identified, that is, reactive machines, machines with limited memory, finite automata, machines with their own theory of mind, and machines with self-awareness. That is not possible here, a limitation that will perhaps be remedied by subsequent authors.

4.2.3. Networking. Networking is now a core feature of modern life with global connections among people, organizations, and even refrigerators. Data are ubiquitous.[33] One aspect of this is distributed wargaming and exercising. Another is online gaming, even to the extent of massively parallel recreational games, study of which may yield national-security insights. Such games are not intended to be ‘‘serious,’’ but behaviors observed in them may suggest possibilities and inclinations that would not be recognized in more academic study.[34]

4.2.4. Modularity and purpose-built model composition. It now makes sense to build models as independently useful (i.e. as modules) and to *compose* more complex structures as needed for the problem at hand. Such composition contrasts with DoD's historical preference for standardized large, integrated, monolithic models. Such standardization is much less attractive where uncertainties and disagreements are common as in higher-level M&S or wargaming. Modular designs permit carrying along alternative conceptions of what is being modeled. This can open minds, which is useful for foresight, as with avoiding surprise or preparing for adaptation. It is also possible to routinely compare alternative models to data, in part for routine updating as suggested in Figure 2. Also, modular development facilitates inserting specializations for a particular problem, an approach recommended by the community of modelers and analysts in a mid-2000s DoD workshop.[35]

4.2.5. Data-driven AI/machine learning. The term AI is commonly used today to mean machine learning (ML), which is only one version of AI. ML has advanced substantially and ML models can often be accurate in fitting past data and finding otherwise unrecognized relationships. A review describes progress but also notes limitations—suggesting theory-informed versions of ML that would be more effective in future-oriented work and highlighting what is called adversarial AI, which includes tactics to defeat the opponent's deep-learning algorithms.[36]

4.2.6. Decision-making under deep uncertainty. Fundamental advances have occurred in concepts and technology for planning, discussed under the rubric of *decision-making under deep uncertainty* (DMDU). This moves away from efforts to ‘‘optimize’’ for best-estimate assumptions toward strategies expected to do well across a broad range of possible futures, that is, across many uncertain assumptions. Whereas addressing uncertainty was often paralytic in the past, it need not be so today. These insights and methods have a long history in defense planning[37–39] and social policy analysis[40,41] and should be incorporated in AI and decision aids.

4.2.7. Designing "always-on" systems with increasing intelligence. Technically, most DoD MSGs have been what the AI community calls ‘‘transformational.’’ The model or game has a starting point; it runs and then reports winner and loser. Multiple runs can be made and results aggregated to capture the variance inherent in complex dynamics. Newer AI models are designed differently, modeling systems that are ‘‘always on.’’ This is called *reactive* programming, as distinct from *transformational* programming. These systems never stop and do not just transform input data to output data. Examples include elevator systems and computer operating systems. Defense examples include a cyber warning system, a missile warning system, or an operations center. None of these go ‘‘off.’’ Defense systems are becoming more reactive, and so must the models to represent them. This was foreseen by design of the higher-level Red and Blue Agents of the 1980s RSAS, which would ‘‘wake up’’ after events and assess the situation and options freshly, rather than continuing to follow a script.[14]

In transformational models, events in the environment may trigger the program to take an action sequentially. Reactive models are different. Programs make concurrent changes in the environment. They change together, or nearly together. One interesting example for defense work involves autonomous weapons. The line between human and machine decisions has blurred because the interactions among people and machines in a reactive system may be continuous and intertwined. Reactive systems are a main thrust of US, Chinese, and Russian defense investments. How will drone swarms and cyber warning systems be represented in M&S and wargames? Unless the representations are apt, the value of related AI models in simulation might be counterproductive.

This, however, is only the beginning. How will AI change as machines have better memories and exploit what they have learned, and as they incorporate theories of the world, including theories of adversary mind?[42] One worry is that increased use of AI will increase the prospect of rapid escalation, as discussed by Yuna Wong and colleagues.[42] The risk of this is especially high for AI that focuses on maximizing some relative quantitative measure, rather than more absolute outcomes and the qualitative evaluation thereof. As an example from Cold War experience, analysis that obsessed on who would ‘‘win’’ a global nuclear war by emerging with a superior post-exchange ratio of nuclear weapons was dangerous.[15] Fortunately, decision-makers understood that outcomes would be catastrophic with no meaningful victor. Even the computer Joshua in the 1983 movie *Wargames* was wise enough to conclude ‘‘Nuclear war. A strange game. The only winning move is not to play. How about a nice game of chess?’’ Whatever AI Joshua embodied, it was more than ML about how to win a recreational game by the numbers.

5. Toward architecture

5.1. Functions sought

Developing a full architecture for modern analytic wargaming is beyond the scope of this paper, but suggesting

Table 3. Features of possible architecture.

Feature	RSAS	Suggested for future
Structure	Game structured with Red, Blue, and Green	As earlier, but with more major agents
Resolution	Aggregate, multi-resolution (MRM)	Family of models of varied character and resolution; tools to allow them to be consistent; tools to create higher-level (lower-resolution) models upon demand for particular contexts
Time	Hybrid: variable time steps but event-based for agent actions	As earlier
Man–machine	Interchangeable agents and teams; interruptible physical simulation	As earlier
Representation on human decision-making	Yes	Yes, but richer with enough agent detail to generate emergent phenomena
Uncertainty	Significant treatment of both model and parametric uncertainty	As earlier but with more extensive and advanced treatment of model uncertainty (e.g. agent behavior)
Exploratory analysis	Multiscenario analysis for tens of scenarios	Exploratory analysis based on massive scenario generation (e.g. 10^4–10^6 cases) and modern tools for displaying and analyzing results (e.g. tools from of robust decision-making (RDM)[41] such as the EMA Workbench[43])
Graphics	Hand-made with 1980s technology	Drastically updated to exploit modern tools such as *Tableau* and others
Language	Hybrid of C for simulation and high-level understandable language (RAND-ABEL) for agents. Complex interface between them	Unclear. How valuable is it for agent models to be readable and editable by non-programmers, as in the RSAS?
Agent-based models	Custom built in RAND-ABEL	Built with modern and tools for agent-based modeling
Overall Simulation	Campaign model (somewhat modular when used by experts who could adjust code) with agent direction as in Figure 3	Much more emphasis on modularity and simplified interfaces
Simulation self-awareness	No	Yes, for example, allowing the simulation to change its structure over time, recognizing new entities and agents, and exhibiting emergent phenomena[44]
Standardization?	Yes, but with extensive features for uncertainty analysis and exploration	As earlier, but with even more emphasis on modularity and easy composition
Sources of data	Eclectic, manual	As earlier, but also with mechanisms for automated updating based on either real-world or synthetic data
Frequency of wargaming in mix	Occasional	Frequent, with wargaming at various scales being common in normal work
Managerial concept	Individual installations	Local and networked as for distributed gaming or virtual meetings
Grand managerial concept	Individual	Mix of centralized organization for sensitive work, a more open social-behavioral modeling laboratory (SBML) akin to Figure 2, and “decision laboratories” used to educate high-level executives and military officers.[45,46] AI in such laboratories *must* be comprehensible to be useful

RSAS: RAND strategy assessment system; RDM: robust decision-making; EMA: exploratory modeling and analysis; AI: artificial intelligence.

some directions is possible. Figure 4 sketches a top-level architecture and Table 3 suggests various features in more detail. Figure 4 recognizes the need to address current-era crisis and conflict with in-depth attention to *at least* three major actors when contemplating many possible crises and conflicts. One example might be North Korea, South Korea, the United States, and China. Figure 4 also calls for a modular approach to the military simulation.

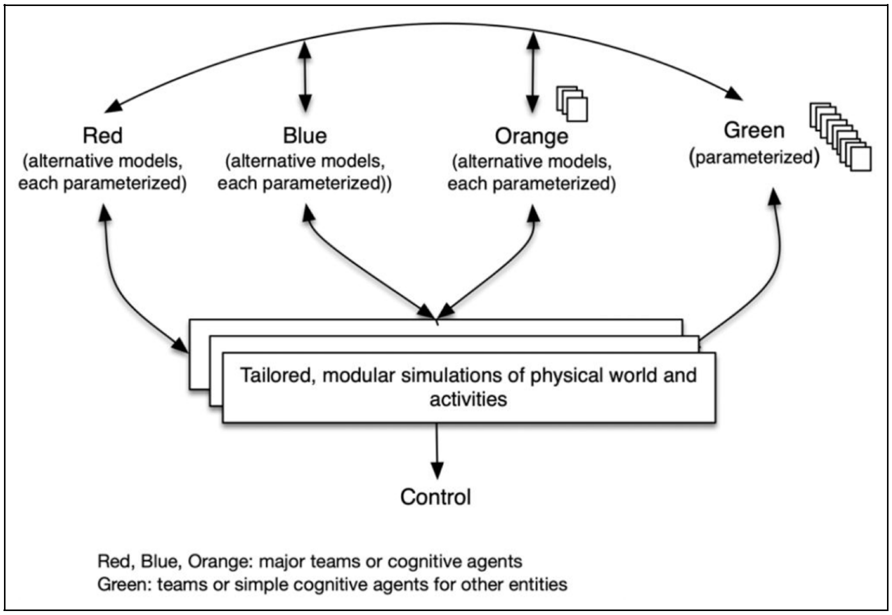

Figure 4. *N*-party game-structured simulation.

As Table 3 indicates, some features of the 1980s RSAS might carry over with modernized versions. Many other features, however, should be quite different. We see Table 3 as an opener for discussion, not an endpoint.

5.2. Limitations of exploratory analysis amidst massive uncertainty and disagreement

Since preparing for massive scenario generation, exploratory analysis, and decision-making under uncertainty is prominent in our discussion, two significant issues need highlighting:

- Exploratory analysis across parameter values is useful only if the simulation is structurally valid (i.e. only if the model itself is valid).[47]
- Drawing conclusions from exploratory analysis can be problematic when the cases (scenarios) examined are not equally probable, their probabilities are correlated, but no good basis exists for assigning probability distributions.

5.2.1. Model validation. As discussed elsewhere, model validity and data validity should be characterized separately for description, explanation, post-diction, exploration, and prediction.[47,48] Also, they must be judged with respect to the particular problem and context. Parametric methods go a long way, but *model* uncertainty has often gotten short shrift and needs more attention as discussed in a recent article.[49] Carrying along adversary models with very different objectives and values is just one example of doing so.

5.2.2. Drawing conclusions from exploratory analysis. With respect to the vexing problem of how to use exploratory analysis without knowing the relative probability of the cases, we suggest that exploratory analysis is very likely to have value for at least the purposes illustrated in Table 4, none of which require probabilities. For each example, the purpose of the exploration is to find *possibilities* (e.g. vulnerabilities or opportunities) motivating measures to prevent them, anticipate them, or prepare for related adaptations. If a critical vulnerability exists, it should be fixed, whether the probability of it being exploited ‘‘seems’’ to be low or high (if its probability was known to be vanishingly small, that would be a different matter).

6. Decision aids and AI

This section discusses *selected* issues that arise in thinking about AI and decision aids for modeling and wargaming. The first discusses decision-aid functions. The next discusses a challenge when envisioning using the ML version of AI to exploit massive scenario generation. The final section discusses one of the fundamental challenges involved in developing ‘‘cognitive AI’’ and related decision aids.

6.1. Decision aids for wargaming

6.1.1. Generic functions. If we ask about the primary function of decision aids based on what we see as important to

Table 4. Sound insights despite unknown probabilities.

Class	Meaning	Examples
Tactics and events	Non-obvious tactics or potential "exogenous" events	Asymmetric forms of escalation.Actions with ambiguous escalatory intent.Surprise attacks.
Phenomena	"Things that could happen"	Technical system failures (e.g. in C^4ISR). Death of a key leader. Shift in adversary's government.
Frictions, noise, fog of crisis, and war	Potential departures from smooth processes	Delays in own or allied decision-making.Delays in reconstructing failed networks.
Failures	Critical components (sometimes unrecognized) of systems or operations	Critical nodes in homeland facilities for global C^4ISR.Potentials for common mode failures.
Alignments	Potential changes in political and military relationships	Abdication of an ally.Emergence of a friendly or adversary coalition.Intelligence sharing of adversary states and other countries.

Table 5. Functions for decision aids.

Items for decision aids to suggest	Description	Example
Frames	Provide structure for reasoning	Escalation ladder; post-exchange power ratio
Understanding	Explain events in cause–effect terms	Misperceptions lead to understandable escalation
Alternative models of adversary and others	Recognize different possibilities for the intentions, perceptions, character, and behavior of adversaries (and others)	Adversary may be aggressive and risk-taking or fearful, stalwart, and risk-avoiding
Possible adversary actions and responses, and other events	Enemy pre-emption, abdication of ally, and types of enemy escalation	Surprise invasion of an ally; other allies refuse to act
Options for self	For restraint, response, and forward-leaning actions	Tit for tat response, proportional escalation, escalation
Factors determining option outcome	Critical factors for success	Reliability of particular allies. Continuity of networking
Factual information	Political and military situation	States of mobilization, withdrawal of diplomats, evacuation of cities, targeting of intelligence
Measures of strategy's robustness	Outcome estimates across assumptions	Measures of "regret" relating to absolute and relative losses and residual capabilities

players, rather than as exciting to AI providers, a number of key functions suggest themselves as in Table 5.[50]

From science fiction, we might expect modern decision aids for gaming to be highly computerized and informed by AI in relatively personalized form as with Isaac Asimov's robots or a less malevolent version of the computer *Hal 9000* in the movie *2001*. The authors' experience to date, however, has been that efforts to ''aid'' humans in games often prove counterproductive, obstructing the quintessentially human free-form discussion. Indeed, the efforts sometimes make the players angry because of the distractions. With that in mind, we discuss

Table 6. Value of decision aid types for different types of games.

Type Aid \ Trained?	Political-military game		Military-political game		Military game		Analytic exercise	
	No	Yes	No	Yes	No	Yes	No	Yes
Instructions and forms	H	H	H	H	H	H	H	H
Checklists and other suggestive lists (objectives, criteria, options, etc.)	H	H	H	H	H	H	H	H
Information tables	H	H	H	H	H	H	H	H
Simple diagrams (e.g. factor trees or simple casual-loop diagrams) or charts	H	H	H	H	H	H	H	H
More complex diagrams or charts	L	L	L	M	M	H	M	H
Layered detail (zooms)	L	L	L	M	M	M	M	H
Interactive model-driven displays	L	L	L	L	L	M	M	H

L, M, and H correspond to low, medium, and high.

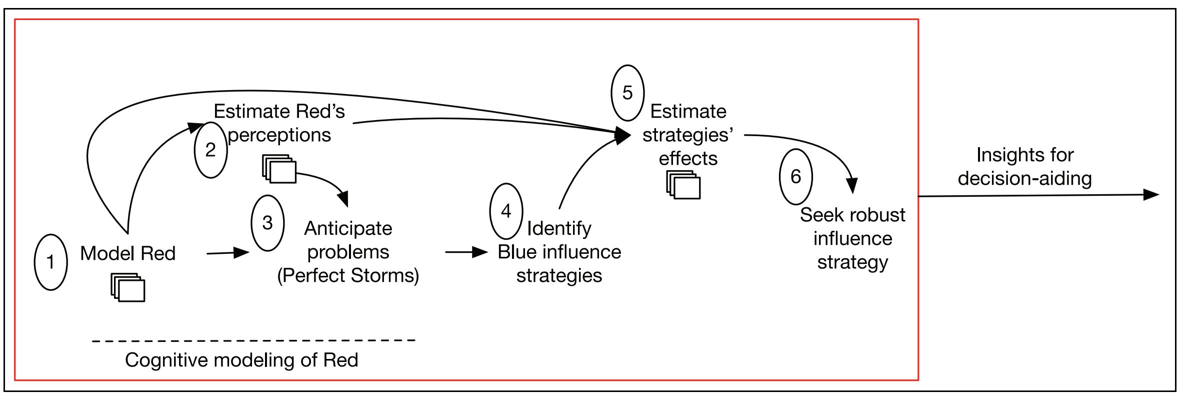

Figure 5. Process for influence exercises.

decision aids separately for practical short-term and more speculative longer-term ambitions.

6.1.2. Near-term ambitions for decision aids. Table 6 provides our subjective estimates, on a scale of low to high, for the value of the simple decision aids shown in the first column. None of these involve AI. Rather, the most valuable aids are in the character of simple viewgraphs with succinct checklists, information tables, or diagrams. The evaluation distinguishes among different types of games or exercises, and also between games in which players have or have not previously been trained with the decision aids. The evaluations were developed after some wargaming experiments conducted by RAND in collaboration with the Korea Institute for Defense Analyses.

Another data point on simple decision aids is *the Strange Game* developed (but not yet published) by RAND colleagues. It is an efficient wargame of nuclear use in which players represent a theater commander and make plays by choosing an appropriate card.[20] The game builds in decision aids that include target categories and a simple linear arithmetic for assessing what targeting option to choose.

As a final example of near-term decision aids, a recent prototype study pursued a low-tech approach to human exercises considering how to influence adversaries in crisis and conflict. The method involved a qualitative method, *uncertainty sensitive cognitive modeling* (UCM), as summarized in Figure 5.[19] The mechanisms were all qualitative, for display and discussion with a real or virtual whiteboard and interactive software. They included factor trees, alternative models of Red with representation of limited rationality, influence diagrams, and tabular comparison of strategies' apparent strengths and weaknesses. None involved AI. It was not evident that AI would even have been helpful. Perhaps that was an important insight, or perhaps it reflected inadequate imagination. Let us now turn to the longer run.

6.1.3. Longer-term ambitions for AI-enhanced decision aids. For the longer term, much more may be possible and we should look for inspiration to, for example, science fiction, electronic recreational gaming, and even the real-time discussion of emerging election results by major television networks. As mere examples of functions plausible in the not-too-distant future, in each of which an AI system responds to queries:

- A team verbally orders up an exploratory analysis of ‘‘paths to success’’ with and without the stalwart cooperation of a particular ally.
- A team asks which alternative models of an adversary continue to be plausible given recent events. The AI report reflects Bayesian-style analysis dependent on subjective likelihood functions, which have been updated to reflect recent history.
- A team contemplating a limited escalation asks about potential responses. The AI helper shows responses observed in prior wargames with players thought to represent actual decision-makers well. It also identifies conditions (as discussed in the next section) under which responses have been bad in simulation, thereby highlighting what aspects of condition need special attention to avoid disaster.

These speculations are minimal, merely to stimulate more creative thinking about how AI could be useful in decision aiding. The field is wide open, as becomes even more evident from the names given to certain types of AI that see a progression from reactive machines to those with limited memory, built-in theory of mind, and self-awareness. Some leading figures, like Pearl and Mackenzie,[51] confidently anticipate that the latter will include consciousness itself. That, however, is for the future. Pearl has characterized current robots as ‘‘as conscious as a slug.’’ That said, swarming weapons will soon be as ‘‘conscious’’ as swarms of birds, fishes, and insects.

Let us turn next to some vexing issues involving AI with M&S. They related to what AI decision aids are feasible.

6.2. Issues for ML that exploits massive scenario generation

The machine learning class of artificial intelligence (AI/ML) has the potential for finding insights by mining the results of massive scenario generation as discussed earlier. Success, however, depends on (1) the quality of the simulations and (2) the methods used to search results.

6.2.1. Are the simulations rich enough to represent essential complexity? The fruits of massive scenario generation may be useful or counterproductive depending on whether the underlying models are sufficiently rich and structurally valid for the purposes of exploration. In studying possible high-end crises, what good is a database of a million scenarios if the underlying models assume perfect rationality, perceptions, alliance relationships, and focus on, say, the post-exchange ratio of nuclear weapons as a measure of outcome? There might be value for military-technical purposes, such as force planning, but probably not for deterrence or anticipating issues in actual conflict or even serious elite wargames.[52,53] Similarly, a database of a million scenarios about a Korea conflict in 2018 would have had little value if the issues in question were sensitive to the unmodeled idiosyncratic features of such national leaders as Kim Jong Un, Donald Trump, and Xi Jinping.

Some aspects of the challenge for model-builders are known, as in recognizing the need for alternative conceptions of the decision-makers (character, personality, health),[21] recognizing the possibility of erroneous perceptions, and allowing for the kinds of non-rational decisions described by Kahneman and Tversky’s Prospect Theory[54,55] and other psychological phenomena. Addressing the challenges is difficult, to say the least, but at least the challenges are recognized.

In contrast, one of the dirty little secrets of military simulation and social-behavioral simulations more generally is that the workhouse models usually do not generate black-swan events, discontinuities, or the kinds of *emergent phenomena* that are a core element in the study of complex adaptive systems and are experienced in the real world[48] and some large and games, such as the ‘‘elite’’ high-level Cold War wargames of the 1950s.[52,53] The reasons are many but often stem from the models being ‘‘scripted,’’ rather than agent based, or—even if they do have agents—on not giving the agents sufficient diversity, degrees of freedom, and incentives to generate *realistic* adaptive behaviors, and on not allowing randomness with long-tail distributions. Doing better on such matters is a grand challenge for social-behavioral simulation generally, and for the simulations intended to connect well to realistic wargaming in particular. Some of the ingredients *are* included in sophisticated wargaming, so that one may observe, for example, disintegration of an alliance and the creation of new groupings that appear to the teams to better serve their national interests. Current-day simulations do not typically allow for such things. Speculatively, we see at least two paths for doing better. If the emergent phenomenon of interest can be anticipated (such as the alliance issues above), then appropriate objects can be built in and the simulation might recognize when to direct them to come into or out of existence. But the most important emergent phenomena (including some that appear in wargames) may *not* be anticipated. Although we do not claim to know what is *necessary*, we observe from the past experiences of complexity research, that emergent phenomena often come about because of complicated bottom-up interactions, diversity, and random events. Traditional higher-level political-military simulations, however, do not have these features. Their value is due in large part to their representing higher-level entities and processes, roughly by analogy with the models of System Dynamics. Our conclusion is that it is important, in moving ahead, to develop multi-resolution *families of models* and methods for relating them to each other. For example, a higher-resolution agent-based model might have adaptive agents for all countries involved in crisis or conflict. Simulation experiments might reveal (as can human games) the kind of emergent behaviors mentioned above, such as *occasional* dissolution of alliances, side-switching, and the popping up of new alliances of convenience. This would be ''insight'' that could then lead to adding new agents to higher-level models, agents to be activated or deactivated depending on circumstances in the simulation. This, however, would require something like the ''self-aware simulations'' discussed in a recent book on social-behavioral modeling,[56] particularly the chapter by Yilmaz[44] who envisions computation that monitors its own state and, as necessary, changes its own structure, and a chapter with debate among authors about emergence.[57]

6.2.2. Extracting insights from massive scenario generation.

If the simulations are sufficiently rich, then meaningful massive scenario generation is possible. But then what? A core challenge in exploratory analysis of simulation data is understanding how to assess the relative significance of different cases. One approach is to assign subjective probability distributions, but where does one find experts who can reliably estimate probabilities without prefacing comments such as ''Well, if tomorrow is like the past.'' Realistically, experts are not good sources for predictions or probabilities, as has been discussed in-depth by Tetlock and colleagues.[58–60]

A variant approach reports how frequently (in percentage terms) results are, for example, good or bad. This can be done with full factorial design or using the Monte Carlo sampling. Unfortunately, the tendency exists to slip into discussion of ''likelihoods'' rather than percentages, even though the cases are not equally likely. Also, for the MSG context, this type of display obscures the reality that actors are constantly looking for obscure ''corners'' of the scenario space where they will gain major advantages. Thus, a case infrequently observed in simulation may be precisely the case that develops.

The approach that we suggest is to eschew assignment of explicit probabilities, but instead to ''look for problems'' or ''look for successes.'' That is, when exploring the vast data generated from exploratory analysis, one may seek to find the *conditions* under which results are, very good, very bad, or whatever. This is called *Scenario Discovery* in the literatures on robust decision-making (RDM) and DMDU.[61]

Going farther, we urge that the AI be given hints in the form of ''aggregation fragments'' motivated from theory, simple models, and subject-area expertise.[62] An example might be ''State of readiness at the time conflict begins.'' The value for that might be the same for drastically different combinations of strategic warning time, tactical warning time, leadership characteristics, prior military readiness, and rate of mobilization. That is, the variable is an aggregation over many more microscopic initial states. Another example (assuming suitable agents) might be psychological state at the time of crisis, with values such as Paranoid, Calm and Rational, and Confidently aggressive.

Given sufficiently rich simulations and theory providing hints for the AI to use during exploratory analysis, we suspect that AI could accomplish a great deal in activities such as identifying circumstances of ''Perfect Storm''—not to predict them, but to note conditions to avoid, much as has been done in a low-tech way with simple wargaming.[19]

6.2.3. Extracting insights from massive data collection.

Another ML application could create algorithms for wargames and M&S from massive intelligence collection on adversary operations, such as those of submarines or ground-mobile missiles.[63] What once took months or years to collect and analyze may now be available in very short periods of time, generating algorithms on operational procedures that could be used in wargames or M&S. As an analogue, consider gaining insights about driving safety. The deepest insights today come from insurance companies (Progressive, GEICO) based on downloadable software that tracks individual operators: their speed, the number of left turns, acceleration patterns, and so on.[64] The data can be integrated with, for example, credit scores and other data. The result can be individualized premium rates. Such data analytics is already today's reality. There should be analogous military and MSG implications. Some, of course, will necessarily be classified and of less significance to the pol-mil focus of this paper than to other applications of MSG.

6.3. Issues for cognitive AI and related decision aids

The discussion above focused on ML-style AI, but the rich simulations that are needed must have agents that reason in more human-like ways, something that may be described as cognitive AI. In this, the decision logic uses factors and reasoning similar to what humans like to believe is the basis of their actual behaviors.

The Red and Blue agents of the 1980s RSAS were early examples. They exploited the broadly accepted escalation-ladder construct to characterize situations, options, and decision choices in nuclear crisis and conflict.

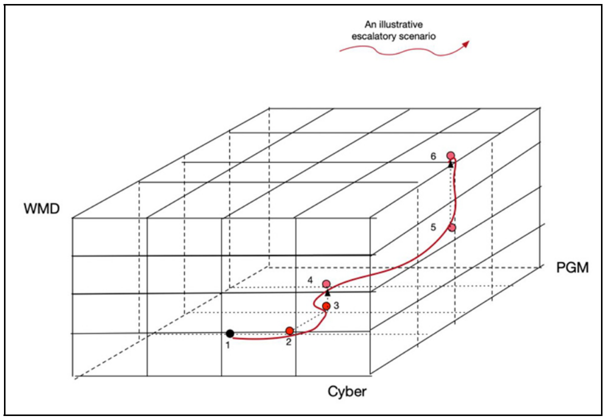

Figure 6. Illustrative escalation in a simplified hyperspace.

Today, we need a new generation of higher-level decision models, but no substitute for the escalation ladder exists. And perhaps no substitute will be found. Complexity increases greatly when going from a two-party game to even a three-party game. A replacement concept will necessarily be more complex—more like an n-dimensional lattice than a ladder—because escalation can involve not just the number of nuclear weapons and their targets, but counts, intensities, and targets relevant to cyberwar, space war, and strategic use of precision fires.

Figure 6 illustrates the notion clumsily, combining several of the dimensions so as to artificially show results in only three dimensions. It shows an illustrative scenario that starts with a tame conventional war (Item 1), but then transitions sequentially to severe cyberattack (Item 2), more extensive use of Precision guided missiles (PGMs) (Item 3), limited nuclear use (a nuclear escalation as indicated by the arrow) (Item 4), even more destructive use of PGMs (e.g. against dams and power grids) (Item 5) and perhaps a slight increase in the level of WMD (perhaps intended merely as tit for tat), and general nuclear war (Item 6). Today, however, no common understanding exists about where a particular kind of attack would appear on a given axis and whether actors would have the same assessment. Not only is the "objective" answer ephemeral at best, perceptions will likely be path dependent, nation dependent, and subject to random influences. *A central issue for planning is whether a protracted non-nuclear war between nuclear-armed near-peer states is plausible*. The issues have become even more troublesome due to the entanglement of command and control systems for conventional and nuclear warfare.[65] It seems that *predictive* models, whether AI-based or not, are not in the cards, although models generating plausible cases to worry about should be.

Many more challenges might be listed for those seeking to build cognitive AI models to represent national decision-makers in crisis, but we hope that our example has whetted appetites.

7. Conclusion and recommendations

The primary suggestion of this paper is to recommend a research agenda that sees modeling, simulation, gaming, and analysis as related and intertwined. In such an integrated view, AI for wargaming would be informed by analysis using models that included agents incorporating AI informed in part by wargaming. This will lead, for example, to agents with AI resembling the decision aids of wargaming as well as more complex algorithms. It will lead to decision aids for wargaming that will resemble the fruits of applying theory-informed ML to "data" generated by exploratory analysis from M&S exploiting AI in the form of decision agents.

With respect to AI per se, we caution against some of the practices common in today's ML. We note the absence of reliably informative empirical data on *future* crises and conflicts. Furthermore, we emphasize the need, in both decision aids and the agents used in models, for *explanation*. This suggests a preference for AI structured by cognitive modeling even if ML is used to fill out and tune that structure.

Finally, we urge great caution in what questions are asked of wargaming (including small-scale activities such as *Day-After Exercises*[66]) as well as of models. Models, simulations, games, and analysis will remain imperfect—sometimes markedly so—but it is possible to use them well to address many issues well, that is, to improve the quality of decision-making. *Anticipating possibilities* has great potential; reliable prediction does not.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

1. Millot MD, Molander RC and Wilson PA. *"The day after ..." study: nuclear proliferation in the post-cold war world: main report*. Santa Monica, CA: RAND Corporation, 1993.
2. Perla P and Curry J. *Peter Perla's the art of wargaming a guide for professionals and hobbyists*. 1st ed. Morrisville, NC: Lulu.com, 2012.
3. Pournelle P and Wong YH. MORS wargaming special meeting October 2016. *Final report, Military Operations Research Society, Arlington, VA*, 31 January 2017.
4. Caffrey MB. *On wargaming*. Newport, RI: Naval War College Press, 2017.
5. Hanley JT. Changing DoD's analysis paradigm: the science of war gaming and combat/campaign simulation. *Nav War Coll Rev* 2017; 70: 64–103.

6. Davis PK. An analysis-centric view of wargaming, modeling, simulation, and analysis. In: Tolk A, Turnitsa C and Blais C (eds) *Simulation and wargaming*. Hoboken, NJ: John Wiley & Sons, 2022, pp. 93–120.
7. Schelling TC. The role of war games and exercises. In: Carter AB, Steinbruner JD and Zraket CA (eds) *Managing nuclear operations*. Washington, DC: Brookings Institution Press, 1987, pp. 426–444.
8. Davis PK, O'Mahony A, Gulden TR, et al. *Priority challenges for social and behavioral research and its modeling* (document no. RR-2208). Santa Monica, CA: RAND Corporation, 2018.
9. Graubard MH and Builder CH. *RAND's strategic assessment center: an overview of the concept* (RAND note N-1583-DNA). Santa Monica, CA: RAND Corporation, 1980.
10. Davis PK and Winnefeld JA. *The RAND strategy assessment center*. Santa Monica, CA: RAND Corporation, 1983.
11. Davis PK. *The role of uncertainty in assessing the NATO/pact central region balance*. Santa Monica, CA: RAND Corporation, 1988.
12. Davis PK. *Toward a conceptual framework for operational arms control in Europe's central region*. Santa Monica, CA: RAND Corporation, 1988.
13. Davis PK and Stan P. *Concepts and models of escalation*. Santa Monica, CA: RAND Corporation, 1984.
14. Davis PK, Bankes SC and Kahan JP. *A new methodology for modeling national command level decisionmaking in war games and simulations*. Santa Monica, CA: RAND Corporation, 1986.
15. Davis PK. *Studying first-strike stability with knowledge-based models of human decisionmaking*. Santa Monica, CA: RAND Corporation, 1989.
16. Millot MD, Molander RC and Wilson PA. *''The day after...'' study: nuclear proliferation in the post-cold war world* (document no. IP-102-AF). Santa Monica, CA: RAND Corporation, 1993.
17. Davis PK and Arquilla J. *Deterring or coercing opponents in crisis: lessons from the war with Saddam Hussein*. Santa Monica, CA: RAND Corporation, 1991.
18. Davis PK. Synthetic cognitive modeling of adversaries for effects-based planning. In: *Proceedings of the SPIE 4716 enabling technologies for simulation science VI*, Orlando, FL, 1–5 April 2002, vol. 4716. Bellingham, WA: SPIE.
19. Davis PK, O'Mahony A, Curriden C, et al. *Influencing adversary states: quelling perfect storms*. Santa Monica, CA: RAND Corporation, 2021.
20. Worman SM, Frelinger DR, Shlapak DA, et al. *Designing a strange game: a nuclear wargame for the 21st century*. Santa Monica, CA: RAND Corporation, 2021.
21. National Research Council (NRC). *US Air Force strategic deterrence analytic capabilities: an assessment of methods, tools, and approaches for the 21st century security environment*. Washington, DC: The National Academies Press, 2014.
22. Work B and Selva GP. Revitalizing wargaming is necessary to be prepared for future wars. *War on the Rocks*, 8 December 2015, https://warontherocks.com/2015/12/revitalizing-wargaming-is-necessary-to-be-prepared-for-future-wars/
23. Shubik M. *Game theory in the social sciences—concepts and solutions*. Cambridge, MA: The MIT Press, 1982.
24. Best M and Bracken J. First-strike stability in a multipolar world. In: Huber R and Avenhaus R (eds) *International stability in a multipolar world: issues and models for analysis*. Baden-Baden: Nomos Verlagsgesellschaft, 1993, pp. 217–222.
25. Liu C. *The three-body problem*. New York: Tor Books, 2014.
26. Kahn H. *On thermonuclear war*. Princeton, NJ: Princeton University Press, 1960.
27. Kahn H. *On escalation: metaphors and scenarios*. Westport, CT: Greenwood Press, 1965.
28. Roberts B. *The case for US nuclear weapons in the 21st century*. Redwood City, CA: Stanford University Press, 2015.
29. Bracken P. *The second nuclear age: strategy, danger, and the new power politics*. New York: Times Books, 2012.
30. Epstein JM and Axtell RL. *Growing artificial societies: social science from the bottom up*. Cambridge, MA: The MIT Press, 1996.
31. Epstein JM. *Agent_zero: toward neurocognitive foundations for generative social science*. Princeton, NJ: Princeton University Press, 2014.
32. Russell S and Norvig P. *Artificial intelligence: a modern approach*. 4th ed. Hoboken, NJ: Pearson Education, 2021.
33. Pfautz J, Davis PK and O'Mahony A. Understanding and improving the human condition: a vision of the future for social-behavioral modeling. In: Davis PK, O'Mahony A and Pfautz J (eds) *Social-behavioral modeling for complex systems*. Hoboken, NJ: John Wiley & Sons, 2019, pp. 3–14.
34. Guarino S, Eusebi L, Bracken B, et al. Using sociocultural data from online gaming and game communities. In: Davis PK, O'Mahony A and Pfautz J (eds) *Social-behavioral modeling for complex systems*. Hoboken, NJ: John Wiley & Sons, 2019, pp. 407–442.
35. Davis PK and Henninger A. *Analysis, analysis practices, and implications for modeling and simulation*. Santa Monica, CA: RAND Corporation, 2007.
36. Osoba O and Davis PK. An artificial intelligence/machine learning perspective on social simulation: new data and new challenges. In: Davis PK, O'Mahony A and Pfautz J (eds) *Social-behavioral modeling for complex systems*. Hoboken, NJ: John Wiley & Sons, 2019, pp. 443–476.
37. Davis PK. Institutionalizing planning for adaptiveness. In: Davis PK (ed.) *New challenges in defense planning: rethinking how much is enough*. Santa Monica, CA: RAND Corporation, 1994, pp. 73–100.
38. Davis PK. *Analytic architecture for capabilities-based planning, mission-system analysis, and transformation*. Santa Monica, CA: RAND Corporation, 2002.
39. Davis PK. *Analysis to inform defense planning despite austerity*. Santa Monica, CA: RAND Corporation, 2014.
40. Lempert RJ, Popper SW and Bankes SC. *Shaping the next one hundred years: new methods for quantitative long-term policy analysis*. Santa Monica, CA: RAND Corporation, 2003.

41. Marchau VAWJ, Walker WE, Bloemen PJT, et al. *Decision making under deep uncertainty: from theory to practice*. Cham: Springer, 2019.
42. Wong YH, Urchak JM, Button RW, et al. *Deterrence in the age of thinking machines*. Santa Monica, CA: RAND Corporation, 2020.
43. Kwakkel JH. EMA workbench documentation: exploratory modelling and analysis (EMA) workbench, 2021, https://emaworkbench.readthedocs.io/en/latest/
44. Yilmaz L. Toward self-aware models as cognitive adaptive instruments for social and behavioral modeling. In: Davis PK, O'Mahony A and Pfautz J (eds) *Social-behavioral modeling for complex systems*. Hoboken, NJ: John Wiley & Sons, 2019, pp. 569–586.
45. Rouse WB. Human-centered design of model-based decision support for policy and investment decisions. In: Davis PK, O'Mahony A and Pfautz J (eds) *Social-behavioral modeling for complex systems*. Hoboken, NJ: John Wiley & Sons, 2019, pp. 798–808.
46. Rouse WB. *Failure management: malfunctions of technologies, organizations and society*. Oxford: Oxford University Press, 2020.
47. Thissen WA and Walker WE (eds). *Public policy analysis: new developments*. Cham: Springer, 2013.
48. Davis PK and O'Mahony A. Improving social-behavioral modeling. In: Davis PK, O'Mahony A and Pfautz J (eds) *Social-behavioral modeling for complex systems*. Hoboken, NJ: John Wiley & Sons, 2019, pp. 15–48.
49. Davis PK and Popper SW. Confronting model uncertainty in policy analysis for complex systems: what policymakers should demand. *J Policy Compl Syst* 2019; 5: 181–201.
50. Davis PK. *Decision aids for human war games and other exercises*. Santa Monica, CA: RAND Corporation, 2017.
51. Pearl J and Mackenzie D. *The book of why: the new science of cause and effect*. New York: Basic Books, 2018.
52. Pauly R. Would US leaders push the button? Wargames and the sources of nuclear restraint. *Int Security* 2018; 43: 151–192.
53. Emery JR. Moral choices without moral language: 1950s political-military wargaming at the RAND Corporation. *Texas Natl Secur Rev* 2021; 4: 11–31.
54. Kahneman D and Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* 1979; 47: 263–291.
55. Biggs AT and Pettijohn KA. Prospect theory and its implications for adversarial decision-making. *J Def Model Simul* 2021; 18: 125–134.
56. Davis PK, O'Mahony A, Pfautz J, et al. *Social-behavioral modeling for complex systems*. Hoboken, NJ: John Wiley & Sons, 2019.
57. O'Mahony A, Davis PK, Appling S, et al. Panel discussion: moving social-behavioral modeling forward. In: Davis PK, O'Mahony A and Pfautz J (eds) *Social-behavioral modeling for complex systems*. Hoboken, NJ: John Wiely & Sons, 2019, pp. 753–787.
58. Tetlock PE. *Expert political judgment: how good is it? How can we know?* (new edition). Princeton, NJ: Princeton University Press, 2017.
59. Lustick IS and Tetlock PE. The simulation manifesto: the limits of brute-force empiricism in geopolitical forecasting. *Futures Foresight Sci* 2021; 3: e64.
60. Popper SW, Lempert RJ, Davis PK, et al. Forecasts and decisions: a commentary on Lustick and Tetlock 2021. *Futures Foresight Sci* 2021; 3: e81.
61. Groves DG and Lempert RJ. A new analytic method for finding policy-relevant scenarios. *Global Environ Chang* 2007; 17: 73–85.
62. Davis PK. Lessons on decision aiding for social-behavioral modeling. In: Davis PK, O'Mahony A and Pfautz J (eds) *Social-behavioral modeling for complex systems*. Hoboken, NJ: John Wiley & Sons, 2019, pp. 899–926.
63. Bracken P. *The hunt for mobile missiles: nuclear weapons, AI, and the new arms race*. Philadelphia, PA: Foreign Policy Research Institute, 2020.
64. Sennaar K. How America's top 4 insurance companies are using machine learning, 2020, https://emerj.com/ai-sector-overviews/machine-learning-at-insurance-companies/
65. Acton JM. For better or for worse: the future of C3I entanglement, 2019, https://nautilus.org/napsnet/napsnet-special-reports/for-better-or-for-worse-the-future-of-c3i-entanglement/
66. Molander RC, Wilson PA, Mussington D, et al. *Strategic information warfare rising* (document no. MR-964). Santa Monica, CA: RAND Corporation, 1998.
67. Marshall AW. A program to improve analytic methods related to strategic forces. *Policy Sci* 1982; 15: 47–50.
68. Graubard MH and Builder CH. New methods for strategic analysis: automating the wargame. *Policy Sci* 1982; 15: 71–84.
69. Steeb R and Gillogly JJ. *Design for an advanced red agent for the RAND strategy assessment center*. Santa Monica, CA: RAND Corporation, 1983.
70. Schwabe WL. *Analytic war plans: adaptive force-employment logic in the RAND Strategy Assessment System (RSAS)* (document no. N-3051-OSD). Santa Monica, CA: RAND Corporation, 1990.
71. Shlapak DA, Schwabe WL, Lorell M, et al. *The RAND strategy assessment system's green agent model of third-country behavior in superpower crises and conflict*. Santa Monica, CA: RAND Corporation, 1986.
72. Bennett BW, Jones CM, Bullock AM, et al. *Main theater warfare modeling in the RAND strategy assessment system (3.0)*. Santa Monica, CA: RAND Corporation, 1988.
73. Allen P. *The secondary land theater model*. Santa Monica, CA: RAND Corporation, 1987.
74. Davis PK and Hall HE. *Overview of system software in the RAND strategy assessment system*. Santa Monica, CA: RAND Corporation, 1988.
75. Shapiro N, Hall HE, Anderson RH, et al. *The RAND-ABEL programming language: history, rationale, and design*. Santa Monica, CA: RAND Corporation, 1985.
76. Davis PK. *An analyst's primer for the RAND/ABEL programming language*. Santa Monica, CA: RAND Corporation, 1990.

Table 7. Key published documents on the RSAS.

Topic	References
DoD's challenge to the community	Marshall[67]
Automated wargaming with AI models (agents)	Graubard and Builder[68]
RSAS design	Davis and Winnefeld[10]
Red and Blue agents (national-command and military-level) using structured expert-system and slotted-script methods	Steeb and Gillogly[69]Davis et al.[14]Davis and Stan[13]Schwabe[70]
Green agent (parameterized rule-based models for 3D countries)	Shlapak et al.[71]
Global simulation (geography, forces, mobilization, combat)	Bennett et al.[72]
Simplified and adaptive theater modeling for secondary theaters	Allen[73]
Software architecture	Davis and Hall[74]
RAND-ABEL high-level language	Shapiro et al.[75]Davis[76]

AI: artificial intelligence; RSAS: RAND Strategy Assessment System.

Appendix 1

Selected pointers to RSAS documentation

Table 7 points the reader to key elements of RSAS documentation, which is even more extensive and also includes materials not in the public domain. Interested readers can request a compilation of the materials in Table 7 (at least 50 MB in size).

Author biographies

Paul K Davis is a senior principal researcher (retired, adjunct) at RAND Corporation and a Professor of policy analysis in the Pardee RAND Graduate School. He received his bachelor's degree from the University of Michigan and his PhD in chemical physics from the Massachusetts Institute of Technology. After research in defense-related strategic technology, he worked in government on strategic arms control, strategic nuclear planning, and—as a Senior Executive—on defense strategy and program analysis. He then moved to RAND where his research has involved strategic planning, modeling and gaming, deterrence and influence, counterterrorism theory, and causal social science modeling for policy analysis. Dr P.K.D. has served on many national panels and been an associate editor of several professional journals. His most recent book is the co-edited volume *Social Behavioral Modeling for Complex Systems.*

Paul Bracken is a professor of political science and business at Yale University. He received his bachelor's degree in engineering from Columbia University and his PhD in operations research from Yale University. He worked with Herman Kahn on futurology for business and defense for 10 years at the Hudson Institute. He then joined Yale where he has focused on global competition, the strategic application of technology in business and defense, and senior management challenges when dealing with a changing strategic environment and conditions of intense uncertainty. Dr P.B. is a member of the Council on Foreign Relations, has served on many advisory panels, and has been a visiting scholar at the CIA and Beijing University. His most recent book is *The Second Nuclear Age: Strategy, Danger, and the New Power Politics.*