A (Very) Brief Introduction to (War)Gaming

Prof. Andrew W. Reddie
areddie@berkeley.edu; @areddie89
University of California, Berkeley
Goldman School of Public Policy

Berkeley
UNIVERSITY OF CALIFORNIA

Why are we here?

We have lots of **questions…**

- Will **artificial intelligence** technologies revolutionize warfare?
- Can the threat of **cyber attacks** be used to deter adversaries?
- How might the proliferation of **hypersonic** weapons alter strategic stability?

Why are we here?

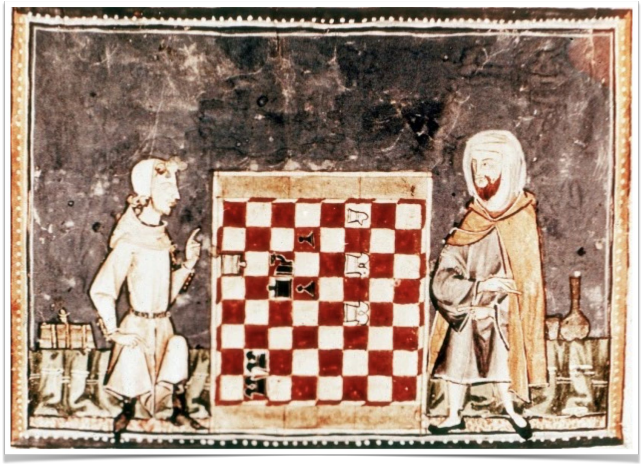
- Will **artificial intelligence** technologies revolutionize warfare?
- Can the threat of **cyber attacks** be used to deter adversaries?
- How might the proliferation of **hypersonic** weapons alter strategic stability?

We have lots of **questions…**

And few **answers…**

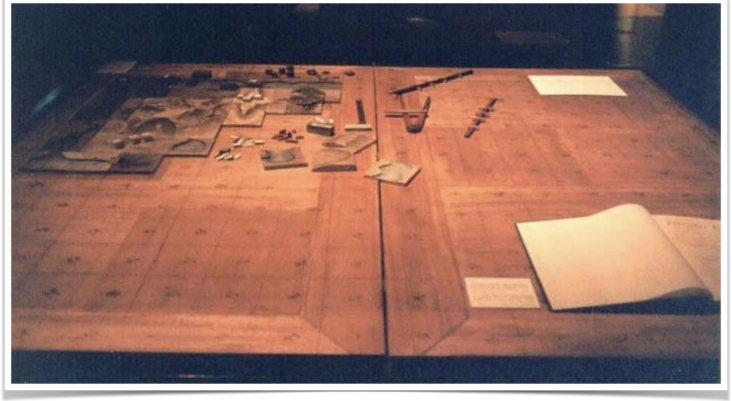
The Long History of Wargames

- Wargames appear in antiquity as early as the 5th century BCE:
  - *petteia* (“pebbles”) or *ludus latrunculorum* (“game of little soldiers”)
  - Central to both Greek and Roman military education… and entertainment
- By the 6th century AD, we have c*haturunga*—better known as…
- These strategy games were primarily used for military training—something that their successors games have also been designed around…



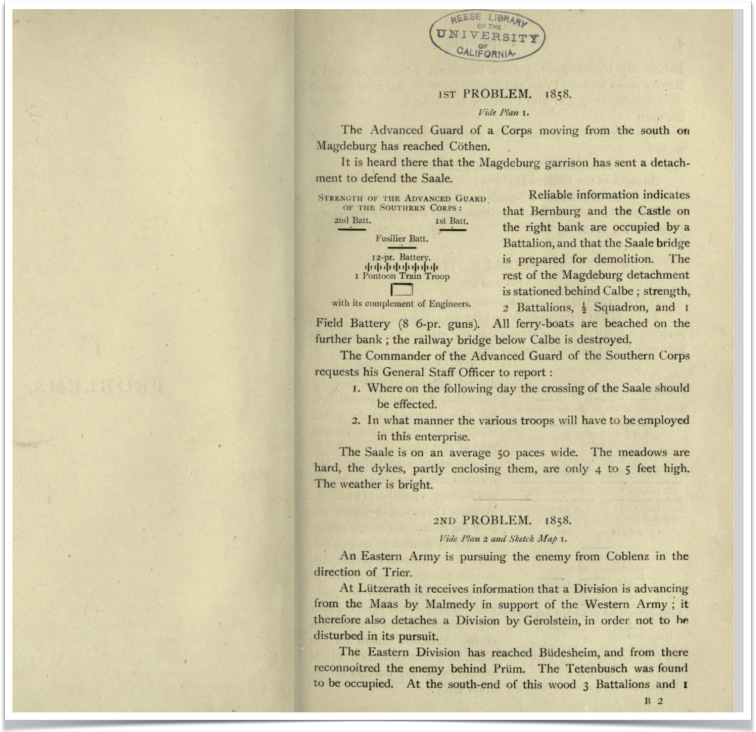
German Origins: Reisswitz, Moltke, and Kriegsspiel

- Against the backdrop of military officers manipulating chess, a Prussian officer (Georg von Reisswitz) develops a map-based war-game using varied pieces to denote unit type (infantry; cavalry) and using data from the Napoleonic wars to model unit effects
  - “Hitpoint tracking”
  - No clear “win conditions”
  - “Free Kriegsspiel” (the White Cell)
- Came to be known as *Kriegsspiel*, German for…



German Origins

- Reisswitz's design became a central aspect of Prussian military training
- Institutionalized by Helmuth von Moltke the Elder (Prussian field marshal)
  - Also a fan of "map-games"
- Victory in the Franco-Prussian War of 1870 was attributed to wargaming
  - Spreads to the Kriegsspiel Club at Oxford University in 1873, and to the United States in the 1880s (more on this later)



1ST PROBLEM. 1858.

Vide Plan 1.

The Advanced Guard of a Corps moving from the south on Magdeburg has reached Cöthen.

It is heard there that the Magdeburg garrison has sent a detachment to defend the Saale.

STRENGTH OF THE ADVANCED GUARD OF THE SOUTHERN CORPS:

2nd Batt. 1st Batt.

Fusilier Batt.

12-pr. Battery.

1 Pontoon Train Troop

with its complement of Engineers.

Reliable information indicates that Bernburg and the Castle on the right bank are occupied by a Battalion, and that the Saale bridge is prepared for demolition. The rest of the Magdeburg detachment is stationed behind Calbe; strength, 2 Battalions, ½ Squadron, and 1 Field Battery (8 6-pr. guns). All ferry-boats are beached on the further bank; the railway bridge below Calbe is destroyed.

The Commander of the Advanced Guard of the Southern Corps requests his General Staff Officer to report:

1. Where on the following day the crossing of the Saale should be effected.
2. In what manner the various troops will have to be employed in this enterprise.

The Saale is on an average 50 paces wide. The meadows are hard, the dykes, partly enclosing them, are only 4 to 5 feet high. The weather is bright.

2ND PROBLEM. 1858.

Vide Plan 2 and Sketch Map 1.

An Eastern Army is pursuing the enemy from Coblenz in the direction of Trier.

At Lützerath it receives information that a Division is advancing from the Maas by Malmedy in support of the Western Army; it therefore also detaches a Division by Gerolstein, in order not to be disturbed in its pursuit.

The Eastern Division has reached Büdesheim, and from there reconnoitred the enemy behind Prüm. The Tetenbusch was found to be occupied. At the south-end of this wood 3 Battalions and 1

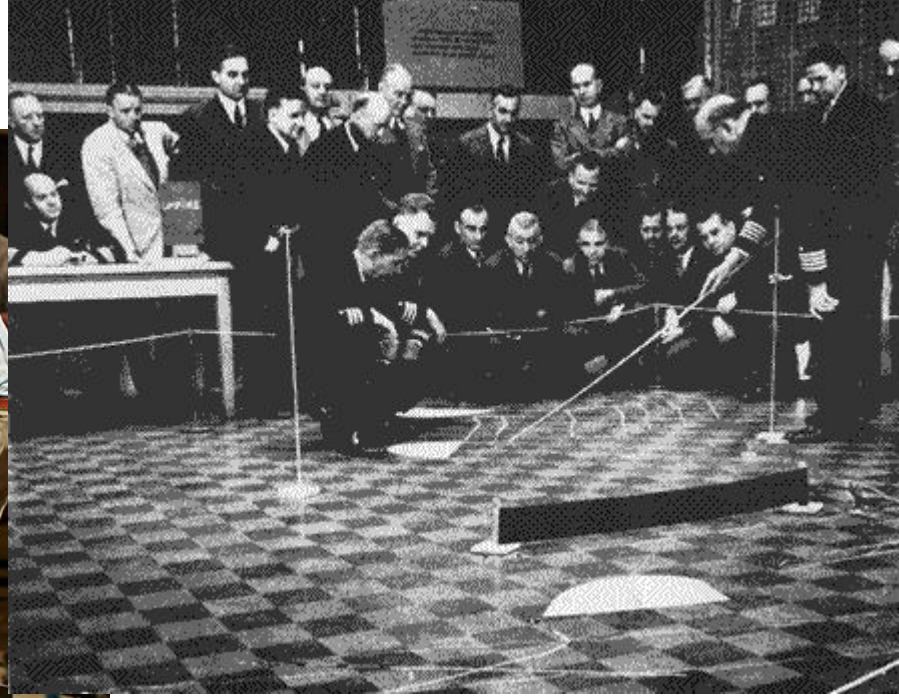
B 2

German Origins

- Reisswitz's design became a central aspect of Prussian military training
- Institutionalized by Helmuth von Moltke the Elder (Prussian field marshal)
  - Also a fan of "map-games"
- Victory in the Franco-Prussian War of 1870 was attributed to wargaming
  - Spreads to the Kriegsspiel Club at Oxford University in 1873, and to the United States in the 1880s (more on this later)

The Naval War College

- Lt. William McCarty Little introduces wargaming at the NWC in 1886 before retiring in 1915
- Wargaming at the NWC was separated into tactics (“board maneuvers”) and strategy (“chart maneuvers”)
- For example, Tactical Game 94 of 1923 demonstrates the importance of reconnaissance and detection of the enemy’s forces
- ORANGE vs. Japan; RED vs. UK
- Adm. Nimitz: Wargames predicted every tactic of the Japanese outside of kamikaze pilots
    - Submarine doctrine…



Wargaming and The Cold War

- A turn toward wargaming as an analytical tool
  - Expansion beyond military science
  - “Political-military” games
  - Seminar, turn, and team-based design with “Control” adjudication between “Blue” and “Red”
- “to simulate the novel problem of fighting a limited war under the shadow of nuclear escalation”
- Examples: RAND; Bloomfield-Schelling DOD games

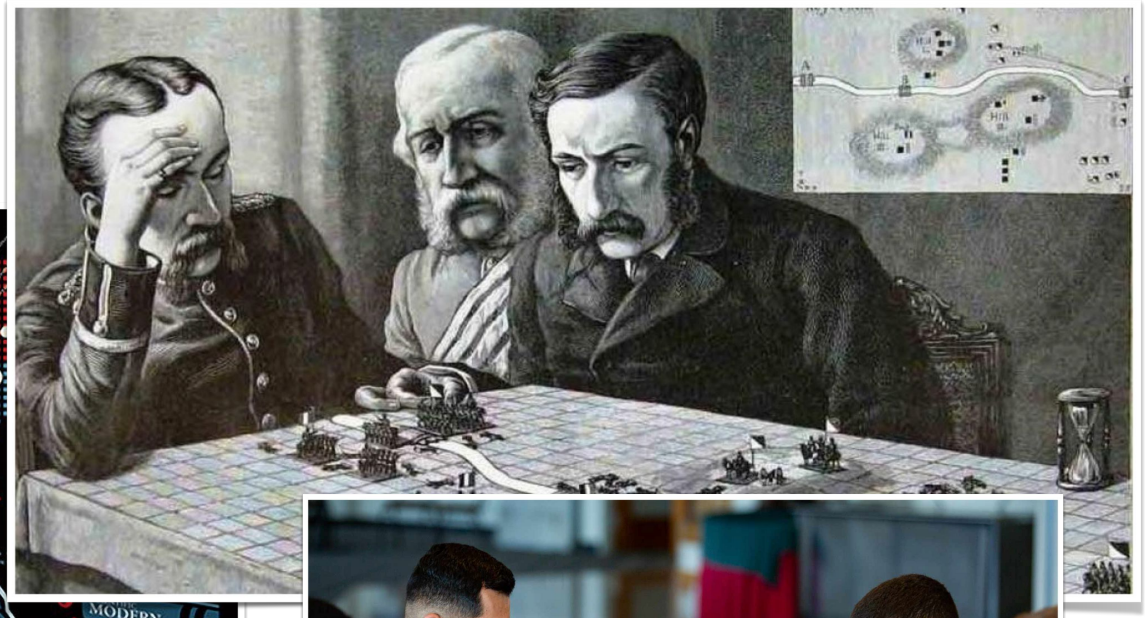
Table 1. Wargames with Strategic Elite Players

Nuclear-Armed Adversaries	Nonnuclear-Armed Adversaries
POLEX I, MIT, September 1958	SIGMA I-62, JWGCG, February 1962
POLEX II, MIT, September 1960	POLEX-DAIS I, MIT, November 1962*
Berlin crisis game I, JWGCG, September 1961	POLEX-DAIS II, MIT, January 1963*
Berlin crisis game II, JWGCG, September–October 1961	POLEX-DAIS III, MIT, February 1963*
EPSILON I-62, JWGCG, September 1962	POLEX-DAIS IV, MIT, March 1963*
DETEX II, MIT, February 1964	DETEX I, MIT, November 1963*
EPSILON I-64, JWGA, April 1964	SIGMA I-64, JWGA, April 1964
DETEX III, MIT, November 1964	SIGMA II-64, JWGA, September 1964
NU I, JWGA, January–February 1966	SIGMA II-65, JWGA, July–August 1965*
NU II, JWGA, January–February 1966	SIGMA I-66, JWGA, September 1966*
BETA I, JWGA, April-May 1967	SIGMA I-67, JWGA, November–December 1967
BETA II, JWGA, April-May 1967	SIGMA II-67, JWGA, November–December 1967
EPSILON 72, SAGA, October 1972	MU I-68, JWGA, April–May 1968*

* Deterrence logic may still operate via nuclear-armed third-party extended deterrence.

NOTE: MIT stands for Massachusetts Institute of Technology; JWGCG stands for Joint War Games Control Group; JWGA stands for Joint War Games Agency; and SAGA stands for Studies, Analysis, and Gaming Agency.

Wargaming Today…



Berkeley
UNIVERSITY OF CALIFORNIA

How Should We Study Emerging Capabilities?

As researchers/policy analysts/academics, we are often faced with a difficult challenge:

- How do you study a phenomenon that has yet to occur?
  - We can't use empirical/case-based observational methods

Synthetic data-generating processes:

- Formal models and subsequent computer-based simulation
- Survey experiments

A Note on branding…

How did we get here?

"Successfully navigating through this complex and dynamic competitive environment will once again require us to push the boundaries of technology while ensuring that innovation remains rooted in operationally realistic doctrine and capabilities. One way to do both is to re-prime and re-stoke the department's wargaming engine."

REVITALIZING WARGAMING IS NECESSARY TO BE PREPARED FOR FUTURE WARS

BOB WORK AND GEN. PAUL SELVA

COMMENTARY

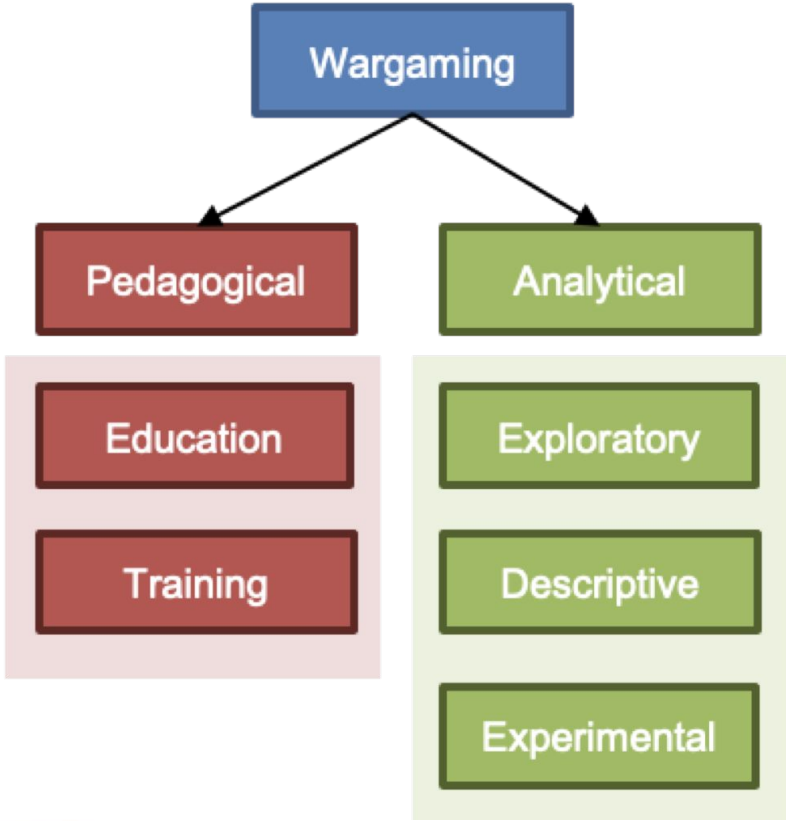
DECEMBER 8, 2015

Applications of (War)gaming Methods

Wargames are particularly useful for thinking about the **tactical**, **operational**, and **strategic** aspects of conflict [AND other behavioral phenomenon that we might be interested in]…

And this has defined the use of wargaming methods across three contexts:

- Education
- Training
- Analysis

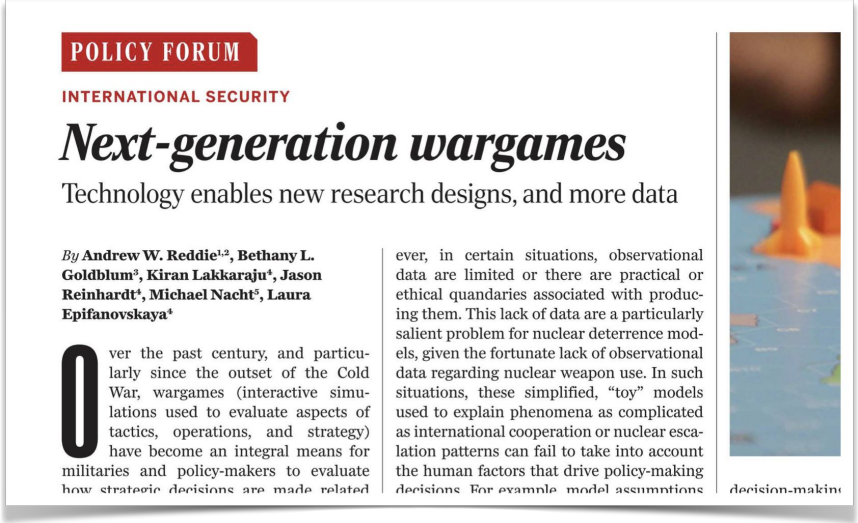


Bringing “Science” to the “Art of Wargaming”

Generally, **traditional wargame designs** are useful for exploring a problem but face clear (and we thought obvious) analytical challenges…

- **Analytical Concerns:**
  - Replication challenges
  - Sample bias
  - White cell adjudication
  - Sponsorship bias

- **Data Issues:**
  - Data collection is often incidental
  - Rapporteurs
  - Data is cloistered
  - Wargame designs are often treated as intellectual property



POLICY FORUM

INTERNATIONAL SECURITY

Next-generation wargames

Technology enables new research designs, and more data

By **Andrew W. Reddie[1,2], Bethany L. Goldblum[3], Kiran Lakkaraju[4], Jason Reinhardt[4], Michael Nacht[5], Laura Epifanovskaya[4]**

Over the past century, and particularly since the outset of the Cold War, wargames (interactive simulations used to evaluate aspects of tactics, operations, and strategy) have become an integral means for militaries and policy-makers to evaluate how strategic decisions are made related

ever, in certain situations, observational data are limited or there are practical or ethical quandaries associated with producing them. This lack of data are a particularly salient problem for nuclear deterrence models, given the fortunate lack of observational data regarding nuclear weapon use. In such situations, these simplified, “toy” models used to explain phenomena as complicated as international cooperation or nuclear escalation patterns can fail to take into account the human factors that drive policy-making decisions. For example, model assumptions

decision-makin

… and we are only beginning to understand these **laboratory effects**

Experimental Gaming Methods

Value Proposition:

- Addresses the limits of existing gaming methods related to causal inference, generalizability, and replicability
- Provides an additional data generating process for theory development and testing (particularly where observational data is limited or non-existent)
- Adds a new type of experimental tool for social science research
- Addresses the complexity of contemporary security environments

017; Lin-Greenberg 2018; Reddie et al. 2018

My Team: Bringing Science to Wargaming

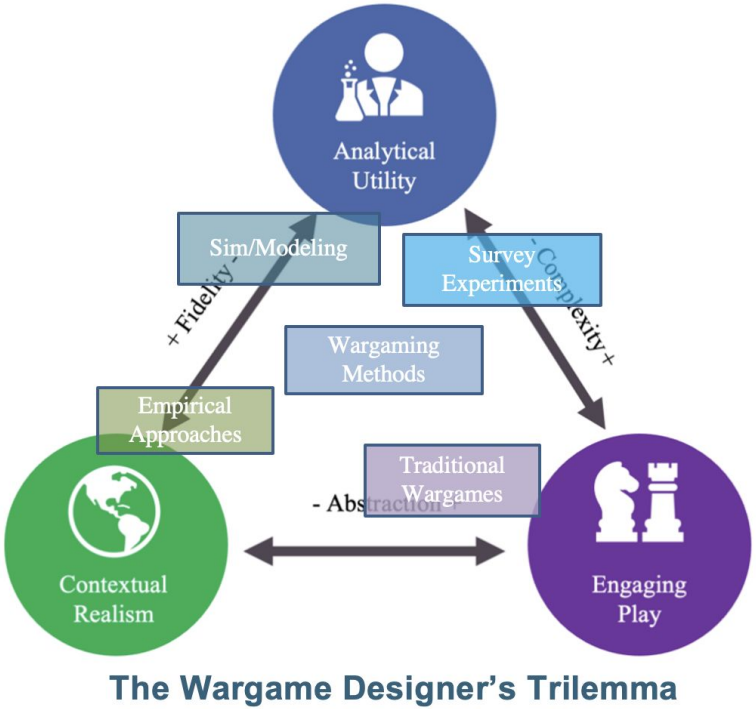
Always starts with the RQ (yields DV(s), IVs, and hypotheses)

1. Experimental design
2. With an emphasis on hypothesis testing
3. Game “elements” that reflect this design
    a. Where we borrow from the “art of wargaming”
4. Alpha version (resources, counters, cards)
5. Beta testing (balancing the trilemma)…
6. Fielding
7. Analysis

Synthetic DGPs in Context

Game design involves making hard (and imperfect) choices…

- How many players are you going to include?
- What type of map should you use?
- Should you use “real” countries?
- How complex (how “real”) should you make the game?
- Does medium matter?
- Who gets to play (your sample)?
- How do you know the results are representative?



The Wargame Designer’s Trilemma

An Example: SIGNAL

SIGNAL TTX, SIGNAL Board SIGNAL Online

- Abstract “states” (Orange, Purple, Green)
- 3-player (with nuclear dyads) game design
- Win Conditions
  - Economic
  - Security
- N-round games
  - Signaling phase *cost
  - Action phase *cost
  - Upkeep phase
- Instrumented for data collection


NORTH SEA
SOUTH SEA


Round
1 of 5
Changes Next Round
Phase
Signaling
20
+11
0
+0
1
+0
3
+0
1
+0
9/12
+0/+0
doqgWUhXbm
NATIONS
0/12
10/12
0/12
Your Rank:
2nd
TRADE
PROPOSALS
Prepare, Time left: 03:06
CHAT & FEED
G - O (Private): Can we negotiate a trade deal to avoid conflict?
Enter text...
SEND
READY

Other Examples:

Data Products

At base, we treat data from wargames in a similar manner to data from survey work…

- Units of analysis:
  - Game
  - Player-Game
  - Player-Round
  - Sample-Player-Game
- Methods:
  - Regression analysis, survival analysis, additional ML methods.

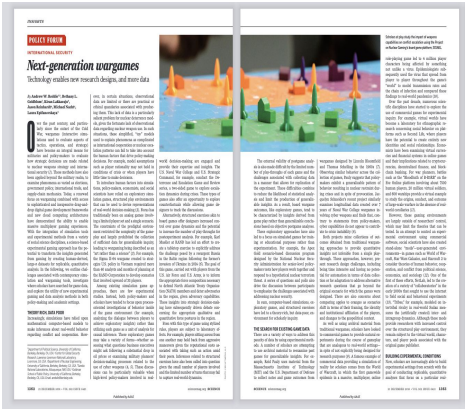
Table 6: Logit models predicting the probability of nuclear use and nuclear first use using the player as the unit of analysis. The values in the table body display the regression coefficients with standard error in parentheses.

	Dependent variable:			
	Nuclear First Use		Nuclear Use	
	(9)	(10)	(11)	(12)
Tailored	0.009	−0.021	0.021	0.017
	(0.143)	(0.144)	(0.154)	(0.156)
Female		−0.334**		−0.119
		(0.141)		(0.146)
College Degree		−0.163		−0.018
		(0.168)		(0.181)
Age > 29		0.078		−0.095
		(0.160)		(0.173)
National Security		0.062		−0.005
		(0.214)		(0.227)
More Conservative		−0.110		0.020
		(0.150)		(0.161)
Reported Knowledge		−0.303*		−0.147
		(0.171)		(0.181)
Constant	−0.560	−0.197	0.981	1.133
	(0.102)	(0.178)	(0.110)	(0.194)
Observations	850	843	850	843
Log Likelihood	−557.664	−547.565	−496.257	−491.282
Note:				*p<0.1; **p<0.05; ***p<0.01

*Wargame data also capture strategic interaction that surveys do not.

Open Questions

- **Laboratory Effects and Game Design**
  - Design concerns (abstract, # of players)
  - Game medium (digital vs. analog)
  - Sampling: “Elite” vs. “non-elite” play (Kertzer et al. 2022)
- **Methodological Extensions**
  - Human-machine play and experimentation
  - The wargame designer’s “sandbox”
  - Combining behavioral and physics-based models in gameplay
- **Substantive Areas for Future Research**
  - Nuclear proliferation (“Left of SIGNAL”)
  - Cyber deterrence
  - Alliance/Wedging dynamics



POLICY FORUM

INTERNATIONAL SECURITY

Next-generation wargames

Technology enables new research designs, and more data

Some More Reading…

Emery, John R. "Moral Choices Without Moral Language: 1950s Political-Military Wargaming at the RAND Corporation (Fall 2021)." *Texas National Security Review* (2021).

Lin-Greenberg, Erik, Reid BC Pauly, and Jacquelyn G. Schneider. "Wargaming for international relations research." *European Journal of International Relations* 28, no. 1 (2022): 83-109.

Reddie, Andrew W. and Bethany L. Goldblum (2023). Integrating the Art and Science of Wargaming, Lawfare

Reddie, Andrew W., and Bethany L. Goldblum. "Evidence of the unthinkable: Experimental wargaming at the nuclear threshold." *Journal of Peace Research* 60, no. 5 (2023): 760-776.

Schechter, Benjamin, Jacquelyn Schneider, and Rachael Shaffer. "Wargaming as a methodology: the international crisis wargame and experimental wargaming." *Simulation & Gaming* 52, no. 4 (2021): 513-526.

Thank You!

areddie@berkeley.edu;
@areddie89

Next-Generation Wargames

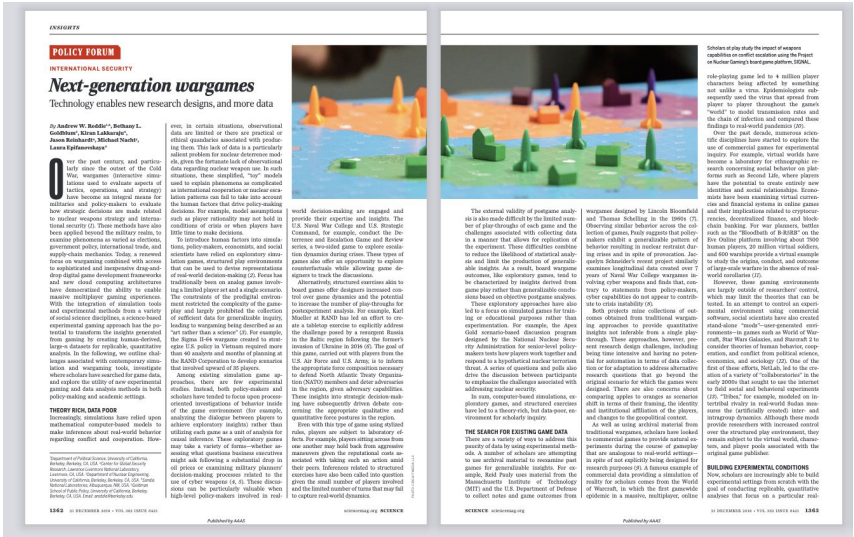
Many of the questions that we seek to answer as analysts of national and international security importance suffer from a dearth of empirical data…

Examples include considering the effects of emerging capabilities and/or new strategies…

How should we, as analysts, think about addressing this challenge?

- Synthetic data-generating processes (DGPs)

The example of the tailored nuclear weapons debate (*Science* and forthcoming in the *Journal of Peace Research*)



INSIGHTS

POLICY FORUM

INTERNATIONAL SECURITY

Next-generation wargames

Technology enables new research designs, and more data

By Andrew W. Reddie[1]*, Bethany L. Goldblum[2], Kiran Lakkaraju[3], Jason Reinhardt[3], Michael Nacht[1], Laura Epifanovskaya[4]

Over the past century, and particularly since the outset of the Cold War, wargames (interactive simulations used to evaluate aspects of tactics, operations, and strategy) have become an integral means for militaries and policy-makers to evaluate how strategic decisions are made related to nuclear weapons strategy and international security (1). These methods have also been applied beyond the military realm, to examine phenomena as varied as elections, government policy, international trade, and supply-chain mechanics. Today, a renewed focus on wargaming combined with access to sophisticated and inexpensive drag-and-drop digital game development frameworks and new cloud computing architectures have democratized the ability to enable massive multiplayer gaming experiences. With the integration of simulation tools and experimental methods from a variety of social science disciplines, a science-based experimental gaming approach has the potential to transform the insights generated from gaming by creating human-derived, large-n datasets for replicable, quantitative analysis. In the following, we outline challenges associated with contemporary simulation and wargaming tools, investigate where scholars have searched for game data, and explore the utility of new experimental gaming and data analysis methods in both policy-making and academic settings.

THEORY RICH, DATA POOR

Increasingly, simulations have relied upon mathematical computer-based models to make inferences about real-world behavior regarding conflict and cooperation. However, in certain situations, observational data are limited or there are practical or ethical quandaries associated with producing them. This lack of data is a particularly salient problem for nuclear deterrence models, given the fortunate lack of observational data regarding nuclear weapon use. In such situations, these simplified, "toy" models used to explain phenomena as complicated as international cooperation or nuclear escalation patterns can fail to take into account the human factors that drive policy-making decisions. For example, model assumptions such as player rationality may not hold in conditions of crisis or when players have little time to make decisions.

To introduce human factors into simulations, policy-makers, economists, and social scientists have relied on exploratory simulation games, structured play environments that can be used to derive representations of real-world decision-making (2). Focus has traditionally been on analog games involving a limited player set and a single scenario. The constraints of the prodigital environment restricted the complexity of the gameplay and largely prohibited the collection of sufficient data for generalizable inquiry, leading to wargaming being described as an "art rather than a science" (3). For example, the Sigma II-64 wargame created to strategize U.S. policy in Vietnam required more than 80 analysts and months of planning at the RAND Corporation to develop scenarios that involved upward of 20 players.

Among existing simulation game approaches, there are few experimental studies. Instead, both policy-makers and scholars have tended to focus upon process-oriented investigations of behavior inside of the game environment (for example, analyzing the dialogue between players to achieve exploratory insights) rather than utilizing each game as a unit of analysis for causal inference. These exploratory games may take a variety of forms—whether assessing what questions business executives might ask following a substantial drop in oil prices or examining military planners' decision-making processes related to the use of cyber weapons (4, 5). These discussions can be particularly valuable when high-level policy-makers involved in real-world decision-making are engaged and provide their expertise and insights. The U.S. Naval War College and U.S. Strategic Command, for example, conduct the Deterrence and Escalation Game and Review series, a two-sided game to explore escalation dynamics during crises. These types of games also offer an opportunity to explore counterfactuals while allowing game designers to track the discussions.

Alternatively, structured exercises akin to board games offer designers increased control over game dynamics and the potential to increase the number of play-throughs for postexperiment analysis. For example, Karl Mueller at RAND has led an effort to create a tabletop exercise to explicitly address the challenge posed by a resurgent Russia in the Baltic region following the former's invasion of Ukraine in 2016 (6). The goal of this game, carried out with players from the U.S. Air Force and U.S. Army, is to inform the appropriate force composition necessary to defend North Atlantic Treaty Organization (NATO) members and deter adversaries in the region, given adversary capabilities. These insights into strategic decision-making have subsequently driven debate concerning the appropriate qualitative and quantitative force postures in the region.

Even with this type of game using stylized rules, players are subject to laboratory effects. For example, players sitting across from one another may hold back from aggressive maneuvers given the reputational costs associated with taking such an action amid their peers. Inferences related to structured exercises have also been called into question given the small number of players involved and the limited number of turns that may fail to capture real-world dynamics.

The external validity of postgame analysis is also made difficult by the limited number of play-throughs of each game and the challenges associated with collecting data in a manner that allows for replication of the experiment. These difficulties combine to reduce the likelihood of statistical analysis and limit the production of generalizable insights. As a result, board wargame outcomes, like exploratory games, tend to be characterized by insights derived from game play rather than generalizable conclusions based on objective postgame analyses.

These exploratory approaches have also led to a focus on simulated games for training or educational purposes rather than experimentation. For example, the Apex Gold scenario-based discussion program designed by the National Nuclear Security Administration for senior-level policymakers tests how players work together and respond to a hypothetical nuclear terrorism threat. A series of questions and polls also drive the discussion between participants to emphasize the challenges associated with addressing nuclear security.

In sum, computer-based simulations, exploratory games, and structured exercises have led to a theory-rich, but data-poor, environment for scholarly inquiry.

THE SEARCH FOR EXISTING GAME DATA

There are a variety of ways to address this paucity of data by using experimental methods. A number of scholars are attempting to use archival material to reexamine past games for generalizable insights. For example, Reid Pauly uses material from the Massachusetts Institute of Technology (MIT) and the U.S. Department of Defense to collect notes and game outcomes from wargames designed by Lincoln Bloomfield and Thomas Schelling in the 1960s (7). Observing similar behavior across the collection of games, Pauly suggests that policymakers exhibit a generalizable pattern of behavior resulting in nuclear restraint during crises and in spite of provocation. Jacquelyn Schneider's recent project similarly examines longitudinal data created over 7 years of Naval War College wargames involving cyber weapons and finds that, contrary to statements from policy-makers, cyber capabilities do not appear to contribute to crisis instability (8).

Both projects mine collections of outcomes obtained from traditional wargaming approaches to provide quantitative insights not inferable from a single play-through. These approaches, however, present research design challenges, including being time intensive and having no potential for automation in terms of data collection or for adaptation to address alternative research questions that go beyond the original scenario for which the games were designed. There are also concerns about comparing apples to oranges as scenario shift in terms of their framing, the identity and institutional affiliations of the players, and changes to the geopolitical context.

As well as using archival material from traditional wargames, scholars have looked to commercial games to provide natural experiments during the course of gameplay that are analogous to real-world settings—in spite of not explicitly being designed for research purposes (9). A famous example of commercial data providing a simulation of reality for scholars comes from the World of Warcraft, in which the first gamewide epidemic in a massive, multiplayer, online role-playing game led to 4 million player characters being affected by something not unlike a virus. Epidemiologists subsequently used the virus that spread from player to player throughout the game's "world" to model transmission rates and the chain of infection and compared those findings to real-world pandemics (10).

Over the past decade, numerous scientific disciplines have started to explore the use of commercial games for experimental inquiry. For example, virtual worlds have become a laboratory for ethnographic research concerning social behavior on platforms such as Second Life, where players have the potential to create entirely new identities and social relationships. Economists have been examining virtual currencies and financial systems in online games and their implications related to cryptocurrencies, decentralized finance, and blockchain banking. For war planners, battles such as the "Bloodbath of B-R3RB" on the Eve Online platform involving about 7000 human players, 20 million virtual soldiers, and 600 warships provide a virtuous example to study the origins, conduct, and outcome of large-scale warfare in the absence of real-world corollaries (11).

However, these gaming environments are largely outside of researchers' control, which may limit the theories that can be tested. In an attempt to control an experimental environment using commercial software, social scientist have also created stand-alone "mods"—user-generated environments—in games such as World of Warcraft, Star Wars Galaxies, and Starcraft 2 to consider theories of human behavior, cooperation, and conflict from political science, economics, and sociology (12). One of the first of these efforts, NetLab, led to the creation of a variety of "collaboratories" in the early 2000s that sought to use the internet to field social and behavioral experiments (13). "Tribes," for example, modeled on intertribal rivalry in real-world Sudan measure the (artificially created) inter- and intragroup dynamics. Although these mods provide researchers with increased control over the structured play environment, they remain subject to the virtual world, characters, and player pools associated with the original game publisher.

BUILDING EXPERIMENTAL CONDITIONS

Now, scholars are increasingly able to build experimental settings from scratch with the goal of conducting replicable, quantitative analyses that focus on a particular real-

Scholars at play study the impact of weapons capabilities on conflict escalation using the Project on Nuclear Gaming's board game platform, SIGNAL.

[1]Department of Political Science, University of California, Berkeley, Berkeley, CA, USA. [2]Global Security Research, Lawrence Livermore National Laboratory, Livermore, CA, USA. [3]Department of Human Engineering, University of California, Berkeley, Berkeley, CA, USA. [4]Sandia National Laboratories, Albuquerque, NM, USA. *Corresponding author. Email: areddie@berkeley.edu

Wargames as Experiments: A New Tool in the Analysis Toolkit

Substantive Concerns:

- Cyber deterrence
- Hypersonics
- HPLY nuclear weapons

Methodological Concerns:

- Laboratory Effects

Table 6: Logit models predicting the probability of nuclear use and nuclear first use using the player as the unit of analysis. The values in the table body display the regression coefficients with standard error in parentheses.

	Dependent variable:			
	Nuclear First Use		Nuclear Use	
	(9)	(10)	(11)	(12)
Tailored	0.009	−0.021	0.021	0.017
	(0.143)	(0.144)	(0.154)	(0.156)
Female		−0.334**		−0.119
		(0.141)		(0.146)
College Degree		−0.163		−0.018
		(0.168)		(0.181)
Age > 29		0.078		−0.095
		(0.160)		(0.173)
National Security		0.062		−0.005
		(0.214)		(0.227)
More Conservative		−0.110		0.020
Reported Knowledge				
Constant				
Observations				
Log Likelihood	−5			
Note:				

Berkeley
UNIVERSITY OF CALIFORNIA