“Archetype of Information Produced by Analytical Games”

Elizabeth M. Bartels, Pardee RAND Graduate School

Abstract:

This paper argues that for policy gaming to be more impactful in research communities, it must be better able to expose the logic of design and analysis to outside scrutiny. Because policy games differ in key ways from mainstream research techniques in the social sciences, we must develop a gaming-specific set of logics to do this work. This paper presents a set of archetypical types of information that can be generated from a game, based on an iterated expert validation approach. It then delves into the logic of each—detailing what differentiates each type from the others and discussing typical tradeoffs make in the design of games of each type.

.

What and how do we learn from games? Surprisingly, this is not a settled question among policy gaming practitioners. Over the course of several years, gamers have debated whether games generate knowledge through deductive, inductive, or abductive processes, reaching little agreement.[1] A similar debate has also raged over whether game can best be considered an art or a science, and what that might mean.[2] Advocates for seeing games as "art" emphasize the experience of the participants of a game, and how games have the ability to cultivate new thoughts and build new understanding in the minds of players.[3] Defenders of the position that games are a science argue that, if games are going to contribute to analytic projects, they should be held to the same standards as other types of research.[4] While both positions are represented among professional gamers, based on analysis of a recent survey of gamers, those using artistic language and design principles like "player enjoyment" outnumber the supporters of viewing games from a scientific perspective.[5]

The commitment to treating policy gaming as an art comes with costs. United States Department of Defense officials have stated that greater systemization of gaming is necessary for games to have their desired impact on policy.[6] Spend time at a gathering of gamers, and it is normal to hear concerns that the results of games are warped or dismissed by sponsors and stakeholders,[7] and that inexperienced gamers are developing unsound products.[8] At the same time, newer gamers argue there are not sufficient tools to learn their trade.[9] Without the discipline of science, too often games fail to meaningful inform what we know about the world in credible ways.

So why is there so much resistance to viewing games as a scientific approach? Ivanka Barzashka's recent article in the *Bulletin of the Atomic Scientist* calling for a more scientific approach to gaming makes several important points that align with my own observations of the

---

[1] Jonathan Compton, "Towards an Epistimology of Wargaming--a Drunkard's Walk" (paper presented at the Military Operations Research Society Wargaming Community of Practice, Alexandria, VA, 2015).

[2] Peter Perla, "The Art and Science of Wargaming to Innovate and Educated in an Era of Strategic Competition," in *King's College London Wargaming Network Lecture* (London, UK2018).

[3] Peter P. Perla and ED McGrady, "Why Wargaming Works," *Naval War College Review* 64, no. 3 (2011).

[4] Ivanka Barzashka, "Wargaming: How to Turn Vogue into Science," *Bulletin of the Atomic Scientists*2019.

[5] Elizabeth M Bartels, "Insights from a Survey of the Wargaming Community," in *Military Operations Research Society Wargaming Community of Practice* (Alexandria, VA2017).

[6] Robert Work and Paul Selva, "Revitalizing Wargaming Is Necessary to Be Prepared for Future Wars," *War on the Rocks*, December 8 2015.

[7] Stephen Downes-Martin, "Your Boss, Players and Sponsor: The Three Witches of War Gaming," *Naval War College Review* 67, no. 1 (2014).

[8] Stacie L. Pettyjohn and David A. Shlapak, "Gaming the System: Obstacles to Reinvigorating Defense Wargaming," *War on the Rocks*, February 18, 2016 2016.

[9] Elizabeth M Bartels, "Building a Pipeline of Wargaming Talent: A Two-Track Solution," ibid.2018.

field.[10] The most important of these is that many gamers have a very narrow understanding of how science is practiced, strongly influenced by military operations research[11] and associated approaches. These approaches tend to define science in quantitative terms, featuring "definite methodologies of attacking new problems and finding definite solutions."[12] Unsurprisingly, this approach to science is a poor fit with the human-centric, exploratory nature of games.

Like Barzashka, I find that the more expansive view of science put forward by scholars like Patrick Thaddeus Jackson to offer a far more useful description for the type of science I think gamers can (and should) practice.[13] Engaging with similar debates over the merit of science in international relations, Jackson argues for a broad definition of science that recognizes multiple possible objectives, including both descriptive and inferential tasks, and multiple viable rules of inference. Instead, a scientific approach is one that takes a systematic, transparent approach. Scientific work should enable an observer to determine whether: "given our assumptions, our conclusions follow rigorously for the evidence and logical argumentation we provide"[14] As Barzashka frame it, this requires work to be "systematic, public, and intended to produce worldly knowledge."[15]

The question then turns to how these values can be instantiated in the work of policy gaming. By the practical nature of the field, which generally depends on policy stakeholders to sponsor work on questions of interest to them, policy games tend to be tied to real world problems. The public conduct of games is somewhat more fraught, since the sponsorship of work often leads to sensitivities that discourage publication. However, policy gaming has made strides to increase the documentation and accessibility of its work. For example, the United States Department of Defense has begun to track games in a central database to make access across offices easier.[16] Major wargaming conferences like the Connection Wargaming Conferences post presentations and reports to their respective websites for public access.[17] Efforts to publish book length studies

---

[10] Barzashka, "Wargaming: How to Turn Vogue into Science."

[11] It is perhaps worth noting that there are alternatives to this vision of operations research, particularly in older texts, that is still practice, particularly in the UK. For a discussion of alternative approaches to operations research, and their relationship to gaming, please see: Yuna Wong, "Preparting for Contemporary Analytic Challenges," *Phalanx* 47, no. 4 (2014).

[12] Philip M. Morse and George E Kimball, *Methods of Operations Research*, 3rd ed. (Mineola, NY: Dover Publications, Inc, 2003).

[13] Patrick Thaddeus Jackson, "The Conduct of Inquiry in International Relations: Philosophy of Science and Its Implications for the Study of World Politics," (New York, NY: Routledge, 2011).

[14] Ibid. p 22

[15] Barzashka, "Wargaming: How to Turn Vogue into Science."

[16] Oleg Svet and Garrett Heath, "How the Joint Staff Calculated a Defense Program's Return on Investment," *Defense One* 2018.

[17] For the US conference website, see: https://connections-wargaming.com/ for the UK conference see: http://www.professionalwargaming.co.uk/

of wargaming by established practitioners are also on the rise—after years of pointing to a single major treatment of the past 30 years,[18] several new books are expected in the coming years. Addition work from academic scholars with greater incentives to publish is also shifting this dynamic.[19] In short, the field has begun to operate in a more public manner, and already is incentivized to research real-world issues.

It is in the area of developing "systematic" approaches where less progress seems to be made. Existing texts on game design stress the importance of linking design to purpose but offer very little advice on how to achieve this goal. The most often-cited handbooks on the design of games identify a range of different options for each of these elements and discuss the advantages and draw-back of each.[20] These works also stress the importance of linking the choice of design elements to the purpose of the game. However, when it comes to *how* to make the linkage, these texts are largely silent. Even the best-respected book on game design states: "There is no recipe for translating a game's objectives into its mechanics… ultimately the designer's talent dictates how and how well the translations from objectives to mechanics works"[21] In other words the core task of the designer is surprisingly under-theorized. In part, this is the result of the dominant view of games as art, the proponents of which tend to worry that systematized logics for design at best encourage "cookie cutter" design in which inexperienced gamers replicate a few, not very strong designs,[22] and at worst serve as a straight-jacket to artificially eliminate promising game designs from consideration.

In my mind, much of this concern stems from the overly-narrow definition of science that has often been invoked in discussions of games. Applying Jackson's focus on making the both the observation and logical argumentation to connect the research process to the conclusions generated puts the focus on making the logic of games more readily understandable to those outside the research team. It argues for developing clearer ways to communicate how design choices connect to the information learned from the game—so we can speak of more- or less-

---

[18] Peter P Perla, *The Art of War Gaming : A Guide for Professionals and Hobbyists*, 2nd ed. (History of Wargaming Project, 2011).

[19] For several recently published examples, see: Reid B.C. Pauly, "Would U.S. Leaders Push the Button? Wargames and the Sources of Nuclear Restraint," *International Security* 43, no. 2 (2018); Jacquelyn G. Schneider, 2018; Erik Lin-Greenberg, "Game of Drones: What Experimental Wargames Reveal About Drones and Escalation," *War on the Rocks*2019.

[20] For well-known examples of this approach, see: Shawn Burns, ed. *War Gamer's Handbook: A Guide for Professional War Gamers* (Newport, RI: War Gaming Department, U.S. Naval War College); James F. Dunnigan, *Wargames Handbook: How to Play and Design Commercial and Professional Wargames*, 3rd ed. (IUniverse, 2000); Matthew B. Caffrey, *On Wargaming: How Wargames Have Shaped History and How They May Shape the Future* (Newport, RI: Naval War College Press, Forthcoming).; and Perla, *The Art of War Gaming : A Guide for Professionals and Hobbyists*.

[21] *The Art of War Gaming : A Guide for Professionals and Hobbyists*.

[22] "Now Hear This—Improving Wargaming Is Worthwhile—and Smart," *Proceedings Magazine*, January 2016 2016.

logical design and analysis as being better or worse, rather than arguing about the relating "scientific-ness" of the approach.

One obvious place to look for such logics in in established social science approaches to research, but unfortunately there is not a single clear analogy available. When compared to the traditional tools of positivist policy analysis and social science, games can often seem neither fish nor fowl. Because games occur in an artificially simulated space, they do not produce traditional empirical data based on observation of real-world policy events. At the same time, the involvement of actual humans making choices means that games are to some extant an empirical exercise, in contrast to formal models or modeling and simulation efforts that are purely artificial. In this sense, games' closes analogy may be social science lab experiments, since they too feature observations of humans making choices in a synthetic environment. However, even this analogy comes up somewhat short, since games also feature groups of people making decisions in competition with one another. As one game designer eloquently framed it, games replicate more of the actual decision-making interactions, and thus are more generalizable. The cost of this is that the researcher loses the level of control typically associated with experiments.[23] In short, while games have similarities to other research approaches, there are also important differences that must be accounted for.

Another potential source for logics for games is in the limited, but existing, literature on game design. While several existing frameworks focus on the purpose of the game, none discuss the connection between game purpose and design choices in a deep enough way to really generate a true logic. Perhaps the most influential of these on my thinking is Ed Parson's distinction between games for experimentation, to instruct decision makers, to promote creativity and insights, and for the integration of knowledge.[24] While this approach to categorizing games by purpose stuck me as sensible, the essay is focused on explaining which types of games should be run and which are likely to fail. Missing is a sense of what is at stake in game design—in other words, how do you design a game given that you want to achieve these ends? Two other frameworks that do somewhat similar work are Graham Longley Brown's construction that games can support efforts to understand problems, provide insights, and evaluate policies[25] and Stephan Downes-Martin's distinction between experiential, comparison, and analytic games.[26] While I agree with these categories to a greater degree than the Parson typology, they are even more underdefined. While the seeds of different potential logics of games can be found here,

---

[23] Interview with Jacqueline Schneider, Newport, RI, June 2018.

[24] Edward Parson, "What Can You Learn from a Game?," in *Wise Choices: Decisions, Games, and Negotiations*, ed. Ralpj L. Keeney Richard J. Zeckhauser, James K. Sebenius (Boston: Harvard Business School Press, 1996).

[25] Graham Longley Brown, *Successful Professional Wargames: A Practitioner's Guide*, ed. John Curry (The History of Wargaming Project, Forthcoming). pp 89-91

[26] Ibid. p 91.

there is not enough detail to help an observe trace the logic of others work, and thus not enough to satisfy Jackson's requirement of a scientific approach.

For gaming to find its footing as a scientific approach, I believer it needs a framework that lays out different potential logics of how games can produce information. The remainder of this paper presents the approach and initial framework develop to meet this goal. The next section lays out the approach to framework development, given limits on available data. The remainder of the paper presents a framework composed of four archetypes. In addition to presenting each archetype, this section discusses the features that differentiate the different types. I then consider the different design tradeoffs implied by each type. These are not intended to offer concrete rules or scripted design, but rather to help designers consider the consequences of different tradeoffs in how the game's environment, actors, and rules are manifested for the quality of information that the game can generate. Finally, the paper closes with a discussion of some concluding thoughts.

Methodology for Framework Development

Building on the existing literature, which defines different types of games, this study aims to develop a framework of different types of games that can help define different logics games use to move from observations of the game to useful information. The first task was to determine the type of framework—or classification scheme--that would provide the most utility to game designers. After determining that a set of archetypes would be most productive, I then surveyed the available sources of data to support the population of types. Because of limited records of games, it was most feasible to focus on expert validation, rather than a traditional classification activity, to refine the framework. As a result, I opted to conduct a series of expert interviews in individual and group settings to gather feedback about the ability of the framework to capture expert practice.

One important note is that in this study, I only consider games whose primary purpose is research, or producing new information for the sponsor. Games have many educational and commercial applications,[27] which are designed to inform, persuade, and entertain. This is not to suggest that education and commercial games are not important—wargaming are integrated into many national security education curricula,[28] and represent a major portion of all defense gaming. There is also important crossover between education, commercial, and research games. In some cases, games originally designed for research purposes can later be used in educational settings to teach the new knowledge generated. Often, commercial games develop innovative mechanics that can then be co-opted by research and education gamers to enrich their games.

---

[27] For a discussion of game typologies that includes deeper discussion of educational and commercial games in relation to games for research, please see: Elizabeth M Bartels, "Gaming – Learning at Play," *ORMS Today* 41, no. 4 (2014); Phillip E Pournelle, "Designing Wargames for the Analytic Purpose," *Phalanx* 50, no. 2 (2017).

[28] Chairman of the Joint Chiefs of Staff, "Officer Professional Military Education Policy," (Washington, DC2015).

However, fundamentally the goal of communicating with a game is different from that of learning from a game, and so I have opted to keep my scope narrow to avoid confusion.[29]

Selection of classification scheme

Generally, categorization systems are developed in one of two ways, either top-down by defining theoretical distinctions that can then be tested by sorting the population of interest, or bottom-up. While terminology differs somewhat between fields, the distinction is usually made between typologies, which are driven by theory, and taxonomies that start with an empirical base to define categories. Both approaches argue that classification schemes should allow all instances of the phenomenon to be sorted in one of the established categories.[30] However, both traditionally demand that to be useful, items should be able to be classified into one, and only one category. This requirement is unhelpful when it comes to games, because it is well established that games often have multiple objectives,[31] and thus a single game may be operating under more than one logic at a time. Simply put, a rigid classification system is unlikely to be useful given the diversity and flexibility of games.

A variant of typologies, archetypes, seems more promising. Used in fields ranging from philosophy, psychology, and literary criticism, archetypes feature "ideal forms." In the field of policy analysis, they are perhaps most closely associated with systems thinking, which defines "system archetypes"[32] which describe broad patterns of behavior that reoccur in many different contexts. Beyond these applications, archetypes have a long history of use in policy analysis as a tool for communicating complicated results to broad audiences, suggesting that the approach may help make my findings more accessible to non-expert audiences, particularly game sponsors. Because archetypes are ideals, the expectation is that few if any observed examples will be fully described by the archetype. Rather, it's a tool for identifying patterns, which may occur in combination, and more or less strongly, across cases. As a result, this framework recognizes that games may fall into multiple archetypes, and that over time, new styles of games could be introduced requiring additional models to be added.

The importance of pattern detection in archetypes seemed a particularly good fit given the initial survey I conducted of game designers. When surveyed, expert gamers describe a well-

---

[29] In an earlier iteration of the framework I did include educational games, which created an interesting effect in which any game that had both educational and research objective shared the same characteristics as a game that only had the research objective. However the presence of the educational objective provided an "out" to designers to declare a game successful even if it did not produce the desired information, As a result, I opted to remove it from the framework to more clearly focus attention on designing games to produce desired information.

[30] Kevin B. Smith, "Typologies, Taxonomies, and the Benefits of Policy Classification," *Policy Studies Journal* 30, no. 3 (2002).

[31] Downes-Martin, "Your Boss, Players and Sponsor: The Three Witches of War Gaming."

[32] Peter M. Senge, *The Fifth Discipline: The Art and Practice of the Learning Organization*, Revised and Updated edition ed. (New York, NY: Doubleday, 2006).

designed game as one for which the game design "matches" the purpose of the game, but they fail to explicitly describe what this match might consist of.[33] The limited language that experts do use to describe the design process has interesting parallels in the literature on expert decisionmaking, and in particular the Recognition-Primed Decision Making (RPD) model.[34] This model centers the role of pattern recognition in decisionmaking—that is, rather than comparing competing options or choices, experts draw on a bank of experiences to quickly assess what is typical about a decision and its context and then use this determination, as well as anomalies in the decision-making context, to develop a viable plan of action. In other words, the RPD model would posit that when experienced gamers are presented with a new game's purpose, they quickly identify what games the new project is similar to, based on their experience and how the new project may diverge from "typical" projects, and use that pattern recognition as a basis for making design decisions.

Archetypes play into exactly that pattern by provide examples for the designer to compare to that may be outside of their own experience because they represent a range of extreme types. A designer can use archetypes to compare to the current situation and determine on what dimensions the problem is similar to any given type and use that information to refine a design. It also gives designers a common language the talk about their games—designers may not have observed the same set of games, but the archetypes provide a shared pool of references that designers can point to aide communication.

Data and approach

There are major lacunas in our record of games. Public reports on games represent only a small portion of the total games run. Security classification, sensitivity, and professional cultures that do not prioritize open publication limit the available game documents in ways that cannot be reasonably assumed to be random.[35] Even more problematically, available written reports often include very little information about the research design of the game, often focusing almost exclusively on a narrative of game play and key insights from the game. In part, this may be due to the limited common language gamers have access to that can describe research design and logic of inference type considerations. Other plausible explanations for limited documentation include lack of sponsor interest in technical details, efforts to protect intellectual property by for-profit game designers, and the current lack of technical language for describing research design choices. When technical details are provided, they often focus on specific components of the game, such as the adjudication system, rather than the overarching logic of research and how it contributed to design. In short, exactly the type of data needed to develop or test a traditional

---

[33] Bartels, "Insights from a Survey of the Wargaming Community."

[34] Gary Klein, *Sources of Power: How People Make Decisions* (Cambridge, MA The MIT Press, 1998). pp 147-153.

[35] For an example of thoughtful treatment of bias in a sample of games, see: Pauly, "Would U.S. Leaders Push the Button? Wargames and the Sources of Nuclear Restraint."

classification scheme focused on research design in games is often missing from our written records.

Instead of attempting to draw conclusions from categorizing limited empirical data, I developed an alternative approach which depended on expert validation to inform iterative refinement of the framework. I started by developing a framework, based on my own understanding of game design. I then conducted a range of semi-structured interviews with game designers and sponsors in which I walked participants through the framework and elicited feedback. Generally, feedback took one of three forms: 1) general statements that the framework aligned with their understanding of design, 2) specific concerns with one or more categories, and how they interacted with their own experience, or 3) concerns about how to differentiate two or more categories.[36] Based on feedback of the later two types, I then made revisions to the framework, prior to conducting the next interview, restarting the cycle.

I conducted interviews with a range of game designers and sponsors. Over the course of the study I interviewed over 30 individuals in one-on-one and small group setting. Individuals were recruited from my professional network, recommendations from other subjects, and calls for interested participants made at major gaming conferences. Table 1 summarized the institution and equities represented in these conversations.

Table 1: Affiliation of interview and selected workshop subjects

Equity	Organization	Designers	Sponsor
Joint	Joint Staff Studies, Analysis and Gaming Division	X	
	Office of the Under-Secretary of Defense, Policy	X*†	X*
	Office of Cost Assessment and Program Evaluation, Department of Defense	X	X*
	Office of Net Assessment, Department of Defense	X*†	
	RAND	X†	
Army	Army Command and Staff College	X	
	Center for Army Analysis	X	
Navy	Naval War College (NWC) Wargaming Department	X	
	NWC Strategic and Operational Research Department	X	
	NWC Halsey Group	X	
	CNA	X*†	
Air Force	Air Force Research Laboratory	X	
	RAND	X†	
Marines	Group W	X†	
Interagency	Central Intelligence Agency	X	X
	Defense Intelligence Agency	X	
	Department of Homeland Security	X†	
International	UK Ministry of Defense	X	

* indicates interview was with a former (<5 years out) member of the office
† indicates interview was with a contractor or federally funded research and development center researchers

[36] For the initial decision to focus on distinguishing characteristics, I am indebted to discussions with Stacie Pettyjohn (interview, Arlington, VA, May 2018)

In addition to individual interviews, I also conducted a broader "validation workshop" at the Connection Wargaming Conference in July 2018. This session was held during a workshop track of the conference, where participants could self-select to attend. Nearly 40 individuals opted to participate in the session. While detailed demographic information was not collected during this session, based on visual inspection and participants previously known to me, the group included participants from the US Army, Air Force, and Navy, Intelligence Community, contractors working across the joint and service communities, and UK MoD. The workshop included both a presentation of the framework featuring a somewhat abridged version of the type of feedback offered in the individual interviews.

My aim in selecting this approach was to establish the validity of the framework in the eyes of a range of experts. However, that makes this approach sensitive to my selection of experts. For some of the same reasons as it is difficult to account for what games are run to support national security policy, it is also difficult to characterize the community of gaming practitioners. Because there is no common academic training, certification for wargaming required for practice, or professional credential (such as a bar or board) there is not central repository of professional wargamers. Many different organization practice wargaming, and not all are known to one another. This problem is compounded by the dominance of contractors, which increase the number of organizations involved exponentially, but may not always be publicly visible. As one prominent convener put it, wargamers seeking to engage with the broader field must, in the words of game designer Tim Wilkie, "navigate an archipelago of excellence" in which many different centers work in relative isolation. A range of profession conferences and communities of practice have been established to mitigate these disconnects. However, because of the time and expense associated with conferences, they attract only a portion of the field. All of these efforts bias toward U.S. DoD games, with uneven participation from other sectors. Other biases may well be present—for example, contractors must justify charging overhead, making them less likely to participate unless they can demonstrate concrete benefit. Much like was the case with publications, the scope and extent of these biases is difficult to predict—because we do not know the full shape of the community, it is hard to know who is missing. Thus, my sample of experts may have introduced biases into my framework, by failing to account for the full diversity of practice.

The Framework

This section proceeds in four parts. First, I discuss the framing of the framework. I then present the four types of the framework. The next section highlights characteristics that differentiate the archetypes to help clarify the types. Finally, the section ends with a discussion of design trade-offs for each archetype.

Framing of the Framework

Early in the framework development process, some work was needed to establish how best to frame the archetypes. Initial conversations quickly revealed that the language of logic of inquiry and research design did not translate well across gamers from different research backgrounds. Policy gamers are drawn from a wide variety of fields,[37] and often terminology, particularly ones related to research design like validity, rigor, and generalizability do not have the same meaning and use across different practices.[38] As a result, attempts to import formal terminology were met with confusion.

Turning to look at the gaming literation, past work on what can be learned from a game generally focuses on the purpose of the game.[39] While this makes sense in theory, in practice it is complicated because the purpose and objectives of the game are set in consultation with the game sponsor, leading to the use of unclear language and multiple objectives.[40] Designers often do considerable work to translate from the "office" objects recorded in the game documentation to their own understanding of what is desired from the game, which then drives design.

Instead, on advice from Ed McGrady, I have opted to focus not on the purpose and objectives of the game itself, but rather on the desired end point of the project—what information needs to be produced in the game.[41] While in some senses the difference is semantic, since the purpose and objectives of the game should state what the desired outcomes of the game are, this backward logic of starting with the desired end point resonated with other designers. For example, this approach mirrors current teaching by other gamers, in which data capture plans are developed based on what information should be produced in the game, and the game is designed around these requirements.[42] As a result, I opted to frame the archetypes around the types of information the game is designed to produce, rather than focusing on purpose and objective.

Overview of the Types of the Framework

The framework consists of four archetypes, or ideal types, that describe the type of information that a policy research game can generate. These are system exploration, alterative conditions, innovation, and evaluation. Each type is described in greater detail below.

System Exploration: This archetype highlights games that bring together diverse stakeholders to contribute their understanding of the policy system to generate a thick description. The

---

[37] Bartels, "Insights from a Survey of the Wargaming Community."

[38] I am indebted to my colleagues at the Naval War College Wargaming department, particularly Peter Pellegrino and Hank Brightman, for clarifying these points over the course of several discussions.

[39] For example see: Parson, "What Can You Learn from a Game?."

[40] Downes-Martin, "Your Boss, Players and Sponsor: The Three Witches of War Gaming."

[41] Interview with Ed McGrady, game designer, Arlington, VA, July 2018.

[42] Interview with Jeff Applegate, game designer, Washington, DC, August 2018.

primary goal of such games is to elicit and synthesize designers' and players' mental model of a policy problem, and how it can evolve over time. The output of successful system exploration games is a representation or model of the problem that combines insight from players with the research of the design team to improve the sponsor's understanding of the nature of the problem.

Alternative Conditions: These games aim to detect patterns of decision-making based on similarities and differences in the decision-making environment to help advance causal inference. The game designs of this type try to minimize variation in environment, actors, rules, and model across iterations, while thoughtfully changing the key factors. Successful games of this type produce an understanding of the impact of varying conditions on decision-making.

Innovation: Innovation games seek to develop new decision options that break from the status quo as a form of policy ideation. These games aim to build a model of the world that relaxes constrains, in the hopes that doing so might enable new approaches to problem-solving. In this way, it shares similarities with hypothesis generation and brainstorming activities. The ideal outcome is for this type of game is to generate one or more promising ideas for further consideration.

Evaluation: The evaluation archetype describes games that aim to judge the potential outcomes of player decisions based on a normative standard—in other words, to evaluate policies, courses of action, or interventions. These games focus great attention on adjudication, in order to generate credible outcomes from player decisions. Because the game must project plausible outcomes, it must contain a fairly well-developed theory of causality that allows the game staff to project different counterfactual outcomes based on player actions. The desired outcome of these games is an assessment of the potential gains and losses from following a course of action.

Differentiating the Types

These archetypes differ from one another in a number of ways, however, the four types can most clearly be differentiated by two characteristics: whether the primary focus of investigation is on understanding the problem or a solution, and whether the audience for the information to come from the game is internal to the sponsoring organization or external. As with the overall archetypes, these characteristics are not always mutually exclusive, but represent common "modes."

First, the distinction between focusing on the policy problem and examining potential solutions shapes the focus of game findings. Games that focus on the problem seek to better understand the policy system in question, including how different stakeholders think about the issue and how the system can evolve over time.[43] In contrast, games that focus on solutions

[43] I am indebted to Chris Chivvis (interview, McLean, VA, March 2018) and Margaret McCown (personal correspondence, March 2018) for helping clarify the ways in which my understanding of policy problems in the contest of games depends on a systems approach to understanding problem. For a treatment of systems analysis

develop and judge potential interventions into the system to understand how they might interact with the broader system. In policy contexts, this distinction also implies a difference in game purpose—games that focus on the problem are more likely to be descriptive in nature, while games focusing on solutions are more likely to lean to proscriptive recommendations.

The audience of the game determines what types of decisions the information from the game should inform, and what type of credibility it needs to be impactful. On the other axis, games for internal audiences seek to inform the sponsoring organization—either by shaping the knowledge base of senior leaders within the organization or by shaping future studies or investments under the direct control of the office. Alternatively, the analytic output of games can be designed to produce information that can influence outside of the sponsoring organization by communicating a particular understanding of a policy problem or generating evidence to that can be used to support recommendations. Games that seek to influence outsiders need to do more to provide tracible results, because the game findings must stand on their own to a greater degree than game findings that are influencing individuals directly involved in the design, execution, and analysis of the game.

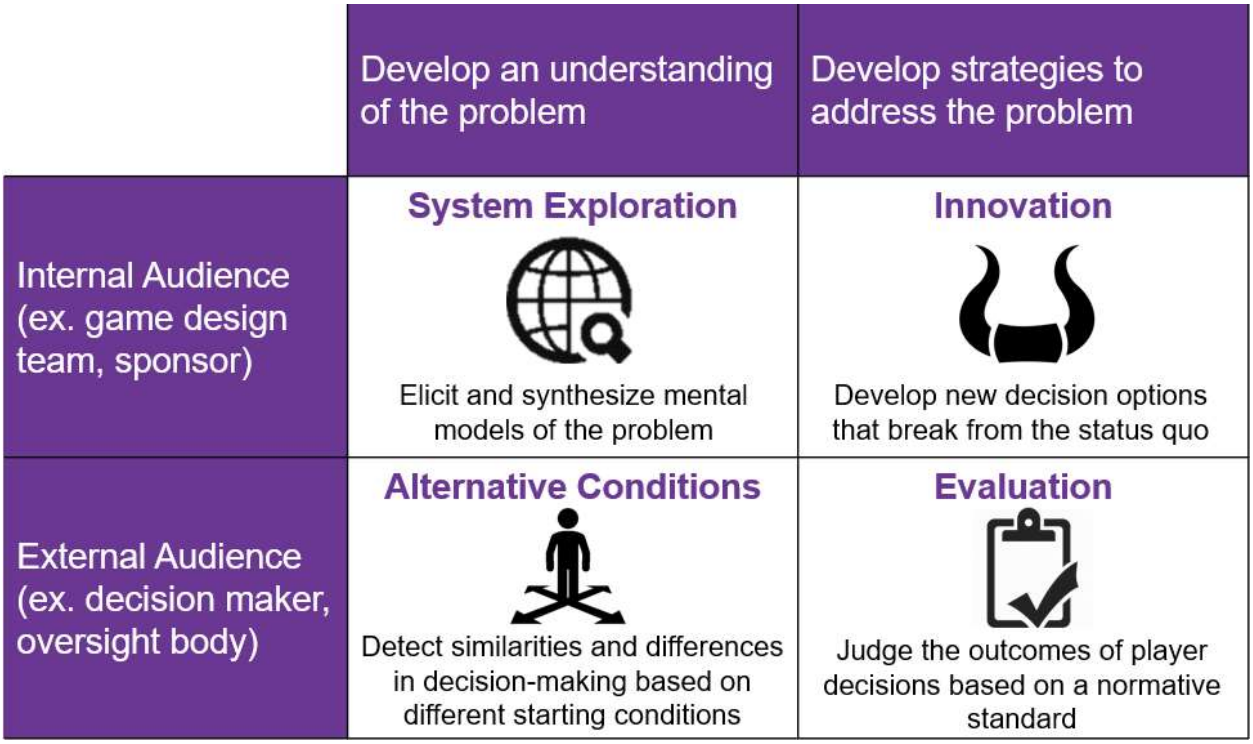
By combining each set of characteristics, we can define the four archetypes. System exploration and alternative conditions games focus attention on developing an understanding of the problem, while innovation and evaluation games focus on potential solutions. System exploration and innovation games are generally intended for internal audiences, while alternative conditions and evaluation games are usually designed to inform external audiences. Figure 1 illustrates how the four archetypes align with these characteristics.

It is important to note that these distinctions come with some important corollaries. Because of the artificial nature of game environments and limited degree of experimental control over players and their interactions, many gamers are deeply uncomfortable using games to support causal or predictive analysis. However, games focused on solutions, and games providing information to external audiences (and most especially evaluation games, that do both) begin to wade into these dangerous territories. For example, because the designer has a great deal of control over how the games environment and rules are shaped, an unscrupulous designer could set up a game designed to produce information favorable to their position in order to persuade outside organizations of the validity of their arguments. Less maliciously, a designer unaware of the potential biases introduced by a specific group of players could over-generalize results of one game to a much broader set of real-world decision contexts, offering poor predictions. As a result, many designers caution that games of these types are difficult to execute well and require careful modesty in the analytical claims to be credible. Thus, most designers would assert that the different types of games vary in difficulty from easier in the upper left (systems exploration) to harder in the lower left (evaluation).

---

applied to policy problems, see: Bob Williams and Richard Hummelbrunner, *Systems Concepts in Action: A Practitioner's Toolkit* (Stanford, CA: Stanford University Press, 2011).

Figure 1: Archetypes organized by defining characteristic



Beyond direct consequences of these two distinguishing characteristics of games, there are several other related characteristics we can use to describe how games of these types differ from one another, summarized in table 2. The first point is that the different types of game present different core design and analysis challenges—that is, because they aim to generate different types of information, there are different tradeoffs that need to be considered. Second, the maturity of the research—that is how developed our understanding of the issue is—tends to be different across game types. Similarly, where in the process of the game is the focus of study will differ. Finally, the target audience of the knowledge generated by the game is different. By unpacking these differences, we can better differentiate between the types.

First, we can describe the core design and analysis challenges that highlight what make each type of game particularly challenging. In the case of systems exploration games, the key concern is how to build a game design that elicits expert understanding in a way that allows it to be captured, and then to transmit that understanding to others in analysis. Alternative conditions games are challenging because of the need to control as many possible sources of confounding variation as possible during design, and to account for the potential effects of variation that could not be controlled for in analysis. The design of innovation games focuses on the challenge of relaxing enough constraints that there is space for new ideas to be thought of. Analysis of

innovation games must struggle to provide a helpful screening of which ideas should be pursued farther: dismiss ideas too quickly and good options could be discarded; present too many harebrained ideas or mild tweaks to the status quo and the sponsor will lose faith in the value of the game. Finally, evaluation games struggle to develop both a credible adjudication system to ensure that outcomes meaningfully reflect potential real-world outcomes, and measure game results in a clear and accurate way. As will be discussed later in this chapter, these differences in design and analysis imply that game designers will often make different trade-offs when setting up the elements of these four types of game.

Table 2: Distinguishing characteristics of each archetype

	System Exploration	Alternative Conditions	Innovation	Evaluation
Core design challenge	Elicit expert thinking in an accurate and transparent way	Maintain control over factors that are not being intentionally varied	Change status quo conditions enough to enable new thinking while retaining realism	Develop adjudication process to generate credible outcomes
Core analysis challenge	Capture expert thinking, accounting for analyst bias	Account for the potential influence of factors on similarities and differences	Initial assessment of idea quality	Develop appropriate measures and instruments
Maturity of research	Early exploratory	Mid-term exploratory	Mid-term exploratory	Late refinement and confirmatory
Focus of findings	Framing the problem system	Decisions	Decisions	Potential Outcomes
Target audience for knowledge	Stakeholders trying to understand the problem set	Decisionmaker responding to or setting conditions	Investor in future research	Policymaker selecting course of action

We can also consider the maturity of research associated with a particular type of game. Games are often combined into broader studies include either multiple games, or games coupled with other techniques.[44] While there is not a hard and fast sequence of where games of different types appear in the cycle of research, generally systems analysis games are run when first trying to understand the nature of a problem, while evaluation games are run later once there is a good understanding of the problem and of potential solutions under consideration. Alternative condition and innovation games fall somewhere in between—they require a somewhat structured understanding of the problem in order to identify factors to manipulate but are only useful when there are still substantial gaps in our understanding of decision-making. As a general rule, if we look at Figure 1 we expect games in the upper left to occur earlier in a research project, then those in the lower right.

[44] For a longer discussion of games in mix-method studies, see: Elizabeth M. Bartels, "Adding Shots on Target: Wargaming Beyond the Game," *War on the Rocks*2017.

Distinctions about where games fall in the cycle of research also suggest some of the tensions that will exist when a game is used to produce more than one type of information. For example, a game that seeks to both develop new solutions and judge their utility is likely to require a fairly advanced understanding of the space for innovation and produce relatively immature judgements because there are still gaps in our understanding. At its extreme, this principle also suggests that it will be very difficult to successfully produce information from the same game that explores a system and evaluates policy options. If there are still questions about the nature of the system fundamental enough to justify a game to explore, its unlikely that we have the necessary level of understanding to produce good judgments about potential interventions in the system. While there may be exceptions to this trend, it certainly holds well enough to serve as a heuristic for designers and sponsors about when a game is being asked to do too much.

Because the different types of games produce different types of information, often one particular aspect of the game process is of particular interest. As a general rule, the action of a game flows in a particular order: players receive information about the decision-making context, they debate what information matters to their decision and why, they make a decision, and then they observe the outcome of the decision in order to understand their new context and begin the cycle again. As the name implies, systems exploration games focus on the nature of the problem, so analytic attention focuses on what aspects of the game context matter to players and why. This can include how stakeholders view the same context differently, as well as how those understanding change over time. Alternative condition and innovation games examine player choices and the processes by which they are made, placing focus on the second two stages of the game. In the case of alternative conditions games, more focus may be placed on how game conditions influence player decisions, while in innovation games often there is a bit more focus on the decision itself though that can vary.[45] Evaluation games focus on the potential outcomes of decisions, and so have a unique focus on the last phase of the game process. While most games will still include all four phases of game play, one or more may be attenuated in design because it is of less importance to the game's focus.

Finally, there is a pattern in the profile of who the information from the game is for, and what they intend to do with it. The results of system exploration games tend to inform stakeholders who are trying to understand a problem set—generally focused on the sponsor (and to a lesser degree, players), these games are about understanding decision-making contexts, rather than about supporting a specific, immediate decision. Alternative conditions game results tend to be most helpful to in decision makers who are responding to or shaping global conditions. In some cases, the information from these games informs outside stakeholders who have influence over key variance, in other cases the goal is to understand the consequences of being in one context instead of another. Results of innovation games are most likely to be useful to investors who are

---

[4545] For example, a game that focused on innovation in the process for decision making might shift the focus of analysis.

determining areas for future research and development. These games help the sponsor decide where to invest in additional research, but on their own are usually insufficient to make the case for major investments. Finally, evaluation game results can inform policymakers trying to select a course of action. The size of decision, and thus the degree to which the information from the game must be persuasive to external audiences, can vary, but generally there is a sizable persuasive element in communicating the results of these games.

As stated earlier, it is possible for a game to resemble more than one archetype. However, several of these differentiating factors suggest that some combinations are more likely to be harmonious than others. That is not to say that it is not possible to combine types, but suggests that there will be serious tensions that need to be reconciled in the design process in order to make such a hybrid game work. If these tensions are poorly managed, one or both of the desired outcomes may be compromised. This makes it the responsibility of the game designer to communicate the tensions between potentially competing objectives, and example how design choices increase or decrease the chances of producing the desired information to sponsors.

Archetypes and the Logic of Repetition

Another way, perhaps more theoretical way, to differentiate the four archetypical categories of games is to consider why we might choose to run them multiple times. Put differently, what additional information can we obtain by repeating a game, and what does that imply about how we are learning from it? Too often, how often games are run is determined primarily by resources or other logistical considerations, rather than being determined by the research design of the study. There is a general perception that running games multiple times is preferable, but few clear articulations of what can be gained through repetition, how much analytic value it provides, and whether the gains are the same between different games. This section attempts to lay out a logic for what can be gained by repetition in the case of each of the four game archetypes.

In considering a more nuanced argument in favor of repetition, it is useful to draw analogy to other types of analysis, to consider how games are similar (and different) from the logic of other potential approaches to conducting research. In part, this is because relating games to existing approaches can help provide a common language to analysts and sponsors who may have familiarity with other types of analysis. This logic also works in reverse—without a deliberate comparison, consumers may incorrectly apply the standards of other methods to games. For example, a major criticism leveled against games (particularly by practitioners of operations research and statistics) is that because games are generally run only a small number of times, the results of the game are not statistically significant, and thus do not contribute valid knowledge. This criticism misses that the information created by games could (and I believe does) follow a different logic of inquiry then that of a model and simulation effort. Add to this the newly

articulated perspective that the logic of inquiry is not the same for all games,[46] it suggests that not all games need share the same logic of repetition. By exploring why we might want to repeat games, we can help clarify which of these logics apply to what types of games.

However, it is important to be clear that in drawing these comparisons, the unique nature of games remains. As noted earlier, games involve both inherently synthetic elements, and inherently uncontrolled interactions between players. The analogies made in this section are just that—highlighting similarities in logics in some respects. Appropriate consideration and caveats of the overall analysis are still necessary to product credible information from games.

System exploration games are repeated to add more viewpoints, and thus build a more comprehensive understanding of the problem. This process is analogous to conducting interviews—as you interview more people, you gain an understanding of what parts of their understanding are shared, and some suspicions about what might be driving differences. At some point, the marginal returns of conducting more interviews is minimal because you've already captured the vast majority of how experts see the problem, a process sometimes referred to as saturation. So to with system exploration games—additional runs of the game allow you to speak to new experts and watch different combinations of experts interact with one another, until the model built in the game is fairly stable. Depending on the initial game's make-up of players, this means that relative few, or even no repetitions of the game may be valuable—if the 80% solution is enough to inform the sponsor, then expensive iterations to get to 100% saturation may well not be a cost-effective use of a limited research budget. That said, running the game repeatedly increases our confidence that the results of the game mirror expert conscience.

Alternative condition games use repetition in order to explore more variation, without losing control over comparisons. For example, if there are two factors, each with two variations of interest but you can only run two games, you have to accept that the factors will co-vary and be difficult to disentangle. In contract, if you run four games, it is possible to compare games which are only one factor different to see if decision-making changes. This is analogous to the way that variation is treated in case study analysis—cases are often selected to maximize variation on the key variable of interest, while ensuring other variables are as comparable as possible.[47] However, like cases studies, it is critical to be cautious in over-generalizing the difference observed in these games. For one thing, game designers never have complete control over variation between games because of the key role of players. Thus, like case studies, analysis of repeated alternative conditions games should focus attention on exploring the evidence to support alternative hypotheses before making any type of strong claim from the game series. Because the games are

---

[46] Bartels, 2015

[47] Alexander George and Andrew Bennet, *Case Studies and Theory Development in the Social Sciences* (Boston, MA: MIT Press, 2005).

in key ways artificial, it is not possible to use approaches like qualitative comparison analysis,[48] where all possible cases of a class are examined, in order to draw stronger conclusions. Thus, repetition increases the clarity of findings, but may not increase confidence in them.

Innovation games have a very simple reason motivating multiple runs: novel, good ideas are hard to come by, and running the game more times provides more chances for a rare event to occur. Solutions that are likely to occur to stakeholders are also likely (though not always!) to have been discarded for good reason. To come up with a new answer requires looking at the problem differently, being willing to adjust the bureaucratic rules, or envisioning a new tool that has not been seen before. Any and all of these are difficult and may not happen in a particular game, no matter how well designed. Thus, running the game multiple time (particularly incorporating feedback about issues and concerns from past games) may increase the chances of a lightning strike of a good idea. Of course, anyone who has worked with a rare event model will tell you that it is all too possible to invest in multiple games, and never have the needed lucky break—more runs of the game may increase the chances of innovation, but they do not guarantee it.

Evaluation games, like classic models and simulation, are repeated to better understand the central tendencies of the model. However, this comes with two important caveats—the first is that like models and simulations, the results are the central tendency of the model, not of the actual phenomenon. To the extent the game's model may misrepresent key aspects of the phenomenon, the results will not be generalizable. However, unlike traditional models and simulations, the players of the game are a key component of the model, and they will be different during each run of the game. As a result, many runs of the game can build up a picture of what type of decisions players tend to make, and what the outcome of those decisions tends to be, but it is up to a skilled analyst to argue how well those results map onto the real world, and thus how confident we should be in assuming that game results are indicative of real outcomes.

On implication of these different logics is that different types of games vary in how difficult they are to deploy well. Systems exploration games use a fairly straight-forward logic, which we can make clear comparisons to well-known alternative approaches. It is not surprising that these games are generally considered the easiest to design and are very common across national security policy researchers.[49] Alternative condition games involve the intuitive logic of comparison, but the logistical realities of game introduce many caveats which must be managed by a designer. Innovation games have a straight forward logic but depend on a rare occurrence. This makes them straightforward, but risky to run. As a results, alternative condition and innovation games are seen as more difficult to design and are likely run less often then system

---

[48] Charles C. Ragin, *The Comparitive Method: Moving Beyond Qualitative and Quantitiative Strategies* (Oakland, CA: University of California Press, 2014).

[49] As noted above, accounting for games is immature, and thus these statements about relative frequency are based on expert judgement, supported by interview subjects, rather than systematic or numerical accounting.

exploration games. Final, evaluation games are seen as incredibly difficult to conduct in a credible manner and thus more rare still. Expectations should be set accordingly.

Design tradeoffs

As has been pointed out many times, there is no fixed recipe for moving from the game's purpose and objectives to its design. It is left to a designer to assemble mechanics, data, and people together to craft an appealing game. However, that is not to say that no guidance can be offered to the designer to steer them towards better and worse choices to achieve their objectives. Here, it can be helpful to think about the designer's job in terms of trade-off. No matter the purpose of the game, the design process seeks to build a game that instantiates a model of the problem at hand. Design choices can either align with, or deviate from, that model, making for better or worse design.[50] However, because games must be run in the service of practical ends, available resources in a wide variety of areas impose constraints that a designer must also work within. As a result, much of a designer's work requires making tradeoffs between what is dictated by the ideal research approach, and what is feasible given constraints.[51]

Given this frame, one way we can provide guidance to designers is to explore what tradeoffs are likely to be more or less problematic to the usefulness of findings, given a particular goal of a game. Because it is not usually possible to run a game in which no practical compromises are made, identifying the tradeoffs that are most likely to undermine our finding can allow for smarter design choices. When problematic choices cannot be avoided (as is often the case), advanced consideration can sometime develop mitigations within the research design, or at least allow for thoughtful discussion as part of analysis.

I organize the discussion of tradeoffs along the three key design elements that make up the model of the game: the environment, actors, and rules. The environment refers to the setting of the game that frame the central problem players seek to resolve. This includes not only the narrative scenario that traditionally describes the events leading up to the start of the game, but also the information provided to players about the state of the world during the game. This comes in a range of forms, included narrative, visualizations, and databases that together created the player's understanding of the setting.

Second is the actors, represented by players, who have resources they can use in an attempt to resolve the problem to their favor. The modeling of the actor includes the frame provided by the designer, such as the decision of which actors are represented and what level of aggregation, and what guidance is provided about each actor. However, perhaps more important is the human

---

[50] Elizabeth M. Bartels, Margaret McCown, and Timothy Wilkie, "Designing Peace and Conflict Exercises: Level of Analysis, Scenario, and Role Specification," *Simulation & Gaming* 44, no. 1 (2013).

[51] It is important to note that this is not a unique issue in gaming. For example, medical studies are often notable for quite small numbers of participants—researchers use the minimum number needed to test for a particular effect, and no more to save on time and costs of the study. Examples from other fields also come easily to mind.

players who fill the role, whose mental models fundamentally shape what choices they do, and do not, make in the game. Taken together, these components will describe what actors are in the game, their objectives, the resources available to them to pursue those objectives, and what decisions they can make in the course of game play.

Third is the rules that structure how the actors' decision interact with one another and the environment. This includes both rules that structure interaction during the game (for example, whom can communication with whom and how) and those that are used during adjudication to decide the outcomes of decisions. These rules can range widely from very rigid to open, and from complex to simple depending of the needs of the game model. They can also be implicit or defined by the players, such as when players assert that a particular action is not permissible Regardless of format, the game rules will shape how game play evolves, including how players learn information, how the make decisions, and the consequences for both actors and the environment of those decisions.

The following sections detail for each archetype what types of tradeoffs between and within game elements are most appropriate.

System Exploration

System exploration games are strongly defined by their players—since participants in the game provide the mental models to be captured, the quality and diversity of player understanding is absolutely critical for a strong game. The other aspects of the game environment, actors, and rules must be constructed in a way that enables and directs player input. If the structure of the game is too rigid. players do not have sufficient freedom of action to contribute their understanding of the problem. The game results will closely mirror the designer's understanding, and little will have been added by playing the game. Conversely, if the game is under-structured, it can lose focus—players speak to different problems, discussion turns away from decision-making and becomes more general, and key data will not be captured.

Environment

The primary purpose of the environment of a system exploration game is to focus player attention on the correct problem. This is inherently a tricky concern, as often different mental models will frame an issue differently. Picking the wrong scope for the game can cause analysts to miss key aspects of the problem, or in more extreme cases, cause participants to balk at participation because the problem is so mis-framed as to be unrecognizable.

Good design requires research on the part of the designer, not only of general subject matter resources but also on the perspectives of the different equities of the players. The more the different understandings of the players are understood in advance, the more carefully the game's environment can be framed to allow for productive dialog and deeper insights.

Often, because systems exploration games occur fairly early in the research process before the team has developed a very structured understanding of the system, it can be easier to use less

structured approaches to communicate about the environment. For example, it's fairly common for games of this type to focus on a written or briefed scenario, rather than developing a detailed game environment through structured game components like boards. This more open form of communication leaves more space for players to intersperse there own understanding of key concerns.

Actors

The most critical step in the design of a system exploration game's actors is the selection of players. Generally, in a systems exploration game, players should be recruited that have more experience than the designer, since otherwise the designer could simply assemble a model without the hassle and expense of running a game. Diversity is also key--if players all have the same perspective there is less benefit to the interactions and debate between different players. From an analytical perspective, robust debate is useful because it provides a natural opportunity to capture not only what players believe, but why they hold different views. Put differently, lucid debate is one of the easiest ways to elicit different player perspectives. Good facilitation can assist, but players with diverse, deep experience will be able to ask questions and raise issues that even the best-prepared facilitator will not know to ask.

However, selection of players will generally require making tradeoffs between the credibility of the resulting model and logistical realities of gathering individuals whose time is in high demand. The reality of politics, limited budgets, and set time-line will almost always prevent the designer from assembling the ideal group of players. Given that, very often the choice is about which players to prioritize, and where to accept a player who less closely resembles true decisionmakers on the issue.

Depending on the nature of the problem at hand, different types of player experience may be more or less important. For example, one could imagine a game populated entirely by senior interagency officials, and another consisting entirely of academic experts with decades of experience studying the conflict environment. Both bring different experience, and the resulting model would be credible and useful to different applications. We would expect that the model built by interagency figures would include a solid understanding of organizational equities and resources, and well as an understanding of the decision preferences of those higher up in the institution. In contrast, academic subject matter experts would produce a model that may benefit from structure provided by models from different academic disciplines and field and archival research. This model might be more credible with stakeholders outside the government. Ideally, it would be possible to bring the two communities together in a game that synthesized both perspectives. However, often issues like cost, scale, and classification prevent doing so. In these cases, a designer must consider carefully which perspectives will produce information that is of the most value to the sponsoring office. For example, if the sponsor is new to her office, understanding from other parts of the bureaucracy might be valuable in bringing them up to speed quickly on the operations of their organization. In contrast, a more experienced sponsor

might benefit more from voices outside the organization, which they do not have regular access to. Regular discussions with the sponsor about what will make the game valuable to them are a designer best guide to making these tradeoffs.

Beyond the selection of participants, considerations about which actors are explicitly or implicitly represented in the game are critical to systems exploration games, because they are key to scoping the problem. One choice is the degree of specificity at which different actors are modeled in the game. This can range from very specific guidance about what individual or office each player is intended to depict, to establishing broad teams representing countries or departments, to bringing players together with broad guidance such as "consider the relevant actors". Generally, less-structured approaches are used in systems exploration games, again, with the goal of allowing players to add their own expertise and experience. However, even a broad frame must be careful to scope the range of decisionmakers appropriately, or risk that the model emerging from the game will be missing major elements of the problem system. Ensuring that what cues are provided to participants about the scope of their roles, which can include the range of participants invited, the taskings provided to players, and the elements of the environment that are highlighted, can all play a role in shaping scope. Designers must be concisions of these potential effect during design, and during analysis must carefully consider how the game's scope may have produced a specific model.

Rules

Much like the environment, the rules of a systems exploration game are often somewhat unstructured to allow players greater freedom of action. Put simply, if a designer is still trying to understand the problem, it is not likely that they can pre-identify potential actions and their likely effects in enough detail to generate rigid rules. As a result, tools like free-form games,[52] matrix games,[53] card driven games, and other seminar-style techniques that allow a great deal of flexible interaction between different players, as well as players and adjudicators, are the norm. However, this creates an important burden on game analysts to document what rules were created by players and adjudicators, since they are such an important part of the model built during the game. Too often seminar-style approaches are left largely undocumented, rather than using the opportunity of the game to elicit implicit understandings and assumptions about who can (and cannot) produce what influences.

Alternative Conditions

The design of alternative conditions games rests fundamentally on the need to create comparisons that allow the game analysis to highlight similarities and differences between games

---

[52] William M. Jones, *On Free-Form Gaming*, ed. Corporation Rand (Santa Monica, CA: RAND, 1985).

[53] John Curry and Tim Price, *Matrix Games for Modern Wargaming: Developments in Professional and Educational Wargames* (History of Wargaming Project, 2014).

which operate under different conditions. Because games provide the flexibility for designers to make many choices about the set-up of the game, but have limited ability to guarantee variation in player decision, alternative condition games often follow the pattern of a "most similar" comparative designs. In a most similar design, cases are selected to be as similar as possible except for variation in the independent variable, which is then used to explain any differences in the outcome variable.[54] Since a game designer will have a relatively strong ability to ensure variation in starting conditions, but will not be able to ensure a diversity of outcomes, a most similar approach increases the chances that the games will produce data that aligns with the research strategy. The down-side is that the strength of findings will depend to a large degree on the extent to which the analyst can argue that variations in the games beyond the independent variable do not offer an alternative explanation for any variation that occurs in the dependent variable. As is well discussed elsewhere, this type of comparison does not tend to produce strong support for causal claims.[55]

Generally, design of an alternative conditions game starts with the identification of the key variable(s) of interest, which will be changed between games. For example, the environment of the game might be varied by providing scenarios that describe differences in the context of the crisis. Depending on the research question, these might be quite narrow (for example, changing the number of initial casualties in an instigating event) or broad (selecting different environments in which a crisis over water could occur).[56] Similarly, the actors can be varied either by changing the identity, objectives, and resources available in the starting conditions, or through careful selection of the human players representing each role. An important example of this is changing the capabilities available to different forces at the start of the game. Varying the actions available to players can allow for investigations of the impact of processes (such as communication protocols or deliberative protocols). Regardless of where in the game the key variation is located, the goal of the designer is to vary that factor, and only that factor, between plays of the game.

However, this degree of control can be somewhat illusionary, as requirements imposed by both perceptions of plausibility and the logistical realities of running an event with human players can make it exceedingly difficult of maintaining control over variables that could provide an alternative explanation. Specific considerations for the depiction of the environment, actors, and rules follows below.

---

[54] George and Bennet, *Case Studies and Theory Development in the Social Sciences*., p 81

[55] The possible exception being when the cases also take on characteristics of a single case selection criteria like a critical case (again, see: ibid.). For example, if the game is set up as a "most likely case" where we would expect to observe a predicted relationship, and it does not occur, that offers somewhat stronger evidence that the hypothesized relationship is wrong. However, it is important to be clear that the value of the inference is driven less by the comparison between cases, then from observing the flow of events within the game.

[56] Bartels, McCown, and Wilkie, "Designing Peace and Conflict Exercises: Level of Analysis, Scenario, and Role Specification."

Environment

Comparative control over the environment is somewhat more under the control of the designer, but requires carefully thinking though second order effects of variance to create coherent scenarios. Unless a game is looking at a quite small variation in independent variables, it is quite likely that variation in the environment, actors, or available actions will necessitate making other changes to the environment to ensure narrative plausibility. For example, in games comparing two different sets of military capabilities, there may also be changes in the relative economic performance as a result of civilian uses of technology. The key here is to consider both what variation is demanded by plausibility, and what variation threatens to complicate the comparative narrative. If additional variation between environments that is required for coherence could provide an alternative explanation, this should be addressed explicitly in analysis.

Actors

Actors are a particular concern in alternative condition games because the variation between human players makes it difficult to create controls between multiple groups. Common solutions are either to have the same group of players play multiple games, or to attempt to recruit multiple comparable groups of participants. In the first case, it is then necessary to account for learning between the rounds of play. For example, if the same group of players is presented with the same decision context twice, we would expect that their decision would in part be influenced by the outcomes of their decisions in the first round of play. There is also the challenge of maintaining engagement and interest in multiple rounds of truly identical play, which often requires that additional variations be introduced. The second option is to try to recruit multiple groups of players. This can be practically attractive, because it allows multiple games to be run in parallel at the same time. However, for the comparison to be credible, the designer must be able to define what characteristics of the players are salient, and defend why the two groups are similar on these dimensions. This can be quite challenging, given the range of experience that may shape decision making, and the limited knowledge about player background that game designers may have available. Regardless of which option is selected, it is extremely unlikely that difference in player experience, attitudes, and beliefs will not play a role in shaping decision making. As a result, it is generally best to treat these factors as an alternative explanation for differences in decisions, and explicitly discuss why the patterns of discussion support that the independent variable, rather than participants, is driving differences in decisions. Tools like process tracing or analysis of player discourse[57] can be particularly powerful here.

Rules

How much of a concern variation in available rules might be will vary considerably based on how formalized the available actions and means of determining outcomes are. To take one

---

[57] For a recent example of this technique, see: John Derosa and Lauren Kinney, "Narrative Analysis of Wargaming," in *Connections Wargaming Conference US* (Washington, DC2018).

extreme, if players can take any action that occurs to them, player perceptions about what actions are available may differ considerably between different groups of players. If players are also involved in adjudication (as for example, in a matrix game) this concern also extends to the projected outcomes of actions. If on the other hand, available actions and adjudication are highly formalized, it is much easier to ensure consistency between rounds of play. However, the choice of how structured a game system to use is often dictated by the existing knowledge base about the problem. For example, there is a great deal of available information about the capabilities of existing weapons systems, and very little about how well potential future weapons systems might work. Comparing the two will inherently introduce uneven information, and in the case of the future systems expert players may be better positioned to project capabilities than the design team.[58] In this later class of cases, what players can do, and what the impact of those actions are on other actors and the environment, may need to be left less specified, and any resulting variation accounted for in subsequent analysis.

Innovation

The core challenge of innovation games is to increase the chances of players producing an innovative idea. The designer needs to loosen some status quo constraints to create space for new solutions to emerge but must be careful to retain constrains needed to ensure realistic outcomes. At the same time, the designer needs to frame solving the problem to be enticing enough to motivate player creativity and engagement. Often this requires a strong competitive element in the game.[59]

Environment

While some aspects of the game's environment may be helpful to alter to open up space for innovation, as a general rule many aspects of the environment are outside the control of any of the game's stakeholders, and thus are somewhat less likely to be a promising avenue for innovation. Instead, the environment is more likely to be a source of constraints that players must work within. For example the laws of geography and physics limit speed and distance—credible innovations must account for both. What is more important is that the game environment clearly frames the problem in a way that motivates player engagement and problem-solving. A compelling narrative about the environment, and particularly the nature of the central conflict of the game, can help support the stakes of actors.

---

[58]For a more detailed discussion of this problem, see: Stephen Downes-Martin, "Adjudication: The Diabolus in Machina of War Gaming," *Naval War College Review* 66, no. 3 (2013).

[59] For thoughtful discussion about the role of competition in innovations games, I am grateful to Graham Longley-Brown (interview Washington DC, August 2018) and Phil Pournelle (interviewed Washington, DC, March 2019).

Actors

Competition is a key component in this archetype. In order to motivate players to discover new ways of acting, the game must motivate them with a problem. Most often, this comes in the form of a thinking, reacting adversary; however depending on the problem, it could also be found in semi-cooperative relationships in which the different motivations between offices, departments, or partner countries provide tension as players strive to overcome an environmental challenge such as a natural disaster or disease. In part, this motivation is psychological. However, competitive dynamics in the game also aid the task of innovation game analysis, by providing a "test" of the new ideas. Competitive perspectives incentivize players to highlight potential flaws in suggested ideas, and to manifest unintended consequences. This can allow the "dominant strategy" to emerge from a set of potential ideas.[60]

Often when recruiting players for innovation games, it can be most productive to empower individuals who are not currently in positions of power. More junior voices, or those from less powerful organizations, or those who hold unusual perspectives have fewer opportunities to air and implement their own ideas, and thus are somewhat more likely to develop options that have not already been considered.

Rules

Rules, particularly rules that represent the current processes and procedures of an actor, are a particularly fruitful area to loosen constraints. Changing processes and procedures is often within the purview (or at least the realm of possibility) for senior game stakeholders, thus innovations that require changing rules are often evidently feasible, and thus of greater interest than innovations that depend on changes to the environment or actors that fall further outside stakeholder control. However, it is important to note that in any national security issue, ease of changing rules and procedures vary dramatically, so an important caveat to analysis of the game should be a sense of the other consequences of the change.

Evaluation

Evaluation games are most strongly defined by their rules. Because games in this archetype are concerned with producing results that examine the outcomes of player decisions, the credibility of the adjudication process is critical to the usefulness of the information. At the same time, the more-developed understanding of the problem required to build a credible adjudication engine also means that many of the questions about the scope and depiction of the environment and actors are somewhat more settled than in the case of the other archetypes. Generally, these descriptions will be driven more by the information needs of the rules, and thus play a somewhat supporting role in design (though in the case of complex models they can still be quite time intensive to produce in practice).

---

[60] Interview with Phil Pournelle, game designer, Washington, DC, March 2019.

Environment

The environment of most evaluation games is driven by two primary considerations: the information needed to produce credible outcomes under the game's rules, and the contextual information players need to appropriately apply courses of action. As a general rule of thumb, more credible adjudication rules will depend on a more robust environment that can represent the anticipated outcomes of decisions in sufficient detail to support analysis. However, this will depend to a large degree on the preferences and beliefs of the audience

Actors

Because of the high degree of structure of the environment and rules, evaluation games can often draw on more junior players, though familiarity with standard practices in the relevant domains is still key. This is particularly true in cases where the contours of the course of action to be evaluated is provided to players. However, as always, the extent to which designers ought to be willing to sacrifice experience for other considerations will be determined by what types of players are seen as credible by the consumer of analysis. For example, sponsors of operational military games almost always want to see military staff experience from all relevant services among the players.

Rules

The qualities of a "credible" adjudication model are not one size fits all—instead, the designer needs to consider what will be credible to the specific audiences of the game results. One particularly important tradeoff to consider is that between complexity and transparency. There is a long-standing literature with in defense modeling that argues that as models gain complexity, it becomes all too easy for modeling assumptions to interact in ways that produce chaotic results.[61] In other words, as models become more complicated, it becomes more difficult to trace why a particular result is generated. As a result, for an audience that values the ability to explain *why* a game's adjudication rules produced the result that it did, its is generally preferable to use a simpler model that allows the designer to explain the logic of how player actions produced the outcome in question. However, the very existence of a literature on the negative effects of complexity in defense modeling points to the entrenchment of an opposing view, that argues that more complex models are better simulations of real world phenomena. In order to convince stakeholders who hold this perspective of the usefulness of a game's results, it may be necessary to demonstrate the complexity of the games rules.

---

[61] For two well-known critiques of this type, see: Paul K. Davis, *The Base of Sand Problem : A White Paper on the State of Military Combat Modeling*, ed. Don Blumenthal, et al. (Santa Monica, CA: RAND, 1991); J.A. Dewar, J.J. Gillogly, and M.L. Juncosa, "Non-Monotonicity, Chaos, and Combat Models," (Santa Monica, CA: RAND Corporation, 1991).

Concluding Thoughts

Having presented a framework that defines different types of games based around the logic that connects observations from the games to desired types of findings, it is worth taking a step back to consider how this product can benefit key communities. Due to the non-empirical nature of the framework's development, it will always be open to additional refinement, and perhaps with time alternative approaches, the desired end-point is for the framework to be useful to key gaming communities. In particular, I want to close by considering how game designers, academics new to gaming, and policy-makers can leverage this approach.

Perhaps most straightforward is the implications for game designers. The framework offers a logic to help guide design and analysis choices. Hopefully, designers can use these as a reference point for their work by comparing new gaming projects not just to past games, but to these idealized types. For new gamers, I hope these will help expand the range of the possible, For all gamers, I hope the language and concepts can form a common reference point to continue to grow and deepen the body of work devoted to documenting gaming. In short, I hope that the framework contributes to the establishment of a more systematic, and thus more scientific, approach to gaming.

In contrast, I hope academics new to gaming will find this discussion of the logics of inquiry to make games a more accessible and accepted tool for research. By providing logics of research design for different types of games I hope to make it clearer that games can contribute new insights that advance our understanding of core issues in international studies. I also aim to provide some basic guidance on how these approaches compare to more mainstream techniques to provide some reference points between the communities. My hope is that taken together, these advances will make it easier for social scientist to incorporate games appropriately into their work.

Finally, for the policy community I hope this work will provide guideline on what types of questions game can, and perhaps more importantly, cannot answer. Recent senior leader interest in gaming has given rise to concerns that games will over promise and under deliver.[62] Sponsors with a critical eye on games are key to preventing that. I also hope that this approach will help change games from data points into things that are part of broader research programs and policy decisionmaking process. By using games in a smarter way, policymakers will be able to have greater confidence that the information they learn from games is useful.

---

[62] See: Perla, "Now Hear This—Improving Wargaming Is Worthwhile—and Smart."and Pettyjohn and Shlapak, "Gaming the System: Obstacles to Reinvigorating Defense Wargaming."

References

Bartels, Elizabeth M. "Building a Pipeline of Wargaming Talent: A Two-Track Solution." *War on the Rocks*, 2018.

———. "Gaming – Learning at Play." *ORMS Today* 41, no. 4 (2014).

———. "Insights from a Survey of the Wargaming Community." In *Military Operations Research Society Wargaming Community of Practice*. Alexandria, VA, 2017.

Bartels, Elizabeth M. "Adding Shots on Target: Wargaming Beyond the Game." *War on the Rocks*, 2017.

Bartels, Elizabeth M., Margaret McCown, and Timothy Wilkie. "Designing Peace and Conflict Exercises: Level of Analysis, Scenario, and Role Specification." *Simulation & Gaming* 44, no. 1 (2013): 36 - 50.

Barzashka, Ivanka. "Wargaming: How to Turn Vogue into Science." *Bulletin of the Atomic Scientists*, 2019.

Brown, Graham Longley. *Successful Professional Wargames: A Practitioner's Guide.* Edited by John Curry. The History of Wargaming Project, Forthcoming.

Burns, Shawn, ed. *War Gamer's Handbook: A Guide for Professional War Gamers*. Newport, RI: War Gaming Department, U.S. Naval War College.

Caffrey, Matthew B. *On Wargaming: How Wargames Have Shaped History and How They May Shape the Future.* Newport, RI: Naval War College Press, Forthcoming.

Compton, Jonathan. "Towards an Epistimology of Wargaming--a Drunkard's Walk." Paper presented at the Military Operations Research Society Wargaming Community of Practice, Alexandria, VA, 2015.

Curry, John, and Tim Price. *Matrix Games for Modern Wargaming: Developments in Professional and Educational Wargames.* History of Wargaming Project, 2014.

Davis, Paul K. *The Base of Sand Problem : A White Paper on the State of Military Combat Modeling.* Edited by Don Blumenthal, Institute National Defense Research, Corporation Rand, Defense United States. Department of Defense. Office of the Secretary of and Agency United States. Defense Advanced Research Projects. Santa Monica, CA: RAND, 1991.

Derosa, John, and Lauren Kinney. "Narrative Analysis of Wargaming." In *Connections Wargaming Conference US*. Washington, DC, 2018.

Dewar, J.A., J.J. Gillogly, and M.L. Juncosa. "Non-Monotonicity, Chaos, and Combat Models." Santa Monica, CA: RAND Corporation, 1991.

Downes-Martin, Stephen. "Adjudication: The Diabolus in Machina of War Gaming." *Naval War College Review* 66, no. 3 (2013): 67-80.

———. "Your Boss, Players and Sponsor: The Three Witches of War Gaming." *Naval War College Review* 67, no. 1 (Winter 2014 2014): 31-50.

Dunnigan, James F. *Wargames Handbook: How to Play and Design Commercial and Professional Wargames.* 3rd ed.: IUniverse, 2000.

George, Alexander, and Andrew Bennet. *Case Studies and Theory Development in the Social Sciences.* Boston, MA: MIT Press, 2005.

Jackson, Patrick Thaddeus. "The Conduct of Inquiry in International Relations: Philosophy of Science and Its Implications for the Study of World Politics." New York, NY: Routledge, 2011.

Jones, William M. *On Free-Form Gaming.* Edited by Corporation Rand. Santa Monica, CA: RAND, 1985.

Klein, Gary. *Sources of Power: How People Make Decisions.* Cambridge, MA The MIT Press, 1998.
Lin-Greenberg, Erik. "Game of Drones: What Experimental Wargames Reveal About Drones and Escalation." *War on the Rocks*, 2019.
Morse, Philip M., and George E Kimball. *Methods of Operations Research.* 3rd ed. Mineola, NY: Dover Publications, Inc, 2003.
Parson, Edward. "What Can You Learn from a Game?". In *Wise Choices: Decisions, Games, and Negotiations*, edited by Ralpj L. Keeney Richard J. Zeckhauser, James K. Sebenius, 233-52. Boston: Harvard Business School Press, 1996.
Pauly, Reid B.C. "Would U.S. Leaders Push the Button? Wargames and the Sources of Nuclear Restraint." *International Security* 43, no. 2 (2018): 151-92.
Perla, Peter. "The Art and Science of Wargaming to Innovate and Educated in an Era of Strategic Competition." In *King's College London Wargaming Network Lecture*. London, UK, 2018.
Perla, Peter P. *The Art of War Gaming : A Guide for Professionals and Hobbyists.* 2nd ed.: History of Wargaming Project, 2011.
———. "Now Hear This—Improving Wargaming Is Worthwhile—and Smart." *Proceedings Magazine*, January 2016 2016.
Perla, Peter P., and ED McGrady. "Why Wargaming Works." *Naval War College Review* 64, no. 3 (2011).
Pettyjohn, Stacie L., and David A. Shlapak. "Gaming the System: Obstacles to Reinvigorating Defense Wargaming." *War on the Rocks*, February 18, 2016 2016.
Pournelle, Phillip E. "Designing Wargames for the Analytic Purpose." *Phalanx* 50, no. 2 (2017): 48-53.
Ragin, Charles C. *The Comparitive Method: Moving Beyond Qualitative and Quantitiative Strategies.* Oakland, CA: University of California Press, 2014.
Schneider, Jacquelyn G. "What War Games Tell Us About the Use of Cyber Weapons in a Crisis." edited by Council on Foreign Relations, 2018.
Senge, Peter M. *The Fifth Discipline: The Art and Practice of the Learning Organization.* Revised and Updated edition ed. New York, NY: Doubleday, 2006.
Smith, Kevin B. "Typologies, Taxonomies, and the Benefits of Policy Classification." *Policy Studies Journal* 30, no. 3 (2002): 379-95.
Staff, Chairman of the Joint Chiefs of. "Officer Professional Military Education Policy." Washington, DC, 2015.
Svet, Oleg, and Garrett Heath. "How the Joint Staff Calculated a Defense Program's Return on Investment." *Defense One*, 2018.
Williams, Bob, and Richard Hummelbrunner. *Systems Concepts in Action: A Practitioner's Toolkit.* Stanford, CA: Stanford University Press, 2011.
Wong, Yuna. "Preparting for Contemporary Analytic Challenges." *Phalanx* 47, no. 4 (December 2014 2014): 35-39.
Work, Robert, and Paul Selva. "Revitalizing Wargaming Is Necessary to Be Prepared for Future Wars." *War on the Rocks*, December 8 2015.