

Illustrating a Model-Game-Model Paradigm for Using Human Wargames in Analysis

Paul K. Davis

RAND National Security Research Division (NSRD)

WR-1179
February 2017

RAND working papers are intended to share researchers' latest findings. Although this working paper has been peer reviewed and approved for circulation by NSRD, the research should be treated as a work in progress. The paper may be quoted and cited without permission of the author, provided that the source is clearly referred to as a working paper. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors. **RAND**® is a registered trademark.



For more information on this publication, visit www.rand.org/pubs/working_papers/WR1179.html

Published by the RAND Corporation, Santa Monica, Calif.

© Copyright 2017 RAND Corporation

RAND® is a registered trademark

Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of its research documents for commercial use. For information on reprint and linking permissions, please visit www.rand.org/pubs/permissions.html.

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Support RAND

Make a tax-deductible charitable contribution at
www.rand.org/giving/contribute

www.rand.org

Preface

This paper stems from a keynote presentation at a conference that dealt with relating human wargames to other elements of the Defense Department's analytic system. I propose one paradigm for doing so and summarize a small project that applied it in practice. I hope that the paper will be of interest to analysts, gamers, and officials who hope to make better use of human wargaming but are wary of its limitations.

The project that the paper draws on was sponsored by the Korea Institute for Defense Analysis (KIDA). The methods discussed are being further developed, so comments and suggestions are especially welcome. They should be addressed to the author at RAND in Santa Monica, California, or by email to pdavis@rand.org.

Preparing this paper was accomplished within the International Security and Defense Policy Center of the RAND National Defense Research Institute, a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the Unified Combatant Commands, the Navy, the Marine Corps, the defense agencies, and the defense Intelligence Community. For more information on the Center, see <http://www.rand.org/nsrd/ndri/ISDP.html> or contact the director (contact information is provided on the web page).

Abstract

This paper proposes and illustrates an analysis-centric paradigm (model-game-model or what might be better called model-exercise-model in some cases) for relating human wargaming to modeling and analysis. It is especially useful when considerable prior knowledge has already been captured in a model but the model may not adequately address the breadth and richness of issues and options that actual decisionmakers need to consider. Other paradigms are more useful when, for example, no good model exists initially, when the premium is on finding fresh boundary-bursting ideas, or when it is crucial to involve stakeholders in model development from the outset. The model-game-model paradigm was illustrated in an application to crisis planning on the Korean peninsula. It included development of an initial theory-based model, design of a war game to explore qualitative matters (e.g., options, criteria for evaluation, and uncertainty), and execution of such a game in Seoul, South Korea. The game confirmed many aspects of the model but revealed shortcomings that led to model enrichment with additional options and considerations. All of this illustrated successfully one cycle of the model-game-model process. Further cycles are planned.

1 Background, Objectives, and Structure

This paper is adapted from a keynote presentation in a Department of Defense (DoD) workshop on wargaming.¹ The workshop stemmed from a DoD initiative by Deputy Secretary Robert Work to use human wargaming (wargaming for short) as an important part of promoting innovation. Traditional modeling has not proven adequate to the department's needs. The initiative includes efforts to (1) revitalize wargaming, (2) embed wargaming more firmly in DoD's suite of analytical approaches, and (3) do a better job of sharing wargame insights with senior leadership.²

Consistent with guidance from the Deputy Secretary, the specific workshop's objectives were (1) to identify best practices for the the use of wargaming *to support the larger analytic process*, and (2) to help generate wargaming capacity and quality. Wargaming itself has long proven its usefulness, but connecting it to the department's analytic processes has always been a challenge. Analysis organizations often look askance at wargaming because of its lack of rigor and reproducibility. Wargamers often see themselves as doing something distinct from modeling and analysis, although acknowledging a larger cycle of research that also includes analysis and experiments.³ The presentation from which this paper grew set the stage for a large working group tasked with better understanding actual and desirable connections. The immediate audience was a mix of national-security modelers, war gamers, and analysts.

The presentation described *one* paradigm. The working group later examined some others as well.

The paper has two distinct parts. Section 2 is somewhat abstract as it discusses relationships among modeling, gaming, and analysis and then focuses on a particular paradigm that uses gaming to review and enrich models used for more rigorous analysis. This analysis-centric view is only one way to connect gaming to other analytic processes, but it is an important one. Section 3 describes a concrete application involving potential nuclear crisis on the Korean peninsula. It provides a case study on what is involved in actually taking the approach.

2 Relating Modeling, Wargaming, and Analysis Relationships Should Depend On Intentions

Alternative constructs exist for relating modeling, wargaming, and analysis. Which is better depends on purpose and context. Methods would vary, for example, for applications to near-term operations planning, mid-term strategic planning, long-term strategic planning, and crisis planning.⁴ So also, the methods would vary across purposes indicated in Table 1.⁵

Table 1 Table 1 Different Purposes for Wargaming

<ul style="list-style-type: none">• <i>Fill in details</i> when representing concepts of operations in computer models• <i>Test and tighten</i> tentative plans—finding flaws and seeking to make them more robust
<hr/> <ul style="list-style-type: none">• <i>Create</i> ideas, options, concepts of operations, or strategies• Understand trends and possible futures• <i>Broaden and enrich the understanding of objectives, possible objectives, and factors affecting both phenomena and decision, to include understanding potential adversary reasoning</i>
<hr/> <ul style="list-style-type: none">• <i>Sensitize policymakers</i> to known serious problems• <i>Communicate and socialize ideas</i>, resulting in shared understanding and personal relationships. <hr/>

The first group in Table 1 involves improving rigor and adding details with predetermined structures; the second group involves creativity and broad, human thinking that includes numerous qualitative considerations and speculations about the possible; the last is about communicating and sharing ideas and constructs. Other purposes are sometimes mentioned, such as "validating a theory" or "calibrating," but one should be skeptical about those for reasons discussed in the literature.⁶

In this particular paper, I have in mind primarily the bolded function of using gaming to "broaden and enrich analysis-centric models," to include understanding the adversary's potential reasoning. That often improves the quality of analysis more than further increasing "rigor."⁷ I introduce a generic approach for doing so and then describe an first-time-and-therefore-experimental project applying it—treating the project as an early case study to assess feasibility of the approach, identify difficulties, and yield suggestions for future applications. The particular application is to crisis planning on the Korean Peninsula, which was attractive because crisis there is quite possible and because there was a long history of valuable crisis gaming on which to draw.

A Model-Game-Model Paradigm

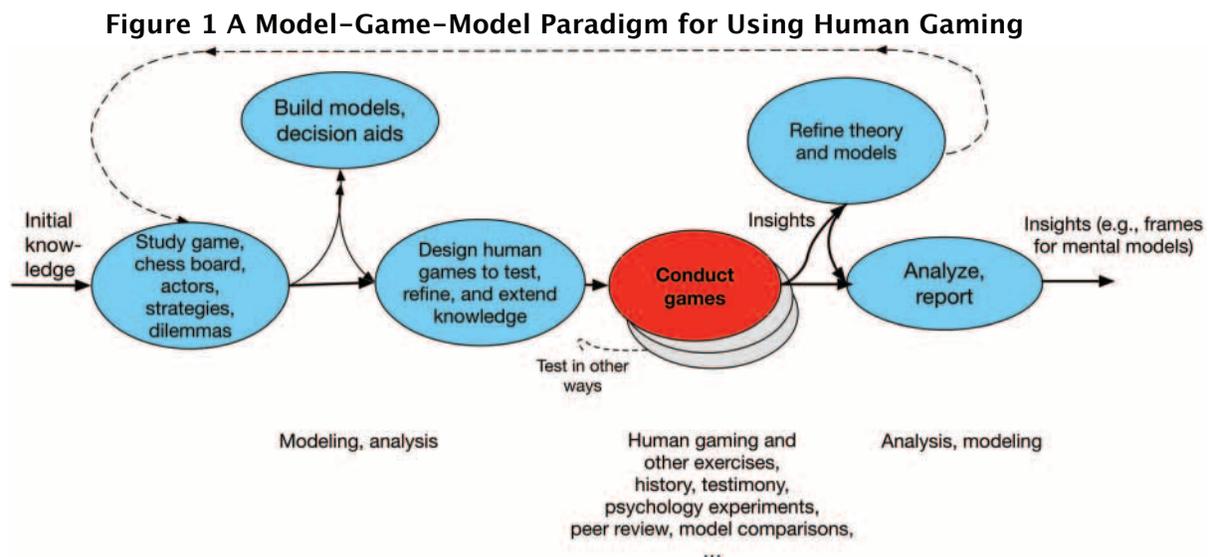
Again, the particular paradigm that I will describe is analysis and model centric. Its intent is to improve the quality of the models being used to deal with a subject. This relates well to the belief that crisis decisionmaking depends on leaders' mental models. Those models should obviously be as well informed as possible, but the term "models" must be interpreted broadly. Models may be (1) qualitative or quantitative; (2) simple or complex; (3) rough and exploratory, or rigorous and predictive; (4) imaginative, or constrained by consensus and conventional wisdom; and (5) open or closed (i.e., with or without human interactions). Further, the function of the models may be narrow prediction, uncertainty-sensitive exploration of possibilities, description, explanation, and/or communication.

The reader who is merely nodding, "of course," should pause to appreciate that—within what is usually referred to as the DoD analytic community—the term "model" is ordinarily construed much more narrowly. Some authors have been notably broader.⁸

With these prefacing remarks, Figure 1 displays the paradigm for the rest of the paper. It is in some respects akin to the theory-test-theory version of the scientific method or to the

model-test-model approach from engineering. Its motivation is that in many problem areas it is *necessary* to reason, design, evaluate, and make decisions based on models: it is simply not feasible to rely on empirical methods because doing so would require innumerable controlled experiments that are neither affordable nor feasible, even in principle (e.g., we cannot conduct controlled experiments on nuclear crises).⁹ In such cases, the models relied upon—especially the largely qualitative mental models—obviously need to be as good as possible, to include dealing with uncertainty and dilemmas. The quality of such models often has to do more with qualitative considerations and framing than with details. How can an analytic process that includes gaming contribute?

Figure 1 uses the metaphor of learning a game such as chess when contemplating a class of possible future crisis. In this metaphor we can try to understand a problem by seeing it as a game between us and an adversary—its rules, the game board, the nature of our adversary (or adversaries), the sides' possible strategies and tactics, and so on. We then build qualitative and quantitative models to represent our understanding.¹⁰



Suppose that we have done this. We believe that we understand the game, or at least partly understand it. We have models that incorporate our understanding. What can we do next? Physical scientists can go to the laboratory or to nature and run controlled experiments. Economists can craft "natural experiments" by exploiting natural variations in how otherwise similar groups of people approach particular problems. Often, however, we do not enjoy the

convenience of having laboratories or other mechanisms for controlled experimentation. Thus, we employ other methods, one of which is to run games and think about them—with cautions discussed below—as experiments.* Often, the primary value of human games is to senior staff and actual leaders who directly participate—often learning rapidly about "the chess game" and achieving remarkable common understanding with other participants (to include an understanding of disagreements).¹¹ If instead the purpose is analysis, then we should design the games to accomplish something analytically useful. We want games that will probe the edges or illuminate matters not understood or recognized.

Such games may extend knowledge, discover errors in previous thinking, or identify items to incorporate in a refined theory and related model. For example, we may gain insights about additional tactics and strategies. We might then do some analysis and report tentative conclusions, including suggestions on how to frame mental models, but we would also want to reenter the cycle of thinking, theorizing, modeling, and looking for information (related to Peter Perla's "cycle of research").¹² The cycle should continue as needed, until—if ever—we arrive at settled theory. That time may never come for national security work because the world keeps changing, as do our competitors and adversaries. Even the nature and rules of the game are unstable.

Before continuing, I note that other approaches are sometimes useful or even preferable. Two examples illustrate the point. Some researchers seek "grounded theory" that is obtained by inductive reasoning from empirical data, rather than seeking to test a pre-existing theory that may not even exist. As a variation, some use abductive reasoning to select a theory from among various candidates, one that seems to be the best explanation.¹³ Another, very different approach, is participatory modeling, which can be valuable in causing stakeholders in a decision to take ownership of the model used for analysis.^{14 15}

* Careful terminology distinguishes between games and exercises. Definitions vary, but a good one is that "a war game is "a warfare model or simulation whose operation does not involve the activities of actual military forces, and whose sequence of events affects and is, in turn, affected by the decisions made by players representing the opposing sides" (Perla, 1990). Human exercises are not so restrictive. For example, they many not have separate teams or they may have discussion but not "moves." Informally, such exercises are often referred to as games and I have found that attempting to enforce the distinction is, for many audiences, distracting. Thus, in this paper, I use the term "game" broadly to include exercises that include separate human teams. They may have computer assistance, but humans make crucial decisions.

Can Human Gaming Truly Serve as "Testing"?

An important question arises at this point: how can human gaming be used to test anything? A game, after all, is not a controlled laboratory experiment. The players are usually not decisionmakers and, even if they are, they may not be the people who will make decisions if the crisis or conflict being gamed actually arises.¹⁶ The game's adjudication of moves may not represent well how real-world actions would play out or what effects they would have. And the conditions posited for the game are merely illustrative. The shortcomings of gaming are many if the measure is reproducing the real world or accurately predicting the future.¹⁷ How, then, can they be used meaningfully in the way Figure 1 posits? The answer is not straightforward.¹⁸

Not surprisingly, Thomas Schelling had important comments on the matter.¹⁹ Schelling addressed the blunt questions, such as "Isn't a game a unique story that may never be repeated?" and "Is there not a danger that participants will be so carried away in this vicarious experience that they identify the game too much with the real thing and learn 'lessons,' perhaps overlearn them, that will prejudice their judgment in the future?" He went on to make a decisive point, which he and I also discussed in a later lunch conversation:

The answer is that games, in this regard, are not different from real experience...

As with tentative "insights" that we gain from real-world events, we must ponder and assess whether the insights are generally valid or merely peculiar to the situation. Schelling goes on to observe something compelling:

Ideas are so hard to come by that one should be ready to take them anywhere one can find them.

So, human games cannot be used to "test" a model in the ways that physical scientists and engineers can. Nor is a human game a good way to calibrate the probabilities expressed in a model.²⁰ What human games *can* do, however, is considerable. For example human games can:

- *Falsify* a model by demonstrating that humans frame issues differently than is allowed for in the model.²¹

- *Enrich* a model by noting additional factors that it should contain because people consider them (e.g., What is driving the adversary's behavior? Internal dissension? Fear? A belief that we will capitulate under pressure?)
- *Creatively change the character* of the problem by noting additional tactics and strategies that competitive people think of. Or perhaps the additional tactics and strategies represent changes of rules or recognition that the adversary has a major vulnerability that 'should not' logically exist (and which the model therefore didn't anticipate), but does (e.g., the edges of the Maginot Line).
- *Identify "frictions"* that need to be represented in a realistic model, frictions such as decision delays caused by the necessity of communication and discussion with allies, or by peacetime rules and procedures that cannot suddenly be put aside. Another class of examples relate to differences between nominal and real capabilities, as when a mission capability exists in principle, but not in reality because "Oh, the unit that could do that hasn't trained or exercised for that mission in years. It will take weeks of preparation before it could execute it again."

Most such contributions from human gaming are qualitative rather than quantitative. I am very skeptical about using games as though they constitute evidence about effects or results.²²

Although I do not dwell on the matter here, wargaming has a very long record of value in military planning, and on related issues of force development.²³ A notable aspect has been its role in historical innovation, to include German gaming of Blitzkrieg and U.S. gaming of how to use newly emerging precision fires.

Possible Principles for Best Practices

If we now ask how to think about human wargaming in a model-centric approach, Table 2 suggests appropriate and inappropriate questions to ask of human gaming. The second column indicates questions that are inappropriate because of inherent uncertainties. Or, to be more precise, it may be legitimate to ask the questions, but one should not imagine that the answers that come from gaming are reliable. A number of papers over the year bear on similar issues.²⁴

Table 2 Questions to Ask and Not Ask of Human Gaming

Good Questions	Bad Questions
What likely and plausible strategies might the adversary employ? Which are more and less plausible?	Cut to the chase. What <i>will</i> he do?
What considerations (factors) have we omitted that need to be included?	What is the full set of factors? What are their precise values?
What could go wrong? What could be much less efficient or effective than assumed?	How precisely will things go?
How might the adversary be reasoning, again considering both likely and plausible?	How will he reason?

A Partial Analogy: the Role of Case Studies in Social Science

Before turning to a specific example of using the paradigm of Figure 1, it may be useful to draw a relationship between the model-game-model paradigm and the use of case studies in social science. Social science includes a great many partial theories about how people and groups behave. For example, economic theories emphasize the role of incentives and a notion that people will act as though they are seeking to maximize their expected utility. Case studies provide opportunities to examine the factors at work in considerable detail empirically, on a case-by-case basis. Do they falsify, enrich, or support the theories?²⁵ Often, this can be addressed only qualitatively, as in concluding that "Yes, that consideration apparently mattered a lot"²⁶ or, as an anthropologist might say "The observed behaviors were rooted in sacred values that don't have materialistic measures. You can try to transform their actual reasoning into your materialistic terms, but, really, you just don't 'get it'."²⁷ Anthropologists were among those who contributed to development of the counterinsurgency field manual associated with Army and Marine Generals David Petraeus and Jim Mattis.

The analogy, then, is that the "testing" (whether by case studies or a human game) is often about constructs, considerations, and other qualitative matters, rather than matters quantitative and precise.

3 Case Study: Deterrence and Stability on the Korean Peninsula

Let me now sketch an experimental case study that adopted the paradigm of Figure 1. This study involved a collaboration with the Korea Institute for Defense Analyses (KIDA), in Seoul, South Korea.²⁸

Background

The background for the study was prior work that reviewed Cold War deterrence theory critically and then applied the concepts to the Korean peninsula. An important conclusion was that U.S.-Korean extended deterrence has *serious* difficulties.²⁹

Maintaining and strengthening the U.S. extended deterrence is no longer straightforward—no matter how fervent and sincere the official statements of assurance. Extended deterrence is inherently difficult when the adversary can strike the extending power with nuclear weapons. NATO faced up to the credibility issue in the 1970s. Something analogous is necessary for Korea.

The conclusion was provocative and upsetting when initially presented to audiences in Korea and the United States, but is being increasingly recognized in US-South-Korean discussions.

The problem in a nutshell is that extended deterrence is easily possible when the deterrer is secure and can severely punish the deteree. It is not so simple when the question arises “Would the United States really trade Washington or Los Angeles for Seoul?” This question is not very different from “Would the United States really trade San Francisco for Paris?” a question that NATO faced in the 1960s. NATO recognized the structural problem and took strides to improve the quality of extended deterrence even though the Soviet Union could readily strike the United States with nuclear weapons.³⁰ What will the United States and South Korea do? The answer remains unclear, even more so due to political turmoil in both nations.

In the meantime, KIDA and I concluded that the focus problem for collaborative work in 2016 would be the following:

How should the Republic of Korea (ROK) and the United States prepare for possible major crisis in which war—even nuclear war—is plausible?

By "major" crisis we meant a crisis significantly worse than those caused by usual North Korean provocations. Our work would attempt to apply the model-game-model paradigm described earlier. Subsequent sections deal, respectively, with (1) model building and game design and (2) executing a human game and learning from it. Thus, our work included just one cycle of Figure 1.

Model Building

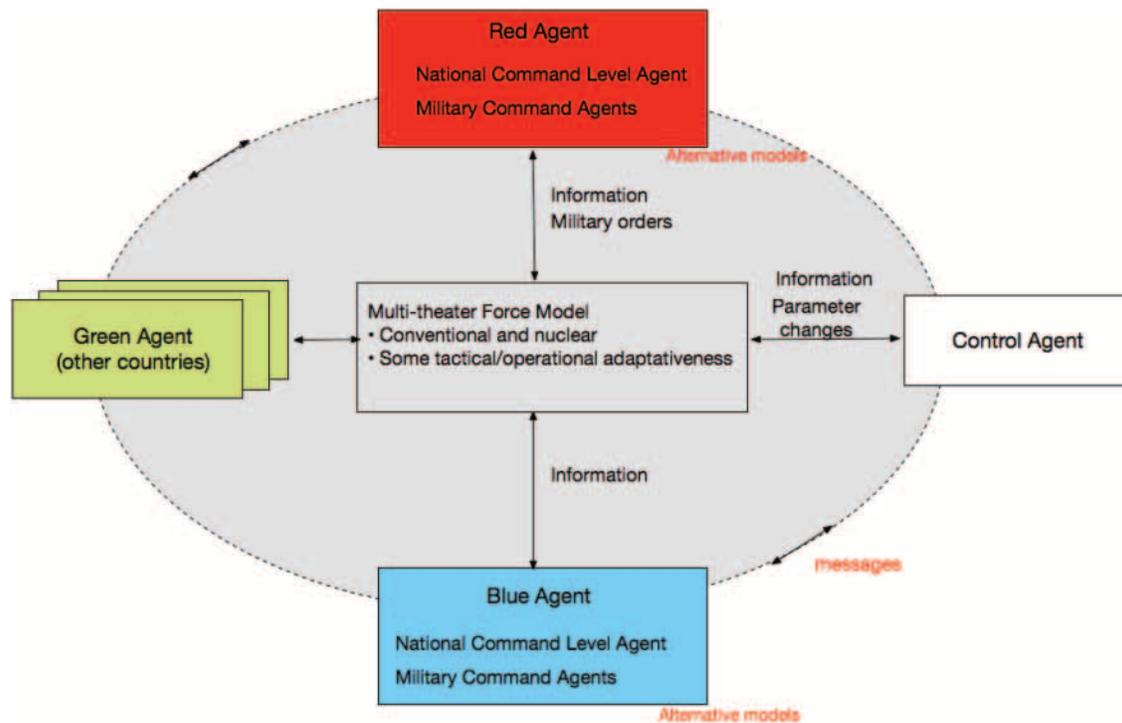
Ideal Methods and Practical Expedients

Ideally, in taking the model-test-model approach, we would have used a sophisticated game-structured simulation as in Figure 2, which shows the architecture of the RAND Strategy Assessment System (RSAS) of 30 years ago. It had a multi theater combat model at its core (center of the diagram), which covered the spectrum from conventional war to general nuclear war, and which promoted uncertainty analysis.³¹ The colored rectangles indicate automated Red, Blue, and Green agents, but human teams were interchangeable with agents.³² The RSAS had *alternative* agents because of fundamental uncertainties about how real-world Soviet/Warsaw Pact and U.S./NATO leaders would reason. The Green Agent consisted of simpler rule-based models for third countries, with some overlaps. For example, France might joint with NATO and be reflected in part through the Blue Agent, but could make independent decisions about its own nuclear weapons.³³ This mattered a great deal in the simulations, a surprise to many of us at the time because of our superpower-centric mindsets.

The RSAS embodied knowledge from political science, psychology, and organizational theory among others. From an artificial intelligence perspective, the agents used a structured version of rule-based modeling at the national-command level and a version of slotted scripts at the level of theater commanders. These agents were quite different from the ones used today in agent-based modeling (ABM), but could relate well, we hoped, to the thinking of real-world leaders. A key point, however, was that even if the models were relatively good, the input uncertainties were enormous. Thus, we stressed "multiscenario analysis" across dif-

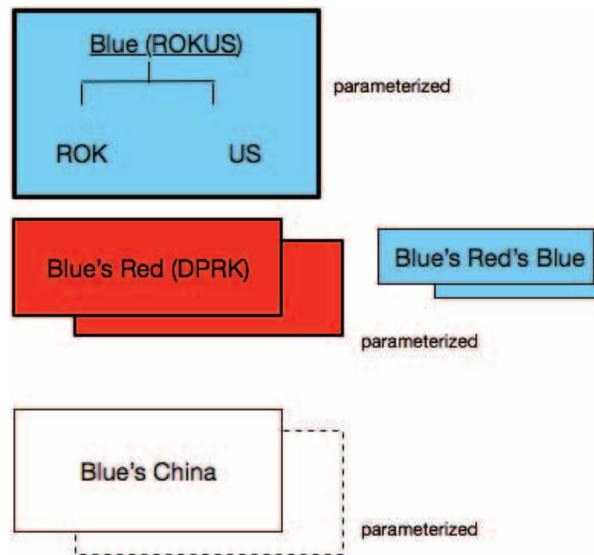
ferent contexts, different assumptions about the thinking of Red, and so on. This evolved in subsequent decades to what I have called *exploratory analysis* and exploratory modeling.³⁴

Figure 2 A Precedent: the RAND Strategy Assessment System



As the 2016 project with KIDA began, neither the RSAS nor anything like it existed. So how could we proceed with the vision of Figure 1? The answer is that we used a poor man’s version of the RSAS concept: no combat model, no treatment of most countries, and only very simple representations of Red, Blue, and China. We included, however, one of the most important aspects of the earlier work: a Blue Agent model with a model of (Blue’s Red), which in turn reflected Blue’s notion of how Red perceived and would respond to Blue (i.e., a primitive Blue’s Red’s Blue). I will merely touch upon some of the primary concepts used.

Figure 3 A Highly Simplified Pol-Mil Model



To actually build models even for our limited purposes it was necessary to address a number of issues described in what follows. The challenges were how to organize thinking about levels of conflict, structure cognitive models, identify the options that the models should consider, identify the considerations (factors) that should inform decisions, and estimate the consequences of executing the options.

Modernizing the Escalation Ladder

As a structuring concept I updated a version of the escalation ladder developed by Herman Kahn in the 1950s and early 1960s.³⁵ Kahn's ladder had 44 rungs. His construct was presented as a metaphor, fully recognizing the ambiguities and subtleties. It was the result of a great deal of thinking, debating, and human gaming. A primary theme was that nuclear war is not like a binary switch. It's not as if one either has peace or general nuclear war. A fundamental problem is that aggression can be small or large, narrow or broad, subtle or crude. How does one deter aggression generally when the threat of massive across-the-board nuclear attack would make no sense and has no credibility? Among those who became policymakers, Henry Kissinger made his initial reputation by writing about such matters, the "problem of limited war."³⁶

Escalation ladders today are more complicated because of the items indicated in Table 3: the "new domains" of space and cyberspace, the asymmetric situation in which some U.S.

adversaries have chemical and biological weapons while the United States has profound capabilities for precision strikes. So also, strategic cyberattack and electromagnetic pulse (EMP) attacks have added new dimensions. The grayed rows show complications that may at first seem new, but that are not actually. As mentioned earlier, extended-deterrence problems were evident in the 1950s.³⁷ So also, history has numerous examples of what we now refer to in terms such as “little green men” and “hybrid warfare.”

Table 3 Modern Complications in an Escalation Ladder Concept

Issue	Comment
Space and cyberspace	Mostly new
Chemical and biological weapons	U.S. sees them as weapons of mass destruction
Precision conventional weapons	Devastatingly effective in conventional warfare
Strategic cyberattack	Role in decaptation, infrastructure attack, and counterforce attack (e.g. disruption of communications)
Electromagnetic pulse (EMP)	New implications (effects on command and control and electronic systems)
Inherent weakness of extended deterrence	Not new. As for NATO in 1970's.
Little green men and hybrid warfare	Not really new
Multiparty crises	Not really new

Although the complications initially suggested that an escalation ladder was no longer feasible (an escalation *lattice* might be better), it turned out that much could be done for the Korea-Peninsula problem with some simplifying assumptions. This included assuming that many adversary actions that the United States would *prefer* to see as “escalations” should be expected and should not merit being on a separate rung.

For initial purposes, and glossing over some subtleties, we used a simplified Korea-specialized escalation ladder as in Table 4. It recognizes that Red and Blue (North Korea and South Korea/U.S.) probably have different escalation ladders. North Korea might find it necessary to use chemical and even biological weapons because of the U.S. advantage in precision weapons. So also, it would not likely treat U.S. space systems as in sanctuary.³⁸ The ladder shown also assumes that the North Korean leaders might—in extremis—prefer to go out with a dramatic last-gasp destruction of its enemies, rather than capitulation. This would not be rational in the normal sense, but might be a consequence of the North Korean leader hav-

ing come to believe the propaganda in which the Kims are almost god-like. I would not predict such reasoning and I see Kim as rational, however misguided, but the possibility cannot be dismissed given the nature of Kim-Dynasty ideology over the decades. In this connection, it is appropriate to mention part of the DPRK's lore, a story to the effect that Kim Il Sung once had a meeting in which he asked his generals what they would do in a war if North Korea were losing. Everyone was shocked and quiet until Kim Join Il spoke up: "Destroy the Earth. What good is this Earth without North Korea?" This was the correct answer and Kim Jong Il rose in his father's estimation and later succeeded him. Although lore is only lore, the words "A World Without North Korea Need Not Survive" appear prominently in the Three Revolutions Exhibition Hall in northwestern Pyongyang.³⁹

Table 4 A Simplified Escalation Ladder for North Korea and for South Korea and the United States

Option	Red	Blue
Wait and see	•	•
Demo nuke	•	•
Limited conventional strike	•	•
Decapitation	NA	•
Conventional war	•	•
Counterforce against nuclear threat	NA	•
Chem/bio use in ground operations	•	NA
Tac nuc use in ground operations	•	•
Lim Strat (limited strategic-nuclear strikes)	•	•
All-out war	•	•
Countervalue attack	•	•
Extended countervalue attack (against the U.S. homeland)	•	NA

NA: not applicable

Cognitive Decision Models

The next challenge was creating simple “cognitive decision models” roughly akin to the agents of the RSAS described earlier. My approach has been to create a structure *allowing*

for rational decisionmaking, but also allowing major departures. Here a "rational structure" implies having multiple options, multiple criteria for evaluating the options, and multiple factors to use in that evaluation. It also means choosing more or less the best option for achieving objectives.

Real people cannot do the related calculations: they lack the information needed and do not know how to do the complex and uncertain probabilistic mathematics. Because of such bounded rationality noted by Herbert Simon a half-century ago, they use heuristics.⁴⁰

The heuristic I have long proposed in representing limited rationality is to have the model assess the outcomes of each option using perceived best-estimate, best-case, and worst-case assumptions.⁴¹ We are all familiar with this style of reasoning, which is perhaps the best that can actually be achieved in most real-world situations. To be sure, all the estimates will be flawed due both to misperceptions, miscalculations, and psychology's cognitive biases,⁴² but at least effort will be thereby made to account for uncertainty by recognizing that options entail risks and may also have upside potential. Memoirs of past decisions by Presidents and top advisors show evidence of their attempting this level of reasoning.

Another key element of the approach is to demand *alternative* adversary models because we don't know how adversaries (or even future Presidents) will think. Our analysis should attempt to bound such matters, rather than focus on alleged best estimates. Indeed, the "tyranny of the best estimate" is one of the afflictions constantly besetting decisionmakers.⁴³

Top-Level Structure

Table 5 shows the essence of the top-level concept, simplified to consider only three options and two models of the adversary, Model A and Model B.

Table 5 Generic Top-Level Structure of Cognitive Decision Model

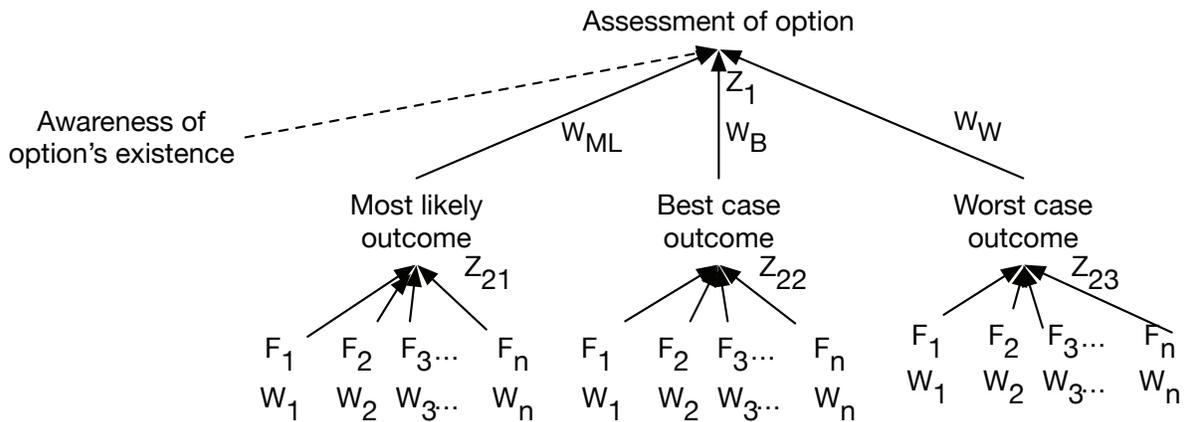
Option	Most Likely Outcome		Best Case Outcome		Worst Case Outcome		Net Assessment	
	Model A	B	A	B	A	B	A	B
1								
2								
3								

In the simplest version of this cognitive model, the net assessments are linear weighted sums of quantified versions of the option assessments. The models may have different weighting factors. For example, a risk-taking model with glorious objectives may give little weight to worst-case possibilities; a risk-averse model concerned primarily with avoiding disaster will give substantial weight to the worst-case assessment. The models may also make different assessments of each of the outcomes shown. That complicates the work analytically, but is important substantively because the "leanings" with respect to weights and with respect to option assessments will *often* be correlated as when risk-takers are also more optimistic about results.⁴⁴ More complex versions of the model are possible. For example, the net assessment can be nonlinear, as when a model rules out an option, rather than just rating it a bit lower, if risks are too high. This is a familiar strategy in real-world decisionmaking.

Lower-Level Structure

The next level of sophistication is to characterize how the models estimate option outcomes given the many errors in their perceptions of reality, including their perceptions of their own adversaries. Figure 4 shows the basic concept. An option is assessed using "most-likely," "best-case," and "worst-case" assumptions. These assessments depend on a number of factors (the F's), but involve different assumptions about factor values. Precisely how the factors combine to generate assessments (represented by the Z functions) could again be as simple as linear weighted sums, or something more complex and nonlinear.⁴⁵ It is likely, in any case, that the assessments will need to give different weights to different factors (the W's).

Figure 4 Multi-Level Factor Tree Assessment of Options



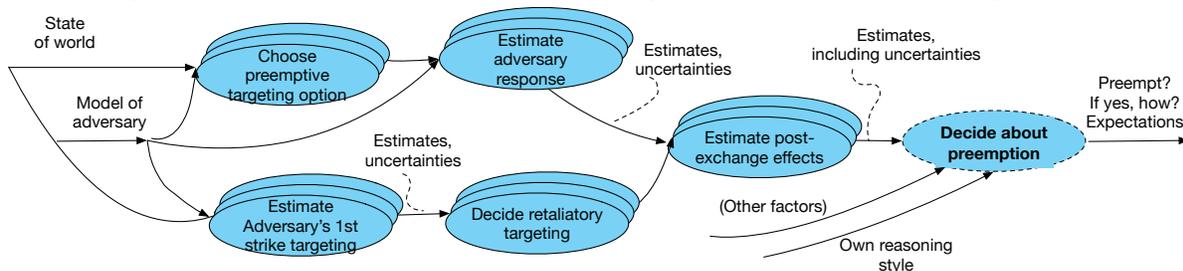
Z_1 = function of personality, psychological domain, desperation, optimism...

Z_{2i} = function of utility function, if any; beliefs, perceptions, personality, psychological domain, desperation, optimism, other cognitive biases...

W_i : weight of factor i

Despite the simplicity of this cognitive-model concept, implementing it is not trivial. Figure 5 illustrates one element of the difficulty. Suppose a model is contemplating a preemptive option, i.e., an option to escalate before the adversary does. The model should consider each preemptive option, estimate what the adversary will do in response (recognizing uncertainties in how the adversary thinks), and characterize the outcome. If he is sure that the adversary is about to escalate (as implied by his referring to “preemption” rather than “escalation”), then he should also estimate the consequences of allowing the adversary to take the next move and then compare results to estimate the “value of preemption.” If he is not certain, then he should account for the probability of mis-estimating the adversary’s intentions.

Figure 5 Features of Even a "Simple" Cognitive Model Considering Preemption

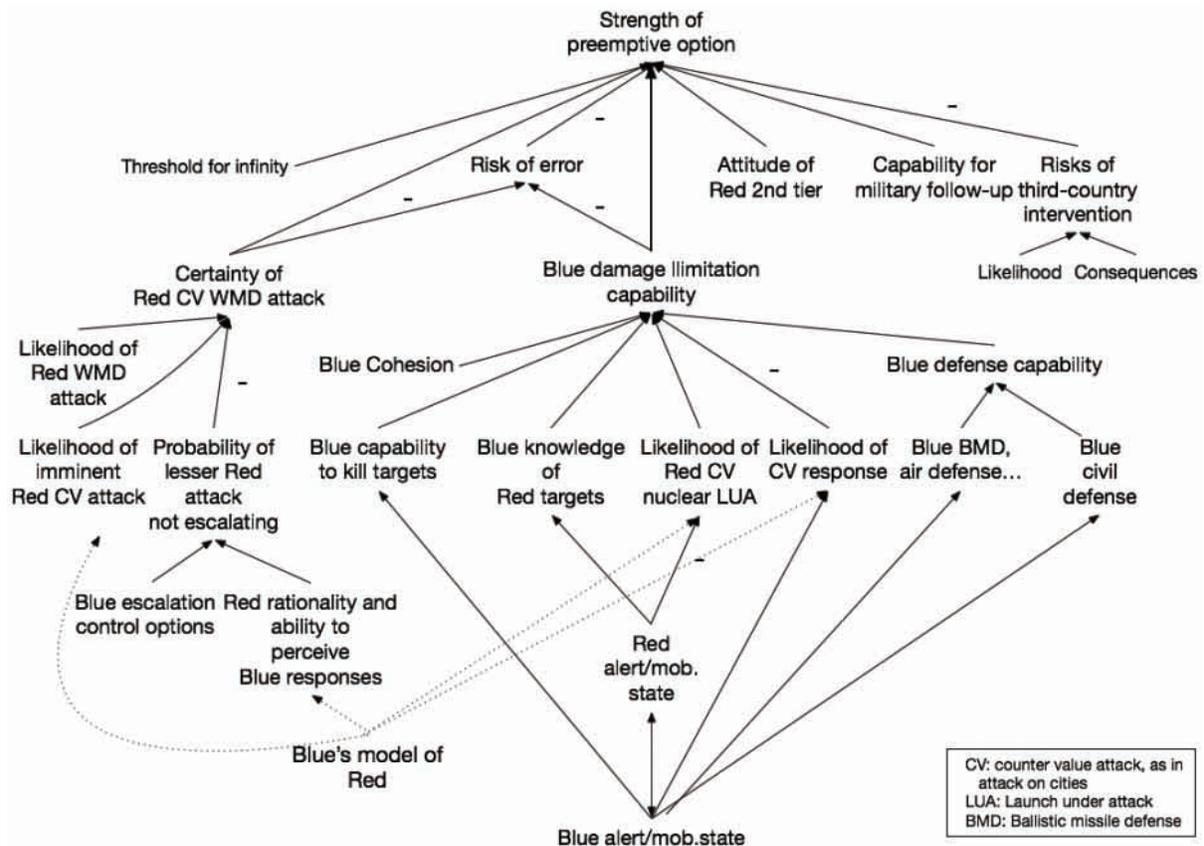


I implemented these ideas in a highly parameterized model written in Analytica®, a product of Lumina Decision Systems. The options that Blue considers, for himself and for his image of Red, correspond to the various levels shown in Table 4. The model is drastically simplified relative to an operational-level simulation or war game in which ground, air, and naval forces would be executing strategies. The premise for our work was that we wanted to focus exclusively on weapons-of-mass-destruction-level issues. Details of conventional operations were not important for our purposes even though they might involve major battles and substantial loss of life (as would a major artillery barrage of Seoul).

Continuing, Figure 6 indicates schematically what I sought initially to include in the model. My KIDA colleagues and I used the graphic for discussion and debate before the model was refined and implemented. Although the graphic is qualitative, the model is semi-quantitative. Fundamentally, although concepts have been mapped to numbers, the figure is still qualitative in the sense of being approximate, soft, and squishy—as with reality. Although Figure 6 shows an along-the-way representation, a more mature version would be what I have called a "factor tree." The methodology for developing factor trees is described with examples elsewhere.⁴⁶

Figure 6 Modeling the Military Value of a Preemptive Option

(original depiction, subsequently altered significantly)



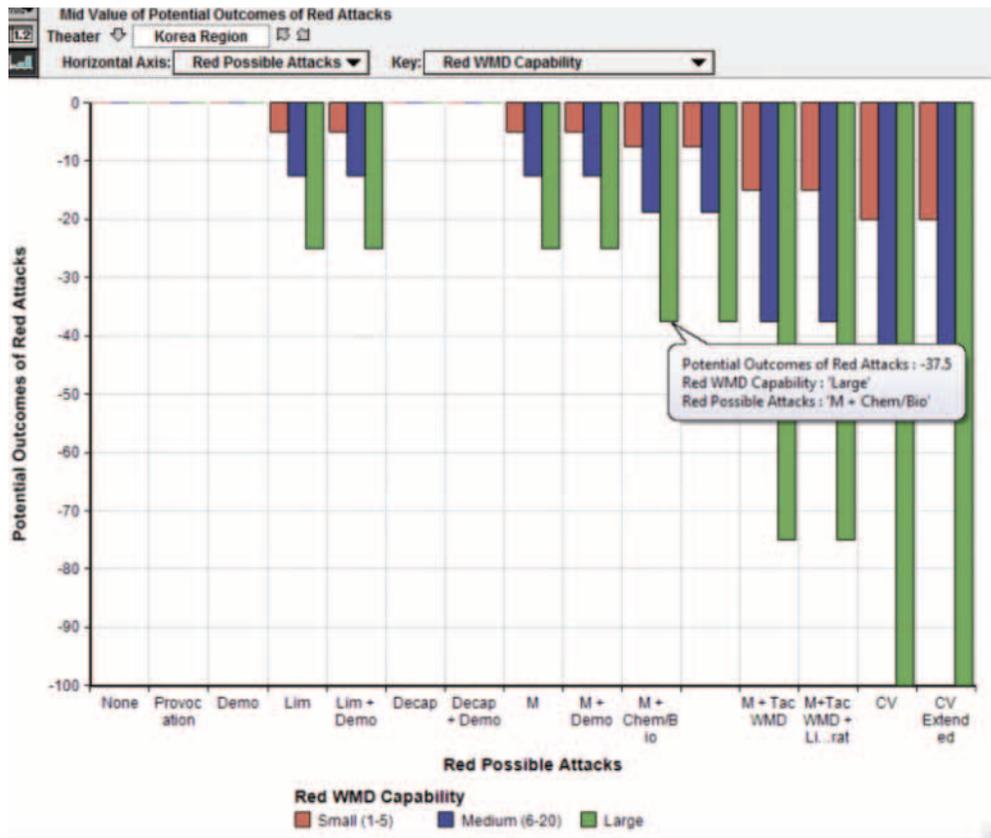
Even in an unabashedly simplified depiction, it was necessary to characterize the consequences of executing various options and receiving Red's response. For simplicity and because it is what matters most, the model focused on damage to Blue. As shown in Figure 6, the model used crude estimates on a 0 to -100 scale for the various types of attacks. The values were subjective inputs, not calculations. They should be seen as nothing more than rough-cut qualitative indicators. In principle, they could have been based on separate detailed studies. They could also have been converted to casualties or losses of economic value, as in Cold War calculations. All such matters were beyond our scope in the pilot study.

At this point, we have seen the essence of the model used: its cognitive structure, the options considered, the factors affecting evaluation of options, and crudely estimated consequences ascribed to various options.

Although more qualitative than not (despite using numbers), the model was useful for thinking. It seemed to incorporate a good of knowledge. For example, it could roughly character-

ize the military consequences of various actions and reactions—as a function of many uncertain parameters relating to alert levels, actual offensive and defensive capabilities, intelligence capabilities, and each side's model of the other. But what could human gaming do to test or enrich the model?

Figure 7 Postulated Damage Data for Different Red Attacks



Note: the numerical figures are purely subjective and illustrative. They are sufficient to make useful qualitative distinctions visually—e.g., between outcomes that are "extremely bad" (-100), those that are bad but do not imply the end of the state attacked (e.g., -30), and those that might arise from much more limited war (e.g., -10). Such methods are semi-quantitative in that they include numbers and allow for graphs, but are ultimately qualitative, i.e., about soft and imprecise matters.

Designing and Executing a Human Game

Having developed a semi-quantitative model that incorporated a good deal of strategic theory, what did we want to accomplish with a human game? Consistent with the ideas described earlier, we sought primarily to look at qualitative matters, such as:

- *Strategies and options* (for both Blue and Red): would teams go beyond those we had identified?

- *Criteria* for evaluating options, such as short-term or long-term benefits and costs, both military and political: would teams use different or additional criteria?
- *Factors* in evaluating options: would teams use different or additional factors?
- *Mindsets*: would teams reason in ways understandable with the model or in very different ways?

As one example in our modeling work, because of insufficient time we had done minimal thinking about China, although it was evident that China would play an important part in crisis thinking by both Red and Blue. Human gaming would, we suspected, jump-start a next iteration of the model that would take China into account substantively. And perhaps Japan as well.

With these constructs in mind, we designed the human war game with some specific features:

- Use an alarming initial scenario unlike "normal" DPRK provocations
- Make war, even nuclear war, appear truly possible (not "unthinkable")
- Have separate teams for Blue (ROK and US), Red (DPRK), and China
- Give special instructions to teams (not necessarily consistent) to sensitize and goad them, but with substantial uncertainties, raising war alarms and issue of rationality.
- Instruct teams to (1) develop options, (2) develop alternative models of their opponent, (3) consider most-likely, best-case, and worst-case outcomes for each option; and (4) *explain* their reasoning.
- Have two rounds of play, with the second round's conditions ratcheting up tensions (first-round moves were *not* adjudicated and used to determine conditions of the second move; thus, this 'game' was more accurately seen as an exercise with different vignettes).⁴⁷

The game was held in June 2016 at KIDA headquarters in Seoul. KIDA organized and conducted the game. I served as an advisor and observer, and gave players an introductory briefing. The game was a day in duration, had no computer support, and had minimal supporting materials. Thus, it was closer to a seminar exercise than an operational-level wargame

with military war planners addressing maps, details of maneuver and logistics, etc. Participants were both civilian and military, largely associated with think tanks or government policy offices or with military political-military divisions, rather than developing war plans. Players were well qualified. The game was, of course, unclassified. It was conducted almost exclusively in Korean, with translation in some instances and with side translations for some Americans individuals such as myself.

Since the game was an experiment, we prepared the format of some grade cards in advance. Table 6 shows our assessment, which was informed by a post-game survey of players conducted by KIDA. The game itself went very well with players reporting favorably on the day’s activities and its usefully provocative nature. The biggest disappointment was that players did not ask for or demand much information, as would surely occur in a real-world crisis environment.

Table 6 Assessment of Game Experiment

Issue: Did teams	Result
Accept scenario?	Yes
Develop options?	Yes
Define options adequately?	Yes
Have in-depth discussion	Yes, at the political level, not much at the military level
Have alternative models of adversary	Yes, sketchy but meaningful
Make most-likely, best-case, and worst-case option assessments?	Yes
Have disagreements?	Yes, but polite
Provide explanations?	Yes
Take seriously the need to address their team's perspective?	Yes
Consider contingent options?	Implicitly at best
Request information?	No

A next kind of grade card was to see how the game related to the model developed before hand. Table 7 describes conclusions about some of the model versus game issues. A primary reason for using human war gaming is that models are often too simplistic and “mechanical.” Human players can quickly point out additional factors and considerations. They

can often be more creative. In this case, however, results were mixed and in some respects the pre-existing theory and model were richer than the game discussion proved to be. The human game tended to *confirm* the prior theory and model. The human teams, however, added options and elaborated reasoning in ways that the original theory phase had not dealt with well.

We were struck that few of the important factors identified in the model (left column) were dealt with by the human teams. We gave a good deal of thought to why. Ultimately, it was the predictable result of the choice of people recruited for the game, and the game's political-military (rather than military-political) nature. The players were outstanding, but reflected their day jobs. An additional factor may have been that the game included both Koreans and Americans. Finally, a big factor was that the model constructed ahead of the game reflected years of experience with nuclear-crisis analysis during the Cold War era, experience that the game participants did not have. For a more typical analysis of possible crisis, prior thinking might be more primitive with no reasonable starting model. Early human gaming may then be especially valuable in suggesting primary issues.

Based on subsequent discussion with other Korean and American figures, including some who have been senior officials or military officers, it is clear that in the same game but with different players, some players would have paid close attention to issues such as “what do you mean by the state of alert; who has done what so far; and which of our options are still feasible?” Also, there would have been discussion of ballistic-missile defense, a sensitive subject in South Korea. Discussion of Chinese options and preferences would have gone differently.

For subsequent work, we are considering mechanisms to enhance games of this type. These include simple decision aids to remind players of factors they might otherwise not think about, charging the teams with much more demanding questions, and more "Control Team" interventions in the midst of team deliberations.

Table 7 Important Factors in Model or Game

Model Variable (Hypothesis)	
Adversary model (Blue's Red, Red's Blue)	Yes, but best estimate prevailed
DPRK alert status	Minimal
ROK-US alert status	Built-in; shallow recognition
"Other" indicators of attack and WMD use	Acknowledged, but not assimilated
Indicators of irrationality	Acknowledged, but not assimilated
Certainty about Red intentions	Irrationality was discounted
Option feasibility govern alerts, intelligence, etc.	No
Blue cohesion	No
ROK damage-limitation capability	No
DPRK launch under attack and delegation possibilities	No
Potential of Chinese intervention	Yes, but little discussion of consequences

Summary

In summary, our initial application of the model-game-model paradigm was successful in demonstrating the feasibility and utility of using a model to (1) capture preliminary knowledge, (2) design a human game that could test or enrich that knowledge, and (3) accommodate information gained. That is, we were able to move successfully through a first cycle of the concept in Figure 1. Doing so was valuable: the human game was more interesting to players and more useful analytically because of the prior design. We also learned a good deal about how to conduct a human war game with a mix of purposes, not just improving analysis.

As it happened, the game confirmed many prior expectations and supported model constructs. It also highlighted some additional options and considerations for inclusion in an iterated model. The exercise also demonstrated shortcomings of individual human games. In our game, the players did not consider numerous factors that—in a real crisis—would surely have been part of discussions and decisionmaking, and that were incorporated in the original model. The particular game was strong on strategic discussion consistent with official thinking, but weaker in its handling of operational issues and weak in its fully addressing the "un-thinkables." Players understandably did not break away fully from prevailing beliefs and en-

gage in the rancorous debates that might occur in real Korean crisis, whether tomorrow or in five years. It was evident that different players from different organizations and perspectives would have generated different discussions and results. Even the same players would have gone down different paths with a different initiating scenario. This reaffirmed a long-recognized principle in human war gaming:

- It is crucial to plan *sets* of games, whether parallel, sequential, or both, and to define the games, recruit players, and establish details so as to probe different corners of the possibility space. Regrettably, we did not have the time or resources to do so.

Corollaries are that

- The insights from a particular game will sharpen understanding of *some* analytic aspects of the problem, but may contribute nothing on others.
- In using a model-game-model approach, a thoughtful pre-existing model will sometimes be more insightful, in some respects, than a human game.

All of this supports another long-standing principle, which has been expressed in diverse ways:

- Valid analysis will only seldom follow directly from human gaming, and results of human games should rarely be considered as "evidence" akin to empirical data. Analysts are responsible for the logic and flow of their analysis, which should be understandable and compelling in a way that transcends the particulars of the human gaming.

I would like to end by reminding the reader that, while this paper has been about a role for gaming in *analysis*, the primary value of gaming is often to those senior figures who directly participate, who quickly assimilate remarkable knowledge about the "chess board, moves, and issues," and who also achieve rapid shared understanding with other participants, sometimes establishing valuable relationships that pay off later.⁴⁸

Our experimentation with the model-game-model approach is continuing. Whereas the first-round model was informal, unverified, and undocumented, a newer version is improved as we contemplate its use in generating useful decision aids for players in human

games. Because this effort is most definitely a work in progress, comments and suggestions are especially welcome.

References

- Atran, Scott, and Robert Axelrod (2008), "In Theory: Reframing Sacred Values," *Negotiation Journal*, July, 221-45.
- Bennett, Bruce (2013), *The Challenge of North Korean Biological Weapons*, Santa Monica, Calif.: RAND Corp.
- Bracken, Paul (2012), *The Second Nuclear Age: Strategy, Danger, and the New Power Politics*, New York: Times Books.
- Caffrey, Matthew B. (2017), *On Wargaming*, Newport, Rhode Island: Naval War College Press.
- Charmaz, Kathy (2014), *Constructing Grounded Theory (Introducing Qualitative Methods Series)*, New York: SAGE Publications Ltd.
- Davis, Paul K. (1988), *The Role of Uncertainty in Assessing the NATO/Pact Central-region Balance*, Santa Monica, Calif.: RAND Corp.
- (1989), *Some Lessons Learned From Building Red Agents in the RAND Strategy Assessment System*, Santa Monica, Calif.: RAND Corp.
- (1994a) "Improving Deterrence in the Post-Cold War Era: Some Theory and Implications for Defense Planning," in *New Challenges in Defense Planning: Rethinking How Much is Enough*, edited by Paul K. Davis, Santa Monica, Calif.: RAND Corp., 197-222.
- (1994b), *New Challenges in Defense Planning: Rethinking How Much is Enough*, Santa Monica, Calif.: RAND Corp.
- (2012), *Some Lessons From RAND's Work on Planning Under Uncertainty for National Security*, Santa Monica Calif.: RAND Corp.
- Davis, Paul K., Steven C. Bankes, and James P. Kahan (1986), *A New Methodology for Modeling National Command Level Decisionmaking in War Games and Simulations*, Santa Monica, Calif.: RAND Corp.
- Davis, Paul K. et al. (2008), *Developing Resource-Informed Strategic Assessments and Recommendations*, Santa Monica, Calif.: RAND Corp.
- Davis, Paul K., and Amy Henninger (2007), *Analysis, Analysis Practices, and Implications for Modeling and Simulation*, Santa Monica, Calif.: RAND Corp., OP-176-OSD.
- Davis, Paul K. et al. (2016), "Deterrence and Stability for the Korean Peninsula," *Korean Journal of Defense Analyses*, 28(1), 1-23.
- Davis, Paul K., Jonathan Kulick, and Michael Egner (2005), *Implications of Modern Decision Science for Military Decision Support Systems*, Santa Monica, Calif.: RAND Corp.

- Davis, Paul K. et al. (2012), *Understanding and Influencing Public Support for Insurgency and Terrorism*, Santa Monica, Calif.: RAND Corp.
- Davis, Paul K., and Angela O'Mahony (2013), *A Computational Model of Public Support for Insurgency and Terrorism: a Prototype for More General Social-Science Modeling*, Santa Monica, Calif.: RAND Corp.
- Davis, Paul K., and Angela O'Mahony (2017), "Representing Qualitative Social Science in Computational Models to Aid Reasoning Under Uncertainty: National Security Examples," *Journal of Defense Modeling and Simulation*, 1-22.
- Davis, Paul K., and James A. Winnefeld (1983), *The RAND Corp. Strategy Assessment Center*, Santa Monica, Calif.: RAND.
- Delpuch, Therese (2012), *Nuclear Deterrence in the 21st Century : Lessons From the Cold War for a New Era of Strategic Piracy*, Santa Monica, CA: RAND.
- Downes-Martin, Stephe (2013), "Adjudication: the *Diabolus in Machina* of War Gaming," *Naval War College Review*, 66(3), 67-80.
- Fox, Daniel B., and Carl M. Jones (1998), *JICM 3.0: Documentation and Tutorial*, Santa Monica, Calif.: RAND.
- George, Alexander L., and Andrew Bennett (2005), *Case Studies and Theory Development in the Social Sciences*, Cambridge, Mass.: MIT Press.
- Graubard, Morlie H., and Carl H. Builder (1982), "New Methods for Strategic: Automating the War-game," *Policy Sciences*, 71-84.
- Hanley, John T. (2017), "Changing DoD's Analysis Paradigm: the Science of War Gaming and Combat/Campaign Simulation," 70(1), 64-103.
- Jervis, Robert, Richard.N. Lebow, and Janet G. Stein (1985), *Psychology and Deterrence*, Baltimore: Johns Hopkins University Press.
- Kahn, Herman (1960), *On Thermonuclear War*, Princeton University Press.
- (1965), *On Escalation: Metaphors and Scenarios*, Westport, Conn.: Greenwood Press.
- Kahneman, Daniel (2011), *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux.
- Kissinger, Henry (1957), *Nuclear Weapons and Foreign Policy*, New York: Council on Foreign Relations.
- Levine, Robert, Thomas C. Schelling, and William Jones (1991), *Crisis Games 27 Years Later: Plus C'est Déjà Vu*, Santa Monica, Calif.: RAND Corp.
- Mingers, John, and Jonathan Rosenhead (eds.) (2011), "Special Issue: Teaching Soft O.R., Problem Structuring Methods, and Multimethodology," *INFORMS Transactions on Education*, 12(1), September, 1 ff.
- Mingers, John, and Jonathan Rosenhead (2004), "Invited Review: Problem Structuring Methods in Action," *European Journal of Operational Research*, 152.

- National Research Council and Naval Studies Board (1996), *Post-Cold War Conflict Deterrence*, Washington, D.C.: National Academy Press.
- National Research Council (2014), *U.S. Air Force Strategic Deterrence Analytic Capabilities: An Assessment of Methods, Tools, and Approaches for the 21st Century Security Environment*, Washington, D.C.: National Academies Press.
- Olabisi, Laura K. Schmitt et al. (2010), “Using Scenario Visioning and Participatory System Dynamics Modeling To Investigate the Future: Lessons from Minnesota 2050,” *Sustainability*, 2(8).
- Pellerin, Cheryl (2015), “Work: Wargaming Critical in Dynamic Security Environment,” <https://www.defense.gov/News/Article/Article/633892/work-wargaming-critical-in-dynamic-security-environment>, (last accessed,
- Perla, Peter P. (1990), *The Art of Wargaming: A Guide for Professionals and Hobbyists*, US Naval Institute Press.
- Perla, Peter P., and Ed McGrady (2009), *Systems Thinking and Wargaming*, Alexandria, Va.: Center for Naval Analyses.
- Pournelle, Phillip (2016), “MORS Wargaming Workshop,” *Phalanx*, 49 (4), December.
- Pournelle, Phillip, and Yuna Huh Wong (2017), “MORS Wargaming Special Meeting October 2016, Final Report,”.
- Pournelle, Phillip E. (2014), “Preparing for War, Keeping the Peace,” *Proceedings Magazine*, 140(9), 339.
- Ritchey, Tom (2011), *Social Messes: Decision Support Modeling With Morphological Analysis*, Springer.
- Roberts, Bradley H. (2015), *The Case for U.S. Nuclear Weapons in the 21st Century*, Redwood City, Calif.: Stanford University Press.
- Rosenhead, Jonathan, and John Mingers (2002a) “A New Paradigm of Analysis,” in *Rational Analysis or a Problematic World Revisited: Problem Structuring Methods for Complexity, Uncertainty and Conflict*, edited by Jonathan Rosenhead, and John Mingers, Chichester, UK: John Wiley & Sons, Inc., 1-19.
- , (eds.) (2002b), *Rational Analysis for a Problematic World Revisited: Problem Structuring Methods for Complexity, Uncertainty and Conflict, 2nd Edition*, New York: John Wiley & Sons.
- Rubel, Robert C. (2006), “The Epistemology of War Gaming,” *Naval War College Review*, 59(2), 108-28.
- SAS-087 (2012), “NATO Guide for Judgement-Based Operational Analysis in Defense Decision Making: Code of Best Practices,” RTO-TR-SAS-087.
- Schelling, Thomas C. (1987) “The Role of War Games and Exercises,” in *Managing Nuclear Operations*, edited by Ashton B. Carter, John D. Steinbruner, and Charles A. Zraket, Washington, D.C.: Brookings, 426-44.

Schwabe, William L. (1992), "Analytic War Plans: Adaptive Force-employment Logic in the RAND Corp. Strategy Assessment System (RSAS)," N-3051-OSD.

Shlapak, David A. et al. (1986), *The RAND Strategy Assessment System's Green Agent Model of Third-country Behavior in Superpower Crises and Conflict*, Santa Monica, Calif.: RAND Corp.

Sik, Kim Hyun (2009), "The Secret History of Kim Jong Il," *Foreign Policy*, October.

Simon, Herbert A. (1978), "Nobel Prize Lecture: Rational Decision-Making in Business Organizations," http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1978/simon-lecture.html, (last accessed, February 3, 2014).

Wong, Yuna Huh (2016), "How Can Gaming Help Test Your Theory?," <https://paxsims.wordpress.com/2016/05/18/how-can-gaming-help-test-your-theory/>, (last accessed, February 5, 2017).

Yin, Robert K. (2015), *Qualitative Research From Start to Finish, Second Edition*, New York: The Guilford Press.

Notes

¹The conference was *Wargaming*, a special meeting of the Military Operations Research Society (MORS), 17-20 October, 2016, in Alexandria, Virginia. A brief overview has been published (Pournelle, 2016) and a fuller report is being prepared by the co-chairs Commander Phillip Pournelle and Dr. Yuna Wong (Pournelle and Wong, 2017).

²Robert Work, memorandum of May 8, 2015, as discussed in a DoD news posting. (Pellerin, 2015)

³ Perla, 1990.

⁴These topics, similar to categories identified in Deputy Secretary Work's memorandum (footnote 3) were discussed in subgroups of the workshop led by James Bexfield, Michale Ottenberg, and Matt Caffrey, respectively.

⁵The workshop report has a related but somewhat different but largely consistent list in the section on results of Working Group 1 (Pournelle and Wong, 2017).

⁶As examples, Yuna Wong discusses how gaming can and cannot help test a theory (Wong, 2016); Stephen Downes-Martin discusses how "Control Teams" do not just objectively adjudicate but in practice are often dominant de facto players, skewing results (Downes-Martin, 2013); Robert Rubel describes the strength and weaknesses of wargaming broadly (Rubel, 2006) and, in doing so, draws on 1991 dissertation work of John Hanley at Yale, which describes wargaming as a weakly structured tool appropriate to weakly structured problems—one that generates weakly structured knowledge, which is conditional and subject to judgment in application (p. 112). In a very recent paper John Hanley lists numerous virtues of wargaming, but *none* of them are what would normally be regarded as proving validity or calibrating parameters (Hanley, 2017).

⁷The dictionary definitions of "rigor" are revealing, with words like "harsh inflexibility" and "being unyielding or inflexible," or "strict precision." (Merriam-Webster on-line dictionary). It is ironic that in the realm of analysis, "rigor" usually has a *positive* connotation associated with carefulness and logic. To be sure, we want analysis to have those characteristics, but not if it means careful logic based on foolish notions (such as

notions that ignore qualitative considerations that humans find crucial.

⁸See discussion of wicked problems and of "soft operations research" in various European papers and books (Mingers and Rosenhead, 2004; Rosenhead and Mingers, 2002b; Mingers and Rosenhead (eds.), 2011). NATO has a related best-practices manual (SAS-087, 2012). Some American authors deal with these issues and use qualitative methods (Ritchey, 2011). Colleagues and I use a related factor-tree methodology (Davis et al., 2012) and extensions to a semi-quantitative computational version (Davis and O'Mahony, 2013)

⁹The primacy of models is broader than sometimes recognized. NASA, in sending astronauts to the moon, and the U.S. military when developing exquisitely accurate ICBMs and SLBMs, has depended fundamentally on models. Flight tests, ultimately, were to validate, sharpen, and calibrate those models.

¹⁰Versions of the graphic appear in earlier publications National Research Council, 2014; Davis and Henninger, 2007.

¹¹The most profound value of human crisis gaming is probably in educating, communicating, and building networks of shared knowledge among senior *participants*. See essays by Thomas Schelling and William Jones in Levine et al., 1991, essays written in 1964. This conclusion about primary value has been corroborated repeatedly in think-tank games and even in games held for senior leaders in the NSC.

¹²Several authors discuss the cycle of research, although usually not in much detail (Perla, 1990; Perla and McGrady, 2009; Pournelle, 2014).

¹³Qualitative theory and grounded theory are discussed in some modern texts (Yin, 2015; Charmaz, 2014).

¹⁴This important point has been made for some years by some European authors who discuss "wicked problems" and "soft operations research" (Rosenhead and Mingers, 2002a; Mingers and Rosenhead, 2004). Participatory modeling has begun to catch on in the United States as well (Olabisi et al., 2010)

¹⁵I thank colleague Yuna Wong for pointing out some of these references.

¹⁶Important exceptions exist, cases in which senior leaders participated or observed from "behind the screen" (Bracken, 2012).

¹⁷A spirited debate about human war gaming occurred fifty years ago. A paper by Robert Levine is strongly negative. Papers by Thomas Schelling and William Jones are highly positive (Levine et al., 1991).

¹⁸A similar philosophy is reflected in a recent paper by Yuna Wong (Wong, 2016)

¹⁹See Schelling, 1987.

²⁰See Levine et al., 1991.

²¹Human games on counterinsurgency, for example, can discredit models that dwell on attrition and ignore attitudes of population.

²²An interesting paper by Stephen Downes-Martin points out that players in discovery games can only reveal their beliefs; he also emphasizes that the Control team, responsible for game adjudication, are often more like participants than objective observers (Downes-Martin, 2013).

²³See Caffrey, 2017.

²⁴See Levine et al., 1991, Wong, 2016 and Downes-Martin, 2013.

²⁵The method of comparative case studies was pioneered by Alexander George (George and Bennett, 2005)

²⁶This "validation" of factors in a qualitative model is a theme of a recent study (Davis et al., 2012)

²⁷For discussion of sacred values, see Atran and Axelrod, 2008.

²⁸The KIDA team was led by Dr. Jaehun Lee, under the general direction of Dr. Yuntae Kim. Team members included Mr. Junho Park and Ms. Sunhee Lee. The KIDA team developed and conducted the human game.

²⁹This is taken from a collaboratively written journal article (Davis et al., 2016). The modern challenges of extended deterrence are also discussed in a recent book (Roberts, 2015).

³⁰For a review of extended deterrence and related debates, see Delpech, 2012, p.30 ff.

³¹The model in Figure 2 became today's JICM (Joint Integrated Contingency Model), which is partly

documented in the public domain (Fox and Jones, 1998).

³²Although not emphasized, the interchangeability of human teams and automated agents was present in RAND's initial 1980 proposal for an automated wargaming system (Graubard and Builder, 1982). The actual RSAS, development of which I led, took a different approach in many respects and emphasized going back and forth between human and automated play, although not with *large* wargames with a mix of human players and agents (Davis and Winnefeld, 1983). We focused on analysis rather than big games, and our man-versus-machine activity typically involved only one or two analysts at a time as they sought to improve and tighten the simulation to support analysis.

³³The RSAS Agents were different in character from the agents of today's agent-based modeling (ABM), the technology for which did not exist. Documentation exists on the Red and Blue agents (Davis et al., 1986), the Green Agent (Shlapak et al., 1986), the military-level agents (Schwabe, 1992), and on some lessons learned (Davis, 1989).

³⁴Multiscenario analysis was introduced to describe the Cold War conventional balance in Europe (Davis, 1988). I later referred to the method as *exploratory analysis* (Davis, 1994b; Davis, 2012).

³⁵See Kahn, 1960; Kahn, 1965.

³⁶See Kissinger, 1957.

³⁷See Kissinger, 1957 and Delpech, 2012.

³⁸See also a study by colleague Bruce Bennett (Bennett, 2013)

³⁹See Sik, 2009. Sik was once a professor in North Korea with Kim Jong Il as student.

⁴⁰A good introduction to Simon's thinking is his Nobel Prize address (Simon, 1978).

⁴¹This approach was first used to understand Saddam Hussein before and during his invasion of Kuwait in 1990, and in thinking about deterrence in the post Cold war era (Davis, 1994a) Excerpts were included in a later national-academy volume (National Research Council and Naval Studies Board, 1996).

⁴²See especially Jervis et al., 1985 for strategic implications and Kahneman, 2011 for a recent summary of decades' worth of research in human psychology.

⁴³See Davis et al., 2005

⁴⁴An analogue occurs in an approach to strategic portfolio analysis that urges recognition of alternative "strategic perspectives" affecting both the weights of criteria and how options rate (Davis et al., 2008).

⁴⁵Based on other research, I expect distinctly non linear combination rules, which can nonetheless be approximated simply, as illustrated elsewhere (Davis and O'Mahony, 2013).

⁴⁶A mature version of the figure would be an example of *factor tree* methodology, first introduced in a study of social science for counterterrorism and then refined in subsequent studies cited in Davis and O'Mahony, 2017.

⁴⁷For the purposes of testing and enriching analysis, this vignette-exercise approach appeared to be more efficient and discerning than a single classic game in which teams take moves, the Control Team adjudicates results, and the teams continue through the course of a single scenario. That might not always be the case, however.

⁴⁸See comments by Thomas Schelling in Levine et al., 1991.