H-Diplo | ISSF

Article Review 157
issforum.org

Editor: Diane Labrosse and Joshua Rovner
Commissioning Editor: Joshua Rovner
Production Editor: George Fujii

Erik Lin-Greenberg, Reid B.C. Pauly, and Jacquelyn G. Schneider. "Wargaming for International Relations Research." *European Journal of International Relations* (17 December 2021). DOI: https://doi.org/10.1177/13540661211064090.

Published by ISSF on 14 April 2022

https://issforum.org/to/iar157

Review by **Rachel Tecott**, U.S. Naval War College

In "Wargaming for International Relations Research," Erik Lin-Greenberg, Reid Pauly, and Jacquelyn Schneider present wargames as a method for international relations research.[1] The article defines and differentiates wargames from other methods, provides guidance for using wargames for research, and concludes with an agenda for future study. The article is a generative work that provides a firm foundation upon which researchers can build methodologically and substantively. It raises far more questions than it answers, but that is not a limitation of the article. It is the hallmark of a major contribution.

The article defines wargames as "interactive events that display four characteristics: human players, immersed in scenarios, bounded by rules, and motivated by consequence-based outcomes" (3). The definition distinguishes wargames from computer simulations, military exercises, brainstorming sessions, and survey experiments. It is worth noting that wargames, by this definition, need not involve war at all (or even foreign policy crises). The authors suggest that researchers can use wargames to study human behavior related to issues ranging from natural disasters to business and economic cooperation.

The authors suggest that the methodological advantage of wargames lies in their ecological validity— "the extent to which behavior under test conditions mirrors real-world behavior." They propose that "(1) wargames are more immersive for research subjects than other approaches, (2) elite players more closely resemble actual decisionmakers than typical survey samples, (3) interactions between participants better represent real-world decision-making, and (4) wargames present players with the consequences of their own decisions" (5). Because of these unique advantages,

[1] The authors build on recent scholarship examining wargames as methodology, including Ellie Bartels, *Building Better Games for National Security Policy Analysis: Towards a Social Scientific Approach*. Santa Monica, CA: RAND Corporation, 2020; Reid Pauly, "Would U.S. Leaders push the button? Wargames and the sources of nuclear restraint, *International Security* 43:2 (2018): 151-192; and Andrew Reddie, Bethany Goldblum Kiran Lakkaraju, Jason Reinhardt, Michael Nacht, and Laura Epifanovskaya, "Next-generation wargames," *Science* 362:6421 (2018): 1362-1364.

researchers can study the thinking and behavior of players in wargames to shed light on the thinking and behavior of real-world actors.

The article outlines five steps in the process of wargame design and execution: game design (e.g. observational versus experimental games, number of iterations); player selection; rule specification (number of moves, number of teams, nature of adjudication); scenario design; and data collection and analysis. The authors emphasize the tradeoffs scholars face at each step. For instance, scholars may make design choices that increase the immersiveness of the game for the players on the one hand, but that decrease the scholars' ability to isolate how any single factor affects player thinking and behavior on the other. The authors do not purport to offer scholars any clear answers regarding how to weigh such design tradeoffs. Instead, they emphasize how decisions should follow from the question guiding the inquiry, as is the case in all social science research.

The authors highlight the potential value of archival game data for scholarly inquiry, discuss historical game selection, and caution researchers that biases in game design, data collection, and declassification may affect the inferences that we can draw.

The article concludes by outlining an exciting agenda for future research. The authors encourage scholars to use wargames as a method to study the microfoundations underpinning international relations theories of deterrence, crisis signaling, and war initiation.[2] They also encourage scholars to employ wargames in combination with other methods as part of mixed-methods research designs. Beyond advocating wargaming as a method, they encourage scholars to study the intersection of wargames and government bureaucracies in order to generate and test theories of bureaucratic politics and organizational theory.

The authors are humble about their propositions, and they invite scholars to test them. They stress the importance of testing their central argument—that the unique characteristics of wargames make it possible for researchers to learn about how real-world actors might behave. This will indeed be an important area for further study. It is one thing to propose that wargames are more ecologically valid than survey experiments, and another thing altogether to conclude that wargames are ecologically valid *enough* to support claims to external validity. After all, wargame designers can do their best to confront players with the consequences of their decisions, but if the real-life stakes represented in games involve the destruction of cities and the deaths of millions, it is still a leap to conclude that players would think and behave as policymakers under conditions of such extreme stress.

The authors ask why scholars might choose wargames over survey experiments, but readers may also be interested to learn more about why scholars turn to synthetic data production in the first place. Wargames are not a research method of first resort. After all, wargames are extremely resource intensive, and the barriers to external validity are significant. Even beyond the challenge of creating game conditions that represent reality with sufficient fidelity, wargame designers have to worry about whether insights into human behavior surfaced in the context of the specific geopolitical scenario ("road to war") provided to the players at the outset of the game would travel to other scenarios. The authors are also right to emphasize the challenge of representing the thinking and behavior of adversaries with the confidence necessary to draw inferences (12). The significance of these challenges varies depending on the research question guiding the inquiry. Still, as a general rule, scholars would be wise not to turn to wargames to answer questions for which real-world data is readily available. Wargames are a better fit for scholars who are interested in questions for which real-world data is scarce.

Wargames, like surveys, also offer opportunities for controlled experiments. Controlled experiments are considered by social scientists the gold standard for causal inference, as they permit researchers to vary a single factor between a treatment group and a control group, making it possible to isolate the effect of that factor on an outcome of interest.

[2] Joshua Kertzer, "Microfoundations in International Relations," *Conflict Management and Peace Science*, 34:1 (2017): 81-97.

Scholars can iterate "treatment" games and "control" games, varying some aspect of the scenario or information injects in order to study the causal effect of the variation (10).[3]

There is an interesting tension at the core of the article. The authors advocate wargames as an innovative method to generate and test international relations theories in the context of limited real-world data, but they also caution that wargames may not be the optimal method for data production and theory testing. The authors suggest that wargames are best used to explain why outcomes occurred, rather than to predict outcomes (5). However, if researchers use wargames to identify and understand the causes of behaviors, do they not improve their ability to predict the outcomes that follow from those behaviors? The authors caution that "designers should strive for realism, not prediction" (12). But isn't the purpose of realistic test conditions to permit researchers to argue that insights about how certain factors that shape human behavior in the game can shed light on how similar factors might shape human behavior in the real world? Isn't the promise of wargames, as advocated by the authors, to generate *and test* theories of the drivers (cause) of human behaviors (effect) within specified contexts? Isn't theory testing closely tied to prediction? Not deterministic prediction, to be sure, but probabilistic prediction?

Certainly, it is wise to be highly skeptical of researchers who claim that a single wargame can be used to predict with confidence the outcomes of interactions as complex as interstate crises and conflicts. However, the authors provide social scientific design guidance throughout the article that is intended to assist scholars in their efforts to test international relations theories. Wouldn't a well-designed experimental wargame iterated multiple times permit scholars to update their priors about the probable effects of X on Y? In cautioning against prediction, the authors may undersell the method they champion.

I would add several topics to the research agenda the authors advance. First, in addition to testing the authors' propositions regarding ecological validity, future research could also explore how to address other challenges to external validity, including the narrowness of the scenarios that set the context for the game and the challenge of representing adversary ("Red") thinking. Second, researchers could study how to incorporate virtual reality into wargames to further enhance ecological validity. Virtual reality has only recently entered the world of political science, but seems a promising avenue for further bolstering ecological validity in wargames.[4]

Third, research could focus on how to optimize adjudication for social science. Adjudication is the process of determining outcomes of game events. Game designers and research assistance or staff often play the role of determining how player moves interact with each other. For instance, if two "sides" in a game initiate military operations against each other, it is up to adjudicators to determine the results of that interaction of military forces. The authors touch only briefly on adjudication in the article, yet adjudicators can play an enormous role in shaping the direction of the game and the kinds of inferences that can reasonably be drawn from the game. Adjudicators often operate under time constraints, in a black box, with their own teaching and research goals in mind. For these reasons, they sometimes push the game in the direction they wish it to go without systematically weighing design tradeoffs or acknowledging their thumb on the scale.

Fourth and relatedly, future research could examine how wargames optimized for social scientific research should be designed differently from wargames optimized for experiential teaching. Wargame designers who want players to come out of the game with, for instance, a healthier concern about the risk that conventional fighting could escalate to nuclear use, might make very different design choices than scholars seeking to examine how cyber attacks could affect nuclear use

[3] Rose McDermott, "Experimental Methodology in Political Science," *Political Analysis* 10:4 (2002): 325-342.

[4] For examples, see Andrew Cesare Miller, "The Information Game: Police-Citizen Cooperation in Communities with Criminal Groups," Ph.D. diss, MIT, 2020, and Julian Borger, "'15 Minutes to Save the World': A Terrifying Journey into the Nuclear Bunker," *The Guardian*, December 14, 2021

risk. Moreover, scholars who examine archival game data need to pay attention to how designers balanced teaching and research goals in game design and execution.

A fifth avenue for future research would integrate wargames and campaign analysis.[5] For wargames that involve the interaction of military forces, adjudication often involves more or less rigorous forms of campaign analysis. Campaign analysis is a method involving the use of a model and techniques for managing uncertainty to answer questions about military operations. Instead of developing a model to analyze the interaction of military forces between game moves, adjudicators operating under time constraints sometimes make rough guesses. If adjudicators have a teaching objective, they may decide to weight the interaction of forces to one side or the other in order to encourage a specific line of thought for the student players.

For their part, campaign analysis modelers often bake 'decision rules' into their models. Modelers attempt to handle uncertainty in a variety of ways. If they are uncertain of the precision of a particular missile, they may set upper and lower bounds on missile precision. For human decision-making, modelers may specify the range of decisions each side could plausibly make, and then assign probabilities to different decisions. Thousands of runs of highly complex models spit out outputs that reflect uncertainty in human decision-making alongside all the other sources of uncertainty complicating analysis of hypothetical military operations.

The modeling communities and the wargaming communities within government tend to regard each other with skepticism, if not downright contempt. Scholarship on campaign analysis and wargaming need not replicate these intellectual siloes. Scholars should nip this impulse in the bud and cross-pollinate the methodologies. Wargame adjudication could be informed by campaign analysis, and campaign analysis decision-rules could be informed by wargames.

Rachel Tecott is an assistant professor at the U.S. Naval War College. Her research on alliances and partnerships, U.S. military strategy and operations, and bureaucratic politics has appeared in *International Security*, *Foreign Affairs*, *The Washington Post*, *War on the Rocks*, and the *Irregular Warfare Initiative*. She earned her Ph.D. in political science from MIT, where she was a member of the Security Studies Program. The views expressed here are her own and do not reflect the official positions of the Naval War College, the Department of the Navy, or the Department of Defense.

[5] Rachel Tecott and Andrew Halterman, "The Case for Campaign Analysis: A Method for Studying Military Operations," *International Security* 45:4 (Spring 2021): 44-83.