Unethical Professional Wargaming

"make your wargame say what you want it to say"

Connections US 2021 Wargaming Conference

US Command and General Staff College (CGSC)
and
The CGSC Foundation

Final Report 13th September 2021

Contributors

Hiroyasu Akutsu, Antoine Bourguilleau, Chad Briggs, Daniel Brooks, Rex Brynen, Kayla Capps, Jon Compton, John Curry, Stephen Downes-Martin, Mark Flanagan, Patrick Garrett, Amna Greaves, Zoltán Harangi-Tóth, Michael Hugos, Montana Hunter, Anne M. Johnson, Nina Kollars, Joel Kurucar, Graham Longley-Brown, Andrew McGee, Ed McGrady, William Owen, Peter Perla, Phillip Reiman, Ian Robinson, Robert Seater, Mark Simpkins, Ben Taylor, Paul Vebber, Kristan Wheaton

Working Group Chair

Stephen Downes-Martin

The contents of this document represent the opinion solely of the contributors and do not represent the policy or position of any organization.

Working Group members maintain full and exclusive intellectual property rights over their contributions. They have granted permission for their contributions to be included in this report.

Any errors, misrepresentation or misinterpretation in this document are the sole responsibility of Stephen Downes-Martin (stephen.downesmartin@gmail.com).

Table of Contents

Executive Summary

Stephen Downes-Martin, PhD
Research Fellow, US Naval War College

The Connections US 2021 Wargaming Conference was hosted by the US Command and General Staff College, with support from the Command and General Staff College Foundation for which the Connections wargaming community is profoundly grateful. As is always the case, the Conference host sets the conference theme, and this year the CGSCF chose "Ethics and Wargaming" as the theme.

In support of the theme the Connections Planning Committee started a Working Group on "Unethical Wargaming", the focus of which was a thought experiment that examined

"how to use unethical practices to make your wargame say what you want it to say"

where by "unethical practice" we mean any practice motivated by a desire to influence the sponsor to make a decision in the best interest of the unethical practitioner instead of the best interest of national security. We know intellectual dishonesty occurs in science and among senior civilian and military leaders. It is irresponsible to assume it does not exist within professional wargaming (or indeed any process that manages any inquiry activity).

This thought experiment is useful in three ways for

1. discovering wargame design principles and malfeasance that wargame designers, practitioners, sponsors, players and other stakeholders might miss if one focused on best practices of well-designed games by well-intentioned competent experts.
2. inoculating wargaming against deliberate and inadvertent manipulation of wargame design by senior stakeholders
3. protecting ourselves from self-deception by our own inadvertent malign practice.

An international group of thirty wargamers with experience and qualifications in Government, Military, Intelligence, Social Science, Industry, Academia, Law and Education started work in October 2020 and wrapped up in July 2021. The nine month period of performance allowed the group time to think, discuss, challenge, write and refine, and to do so in depth. The group produced seven research papers (and nineteen additional detailed discussion topics), all of which examined various aspects of unethical professional military wargaming and proposed methods for ensuring ethical gaming occurred. The papers dealt with:

"An Unethical Wargaming Taxonomy for Practitioners and Observers". Robert Seater and Joel Kurucar examine conscious and unconscious unethical behaviors and produce a classification system for cheating, bias, and sloppy analysis in the form of a taxonomy for cataloguing, organizing, and documenting types of unethical behaviors in wargames.

"How to Build Unethical Wargames". Ed McGrady considers the internal and external factors that affect the ethical environment for a game. He examines issues such as whether the players are treated fairly, is the game objective, and whether the game is used by others in an ethical way.

"Exploit Group Dynamics to Corrupt a Professional Wargame". Stephen Downes-Martin explores methods for corrupting the inner and outer wargame by exploiting the dynamics of small group discussions (that are a common feature of wargames designed for senior officers) in the form of a game of "design moves" between the unethical and the ethical game designer.

"Unethical Wargaming: Let us be Incompetent!" William Owen makes the claim that incompetence is the most common unethical behavior present in the wargaming community since wargames are supposed to influence battlefield decisions, and so if done badly puts missions and friendly lives at risk.

"The Layers of Ethics Surrounding Wargaming". Hiroyasu Akutsu draws on the history of Japanese wargaming from the late 19th Century to present times to create a social hierarchy of ethical layers and then examines how Japanese culture and history interacted to produce different types of unethical wargaming in each layer.

"Unethical from the Beginning". Zoltán Harangi-Tóth examines the lessons learned from wargaming in the Central and Eastern European region during the Soviet occupation, the damaging effects on wargaming of certain cultural proclivities, and how these even today express themselves in recent wargames and exercises in the region.

"Can AI Save Us from Unethical Wargames?" Robert Seater examines three ways in which wargames can run afoul of ethical boundaries and how an impartial Artificial Intelligence (AI) can help mitigate those risks by considering game design as a form of human-machine teaming.

Core Conclusion: Most professional wargames are vulnerable to unethical practice due to the presence of the three established criteria for intellectual fraud. The lack of familiarity by senior officers or civilian executives with the unethical practices described in this report means we cannot say that most DoD wargames are free of unethical practices.

What is to be Done? By examining the interactions between the wargame stakeholders in the external environment, the outer game and the inner game, along with the the three criteria for the presence of intellectual fraud, and taking culture into account, we can increase the value and ethical probity of wargaming and ensure the decisions that the wargames are designed to influence are in the best interest of national security. Details are in this document.

Working Group Papers

An Unethical Wargaming Taxonomy for Practitioners and Observers[1]

Unethical Wargaming Working Group
Connections US 2021 Wargaming Conference

Robert Seater
MIT Lincoln Laboratory
robert.seater@ll.mit.edu

Joel Kurucar
MIT Lincoln Laboratory
joel.kurucar@ll.mit.edu

Abstract

The unique value provided by wargames stems from their open ended nature. Even structured wargames leave room for creativity and innovation in tactics and maneuver. However, the open-ended nature of wargames, inherent to the benefits they provide, also exposes their creation, execution, and analysis to unethical behavior, both conscious and unconscious. Cheating, intentional bias, and sloppy analysis can all distort the conclusions reached from a wargame, doing harm to the field or topic it was created to support. At a bare minimum, the unmitigated potential for such activities undermines confidence in well-run wargames. We broadly classify those types of wrongdoings under the umbrella of unethical behaviors for pragmatic purposes, acknowledging the complex gray areas that surround any formal definition of ethics.

Avoiding unethical behaviors is not just a matter of being aware of them. Combating ethical issues, like any other organizational ill, is about culture, process, and systematic approaches in the face of problems that are fundamentally not formal. As one piece of that solution, we propose a taxonomy for cataloguing, organizing, and documenting types of unethical behaviors in wargames. The aim is twofold:

(a) **Systematic List**. This effort aims to provide a way to frame and understand the myriad ways in which a wargame may become unethical. A structured schema can help one be thorough and avoid careless omissions while also prioritizing, organizing, and visualizing an overwhelming laundry list of potential micro-concerns.
(b) **Accountability**. Outline a practical framework for determining roles and responsibilities to reinforce a culture of self-assessment and pride in rigor.

The provided taxonomy and examples are meant to guide practitioners and those in the wargaming community in critically examining their games for vulnerabilities.

[1] DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.
This material is based upon work supported by the United States Air Force under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Air Force.

1 Goals and Motivation

Wargames are a powerful tool used across the DoD enterprise. A significant part of the value of wargames stems from their open ended nature. Even very structured wargames leave room for creativity and innovation in experimenting with operational art and the decisions made by players, which can manifest as novel tactics, maneuvers, or other related concepts. Most wargames are not fully structured and allow for some form of improvisation and adaptation by the moderators to accommodate, encourage, and elicit new ideas. After all, a primary benefit of a wargame is to discover viable outcomes, not necessarily to predict the most likely outcome.

The goal of this paper is to provide a tool for practical use to help safeguard the creation, execution, and analysis of wargames from unethical behavior. To achieve this we provide a taxonomy to help identify unethical issues in wargaming as well as a set of common roles and responsibilities to foster accountability. Together these ethical controls and guidelines are meant to provide basic quality standards, not provide overly philosophical rhetoric. In addition, these tools are constructed in a general manner so that others can apply these tools to their specific wargaming communities. This work is intended to organize concerns and safeguards discussed elsewhere, particularly from within the working group that published this work, bringing together a wide range of potential violations into a single systematic structure [1, 2, 3, 4, 5, 6].

Our primary claim is that being aware of ethical risks is not sufficient protection against violations; explicit recognition of ethical oversight is also required. Such oversight might come from a single authority or a communal responsibility, but it must be explicitly allocated and associated with appropriate authorities. A taxonomy provides a systematic way to organize and consider unethical wargaming behavior, but explicit roles and responsibilities is a necessary component for putting such a taxonomy into practice. In this paper we will provide a common set of roles and responsibilities for ensuring ethical wargaming, detail our unethical wargaming taxonomy, and provide some examples to help guide wargame practitioners.

We take a practical approach to what actions we label as "unethical" as definitions vary and are based on intangibles and subjectives such as intent. For this analysis, we focus on outcomes since intent is impractical to sense. As such we often label some things as "unethical" that some may not consider such. For example, if a person makes a mistake, is this "unethical"? In this analysis we included mistakes since it's too difficult to tell mistakes apart from bad intentions. Section 4.3, "Is this Ethics?" provides a more detailed discussion on this topic and decision.

2 Taxonomy Structure

2.1 Terms and Definitions

We will classify unethical behavior broadly in terms of the *perpetrating role*, that is, the person or people who may exhibit unethical behavior, and the *temporal phase*, that is, at what stage of the wargame the unethical behavior is likely to occur. Some role-phase combinations are more pertinent than others, so we will focus attention on the most problematic combinations.

2.1.1 Stakeholder Roles

Roles need not be individual humans; it is common for one human to fulfill multiple roles, it is possible for an organization or committee to play some roles, and it is possible for non-human automation to play some roles. A role bears personal responsibility for ethical due diligence, whether that role is a person, a fragment of a person, a group of people, or an algorithm created by a person. *Note: Multiple distinct roles may be listed together. These roles are distinct but during our analysis, we found these roles had significant overlap in terms of unethical behaviors and are therefore grouped accordingly for brevity (to keep the list manageable).*

- **Sponsor**. Provides funding and defines success for the project. Provides access to a recruitment pool. Ultimately responsible for justifying the existence of the game by delivering value to the content consumer.
- **Designer, Developer**. Oversees the creation of a working game, including rules, software support, scenarios, and keeping to time. Ultimately responsible for the game adequately addressing the target topics and providing useful output.
- **Moderator, Adjudicator, White Cell, Game Director**. Oversee the execution of the game, execute the structured rules, provide expert and informal adjudication in unstructured or corner cases. Ultimately responsible for ensuring the game runs as designed.
- **Red, Blue, (Purple, Orange, Green, etc...) Participant**. Active participants in the game. Makes decisions, trying to achieve their stated objectives within their stated constraints. Ultimately responsible for trying their best to win within the bounds of the game.
- **Gray, Pink Participant**. Passive participants in the game with the goal of behaving realistically, but not necessarily with any particular objective or win condition (e.g. as a civilian population or force of nature). Ultimately responsible for providing a realistic and responsive backdrop to the main (red and blue) participants.
- **Prior Participant**. Anyone who participated in a previous iteration of this or a similar game, but who does not have an official role this time around. No official responsibility due to being 'spoiled' with too much knowledge.

- **Analyst, Data Collector, Data Reporter**. Collect data of any form (including actions, opinions, and observations), analyzing that data, and presenting it. Ultimately responsible for accurately and meaningfully summarizing the events of the game.
- **Content Consumer**. Read the official report and listen to unofficial reports from the wargame to apply the findings to some application. Ultimately responsible for deciding how to act on the outcomes of the wargame.
- **Observer**. Observe the wargame as it executes as a 'fly on the wall'; socialize their observations either informally or in a formal role defined in the game rules. No official responsibility but will often overlap with other roles.

It is common for one person to play multiple roles. For example, a content consumer will often also show up in person as an observer, a designer will often help moderate, the sponsor might also be the consumer, a prior participant might be recruited as a gray cell participant, and a pink cell participant (when the red cell is run by the white cell) is by definition overlapping red and white. Doubling up is often unavoidable and often beneficial, although one class of problem we will discuss (Section 3.10, Role Confusion) stems from participants confusing their responsibilities and obligations across multiple roles.

2.1.2 Lifecycle Phases

Phases of a wargame need not be strictly linear, and they might overlap or occur in parallel. A role might participate in one or more phases, and it is rare for a phase to only involve one role.

- **Pre-Creation**. Creating the project, identifying success criteria, allocating funding, and assembling the team.
- **Design**. Designing the game, developing software support, planning logistics, and preparing scenario content.
- **Pre-Briefing**. Recruiting participants, issuing pre-game surveys, providing background materials, and teaching the rules.
- **Execution**. Running the game itself, adjudicating participant decisions, collecting data and observations, and providing feedback to participants.
- **Hotwash**. Running a hotwash with participants, providing quick summaries to observers, observers writing up informal reports for unofficial venues, and any form of reporting on conclusions prior to the official analysis and final report.
- **Analysis and Communication**. Analysis of data and observations collected during the game, interpretation of findings, reporting of findings, *creation and communication* of a final report or brief, and making official recommendations.

- **Repetition**. Running the game multiple times with different participants, creating a campaign of related games for a broader analysis, trying to reproduce earlier findings, or running a series of short stand-alone games in one sitting. *Note: for the purposes of this taxonomy, this phase just covers planning what form of repetition to use, not all the details of the actual execution of those repetitions. Executing a repeat session involves the other phases already listed.*

2.2 Role Matrix

Consider the matrix of roles vs. phases depicted in Figure 1. For each phase role-phase combination (a cell in the matrix), that role might play a primary role in that phase, be involved in that phase, passively observe that phase, or have no connection to (or not exist during) that phase.

- Primary. The role who leads that phase and is primarily responsible for that phase successfully executing. There is exactly one primary role per phase, keeping in mind that that role might comprise several collaborating people.
- Involved. A role who actively contributes to that phase, in support of the primary.
- Passive. A role who observes the phase but does not contribute or intervene.
- Absent. A role who cannot participate in that phase – they either don't exist during the phase or are explicitly excluded. Plain white cells are considered "absent".

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	primary	involved	passive	passive	passive	involved	primary
designer / developer	involved	primary	involved	passive	involved	passive	involved
moderator / adjudicator / white cell		involved	primary	involved	primary		
red / blue participant			involved	primary	involved		
gray / pink participant			involved	involved	involved		
prior participant	involved	passive			passive		passive
analyst			passive	passive	passive	primary	involved
content consumer					passive	involved	passive
observer			passive	passive	passive		passive

Figure 1: This table shows functional responsibilities and level of involvement at each phase. Schematizing functional roles and responsibilities will also inform appropriate assignment of roles and responsibilities for monitoring and managing ethical issues.

2.3 Roles and Responsibilities

Functional roles are not the same as ethical roles, but they are related. For any given phase, we should ask two questions about ethics:

A. Who has responsibility for (avoiding) ethical violations?

B. Who monitors the project to ensure the responsible parties remain ethical?

For each question, we consider three possible answers:

1. The primary functional contributor, on the grounds that they are best qualified (although they may also have the most biased perspective).
2. A shared responsibility of everyone involved, whether leading, contributing or passively observing, regardless of official authority or rank.
3. An external and independent authority (who has no direct bias, but who may be overly harsh in order to justify their existence).

These options are depicted in Figure 2.

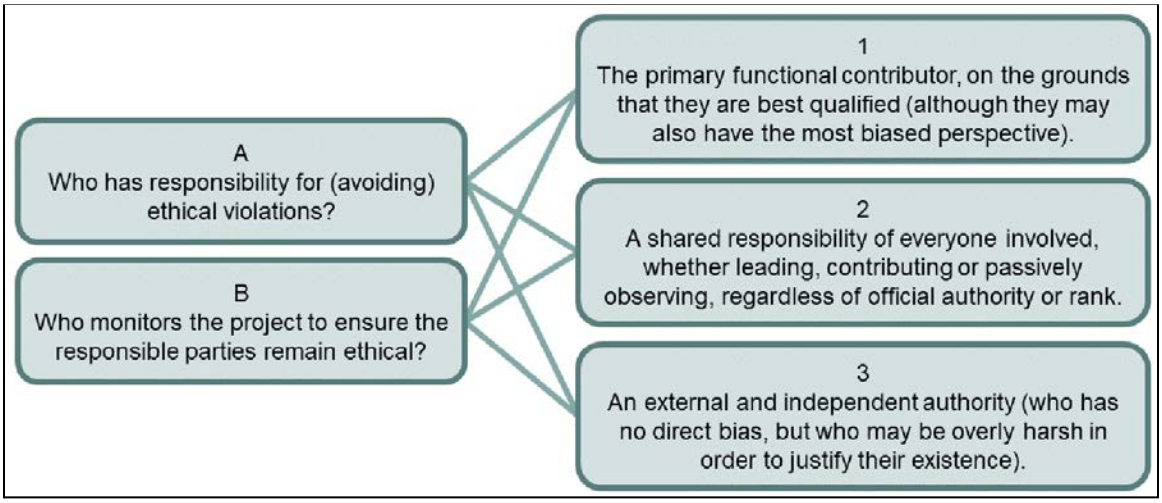


Figure 2: Two key questions and three possible answers when considering roles and responsibilities for improving ethics.

We suggest that a practical answer is A-1 and B-2. That is, the functional lead is the party ultimately responsible for avoiding unethical elements and resolving problems that are identified. However, it is a shared responsibility to make sure those issues are raised, since the functional lead is likely 'too close' to the material to see the problems. So, while responsibility ultimately lies with the leader, the supporting roles should expect to bear a disproportionate burden of the monitoring task.

That could create a cultural clash – the leader is used to being the boss and having the final call on most issues, and yet they are being asked to be subservient to the supporting roles when it comes to ethical considerations. An ideal answer might be A-1-2 and B-2-3. That is, everybody involved in a phase is responsible for addressing ethical concerns, and an external unbiased authority is responsible for monitoring problems. However, while that sounds great in theory, in practice it is likely to be too high of a burden. Perhaps it should become the convention to allocate resources to a more serious solution like the described ideal, but in the present culture and conventions, it is probably a bridge too far.

An external authority could be a colleague not involved in the project, an FFRDC (Federally Funded Research and Development Center), or a wargaming group from a different service in a reciprocity agreement. To make a structure like this work in practice, it would need to provide other benefits to justify the resources expended. A reciprocal agreement might provide both ethical oversight and skill/tool sharing that would help justify the arrangement on paper. While such agreements may be practical, safeguards and controls must be put in place to reduce possible negative effects such as unspoken quid-pro-quos. It is also important that the outside authority not just be involved, but that it be explicitly given the role of ethical oversight (and clear guidelines, such as this taxonomy). Just involving an outside authority does not check the ethical box.

However, it is important to not over-rely on organizational structure and authorities. Ethics, like safety, are best improved by building an appropriate culture [39, 40, 41, 42, 43]. Best practice is to cultivate a culture of shared responsibility, non-judgmental criticism, and low-consequence failure. Contributors who are worried about the long term impacts of mistakes will hide their work to obscure mistakes, rather than share the work transparently to improve it. Culture trumps oversight, but both have a role in successful organizations. So, in practice, A-2 and B-2 is likely to be much more effective than A-1 and B-3, but A-1-2 and B-1-2-3 is likely to be the best of all if resources permit.

3 Unethical Dynamics in Wargames

3.1 Taxonomy Categories

We will consider several broad categories of unethical behavior, each of which can manifest across multiple roles and phases. This list is not exhaustive, but it hopefully covers the common and high risk cases that can arise.

As shown in Figure 3, we organized the unethical dynamics into three broad categories—motivation, execution, and indirect effects—keeping in mind that many of the dynamics bridge those larger categories. For example, desensitization can also pollute the data you collect, independent of any lasting effect it may have on the participant. Hasty conclusions can change the participants' views of the world. The illusion of choice is not just at risk of being unduly influential on participant opinion but is also a form of Structural Bias that corrupts the value of the game. Inappropriate Asymmetry can also produce a Backdrop Messaging effect, sending an unintended implicit message about the nature of a potential adversary.

It is no accident that this breakdown bears a strong resemblance to Walzer's Just War Theory [7]. Walzer's theory is often criticized for only addressing just reasons and just conduct,

not post-war effect and just peace. Our taxonomy attempts to account for parallels of all three, that is, the motivation, the conduct, and the after effects.

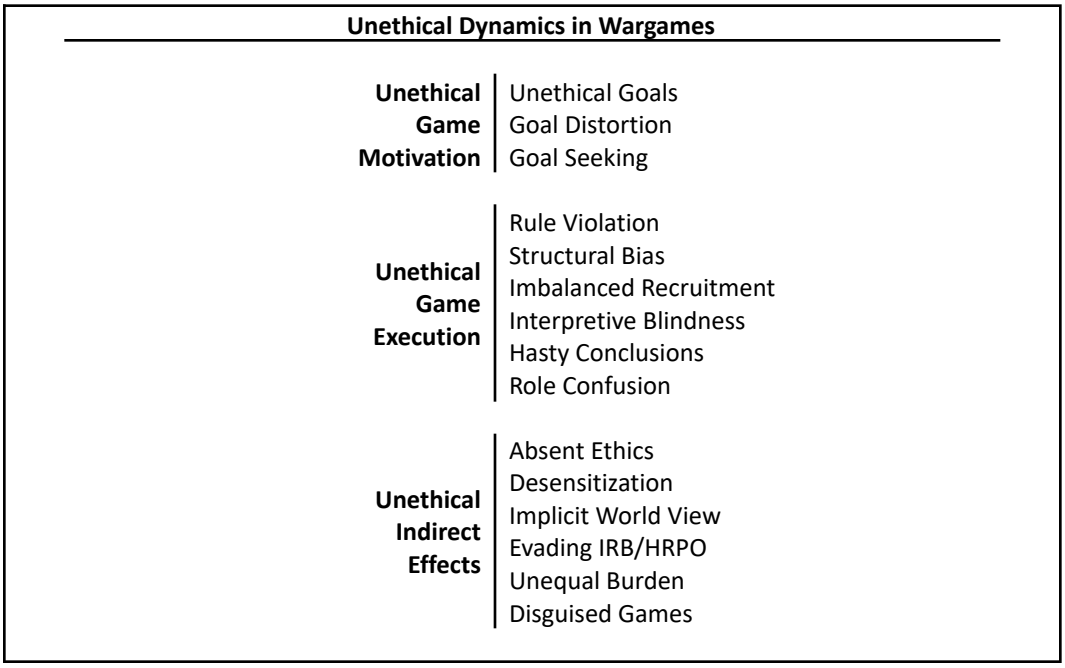


Figure 3: Categories of unethical wargame dynamics discussed in this document.

This section breaks down each category with a definition, a table of risk assessment for each role-phase combination, and a discussion of gray areas. Risk is marked in the role-phase table based on who may conduct or cause the consequence of the unethical behavior. For example, in Figure 3, there is a "low risk" value for the sponsor - pre-creation cell. In this case, the sponsor is susceptible to unethical behavior in the pre-creation phase. Higher risks have high consequences and occur frequently while lower risks have only one of those properties. Risk is assessed as a combination of likelihood (frequency) and severity (consequence), but these ratings are highly subjective; risk may be higher or lower than our estimates depending on context, experience, and organizational culture. We provide these ratings to help convey intuition and our experience, not as a definitive or data-driven judgment. *Note, "unlikely" is used to indicate cells where the risk is considered negligible due to the rarity of the event, even though the consequence may be high.*

Appendix A provides examples for each category, giving both high and low risk cases. Those are generic examples created to illustrate how an ethical violation might occur and how we have observed them to occur. They are not, to our knowledge, reflective of particular real-world events or meant to target or single out particular organizations.

3.2 Unethical Goals

Unethical Goals are motivations for creating a game that are themselves fundamentally inappropriate, regardless of the execution and findings of the game itself. This category is about selecting the goal with unethical intent, prior to the design and execution of the game itself.

This category includes

- creating a game with the intention of obfuscating the truth rather than uncovering the truth.

It does not include a game

- created to study an ethical gray area that is then subverted to produce a pre-conceived result (those cases are covered under *Structural Bias*, *Unequal Recruitment*, and other violations in the *Unethical Game Execution* category), or
- that intentionally explores ethical gray areas with the intent to improve ethics.

For example, commissioning a game to study the most efficient methods of ethnic cleansing (player goal) for the purpose of conducting ethnic cleansing (sponsor goal) is an unethical goal, regardless of the execution and analysis of the game.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	low risk	unlikely	unlikely	unlikely	unlikely	unlikely	high risk
designer / developer	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely
moderator / adjudicator / white cell		unlikely	unlikely	unlikely	unlikely		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	unlikely	unlikely
content consumer					unlikely	unlikely	unlikely
observer			unlikely	unlikely	unlikely		unlikely

Figure 4: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Unethical Goals*. *Unethical Goals* are motivations for creating a game that are themselves fundamentally inappropriate, regardless of the execution and findings of the game itself.

Gray Areas and Abuses. A game can confront ethical gray areas without being itself unethical. Evaluating whether certain policies for urban warfare are ethical might involve simulating actions and policies that are themselves unethical. However, if the game is designed to assess those ethics and produce a more ethical set of policies and doctrine, then the game is ethical despite the unsavory actions modeled.

For example, to contrast to the clearly unethical ethnic cleansing example, a game commissioned to study the most efficient methods of ethnic cleansing (player goal) for the purpose of finding the best way to prevent ethnic cleansing (sponsor goal) *may* be considered an ethical goal, however special controls should be put in place to prevent negative effects on players (e.g. desensitization) and possible exploitation of results by others for unethical purposes (e.g. "falls into the wrong hands"). This example highlights the gray areas.

The boundary can be fuzzy. However, the presence of fuzzy boundaries does not mean that there are no cases that clearly cross that boundary, and it does not mean that the issue should be ignored. It merely means that judgment (possibly independent judgment) is required to properly probe the gray areas. In contrast, creating the game with the intent of casting doubt onto criticisms of certain urban warfare doctrine, rather than to fairly assess them, would cross the gray area into the camp of being unethical by design.

Certainly the fuzzy nature of ethical issues can become an excuse to simply not study ethics. The fear of *Unethical Goals* might lead to *Absent Ethics*. As with many of these categories, concern over one category might lead well-intended people to ignore another category or provide ill-intended people an excuse to intentionally ignore another category. This taxonomy can only help be systematic, not remove the need for judgment, culture, and oversight (as discussed in Section 3.2).

3.3 Goal Distortion

Goal Distortion is when someone intentionally misinterprets the stated goal of the wargame, pushing the game in a direction other than what was intended and stated.

This category includes

- accidental or intentional distortion of the original goals for the wargame.

It does not include

- goals that should have been distorted due to being unethical (those situations are covered by *Unethical Goals*), or
- cases where goals change and evolve with the consent and knowledge of the sponsor, content consumer, or other stakeholders.

For example, in a game designed to find the best way to operate within existing legal constraints, players reinterpreting the rules to find alternative legal approaches that favor their agency and use the game to justify those legal changes is goal distortion.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	low risk	unlikely	unlikely	unlikely	unlikely	unlikely	low risk
designer / developer	unlikely	high risk	unlikely	unlikely	unlikely	unlikely	low risk
moderator / adjudicator / white cell		unlikely	low risk	unlikely	low risk		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	low risk	unlikely
content consumer					unlikely	unlikely	unlikely
observer			unlikely	unlikely	unlikely		unlikely

Figure 5: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Goal Distortion*. *Goal Distortion* is when someone intentionally misinterprets the stated goal of the wargame, pushing the game in a direction other than what was intended and stated.

Gray Areas and Abuses. The risk of distortion could accidentally or intentionally justify failing to evolve a goal that needs to change or keeping an unethical goal in place. Fear of *Goal Distortion* is not justification to continue with *Unethical Goals*. Nor should it be a reason to not update a game whose goals have been overcome by events. New world events, discoveries during the design and creation stage, or changes to doctrine might all legitimately call for the goals of a game to change. Such changes only run afoul of ethics when they are done without involving the parties that have commissioned the game in the first place.

3.4 Goal Seeking

Goal Seeking is a case where the existence of the game produces an inappropriate bias on the resulting analysis, regardless of the game design and execution.

This category includes

- setting the scenario or adversary to inherently bias it towards one outcome

It does not include

- seeking an outcome that is unethical (those cases are covered under *Unethical Goals*), or
- designing the game itself to force a particular outcome (those cases are covered under *Structural Bias*).

For example, creating a game about jungle warfare in the hopes of generating evidence against armored vehicles would be goal seeking.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	high risk	unlikely	unlikely	unlikely	unlikely	unlikely	high risk
designer / developer	low risk	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely
moderator / adjudicator / white cell		unlikely	unlikely	low risk	unlikely		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	low risk	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	high risk	unlikely
content consumer					unlikely	unlikely	unlikely
observer			unlikely	unlikely	unlikely		unlikely

Figure 6: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Goal Seeking*. *Goal Seeking* is a case where the existence of the game produces an inappropriate bias on the resulting analysis, regardless of the game design and execution.

Gray Areas and Abuses. This category is closely related to *Unethical Goals*, except that the outcome being 'forced' is ethically neutral. It is also closely related to *Structural Bias*, except that the bias is injected even before the structure of the game is created. We call it out as a separate category to emphasize that ethics can be violated even when the topic is ethically neutral and that such violations can start extremely early in the lifecycle of the game [8, 9, 10].

3.5 Rule Violation

A *Rule Violation* is when a party intentionally violates the structured rules of the game. For a more exhaustive discussion and categorization of rule breaking, we suggest [56].

This category includes

- all forms of traditional cheating by the players, either by players or moderators [56], and
- unilaterally breaking the rule because you think it is more realistic or valuable to do so, when the game format does not explicitly allow fluid rules.

It does not include

- making a judgment call about a case not covered by the rules or fixing rules that turn out to be broken,
- players suggesting to the moderator that they modify the rules to allow something clever, or non-rigid or matrix-style [11] games where rules are intended to be fluid, or
- players 'gaming the system' by exploiting loopholes in written rules, or players failing to play 'in the spirit' of unwritten rules.

For example, a player peeking at the opposing team's planning map would be a rule violation (assuming that spying is against the rules of the game).

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely
designer / developer	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely
moderator / adjudicator / white cell		unlikely	unlikely	low risk	unlikely		
red / blue participant			unlikely	high risk	unlikely		
gray / pink participant			unlikely	low risk	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	unlikely	unlikely
content consumer					unlikely	unlikely	unlikely
observer			unlikely	low risk	unlikely		unlikely

Figure 7: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Rule Violation*. A *Rule Violation* is when a party intentionally violates the structured rules of the game.

Gray Areas and Abuses. If the goal is to explore threats by a clever enemy, you want to allow small rule violations in the interest of allowing creativity, surprise, and innovation [12]. Real adversaries break rules, so some game formats explicitly encourage players to suggest rule changes. Suggesting a rule change is different than unilaterally breaking the rules. Judgment is required, and it is not a simple line in the sand.

Players worried about breaking rules in an unfamiliar system might be unduly limited in their creativity and effectiveness, so simply punishing rule violation is likely to do more harm than good. Well-intended players learning a complex game are likely to break rules. If that becomes an issue, a moderator interpreting player intent into in-game actions might be a better route than requiring the players to internalize the rules; ethics consider intent as well as consequence.

Rule violations may be especially hard to identify in loosely adjudicated seminar-style games, where player incentives are not to correct the rationalizations of their teammates [13]. Similarly, loosely or under-defined rules may produce ambiguity on what is cheating and what is "initiative." For example, if a player spys on the opposing team and spying is not outlawed in the rules, is the player cheating?

A rule violation may be malicious yet still provide insight [14, 15]. When a rules violation happens, there is a judgment call to be made based on programmatic goals—to correct it, to let it play out, or to give the opposing force an opportunity to make a comparable bend in the rules [16, 17].

3.6 Structural Bias

Structural Bias is when the rules, format, or moderation of the game drive players towards certain choices or outcomes independent of underlying models and player inclination. That is, the game is unrealistic due to the structure of the game, not due to a lack of knowledge of how to be realistic.

This category includes

- what actions the game abstraction allows, such as whether or not a given technology, maneuver, or domain is included,
- how the rules of the game are presented, including what is emphasized during rules explanation, how questions are answered during execution, and what terminology is used,
- how game success, scores, and victory are defined or computed, and
- any judgment calls made about game balance.

It does not include

- a game that biases participants against an action because of realistically modeled drawbacks to that action (that is a bias that is really experience),
- cases where moderators break the rules of their own game in order to bias behavior (such a situation would fall under *Rule Violations*), or
- cases where the designers make a mistake in modeling unknowns about future technology, adversaries, or environments. This category is about the game poorly modeling the best understanding, not about when the best understanding is itself flawed.

For example, adding rules to model mobility restrictions of mud but not the mobility restrictions of wave height would structurally bias the game towards sea-based logistics over land-based logistics.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	unlikely	low risk	unlikely	unlikely	unlikely	unlikely	unlikely
designer / developer	unlikely	high risk	low risk	unlikely	unlikely	unlikely	low risk
moderator / adjudicator / white cell		low risk	low risk	high risk	low risk		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	unlikely	unlikely
content consumer					unlikely	unlikely	unlikely
observer			unlikely	unlikely	unlikely		unlikely

Figure 8: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Structural Bias*. *Structural Bias* is when the rules, format, or moderation of the game drive players towards certain choices or outcomes independent of underlying models and player inclination.

Gray Areas and Abuses. All game designers require modeling abstractions, and any modeling abstraction implicitly creates bias in the player. In fact, game rules exist precisely to influence player action. We give a particular bonus to suppressing fire, public affairs announcements, or use of experienced subordinates because we want players to solve a problem in the presence of those biases. The ethical element comes into play when those rules are distorted intentionally or through sloppy practice, leading to the rules inducing bias that is unrelated to the best available estimates of reality. The fear of this shortfall can easily lead to design paralysis.

3.7 Imbalanced Recruitment

Imbalanced Recruitment is when the set of participants and/or adjudicators are selected from a non-representative pool, thereby influencing the game outcome in a manner that does not reflect the population that appears to have been included.

This category includes

- any role that is recruited, including participants and moderators,
- both demographic diversity and organizational diversity,
- accidental imbalance resulting from volunteer self-selection, and
- reuse of a small pool of available participants for multiple iterations of a campaign of related wargames.

It does not include

- intentionally and transparently targeting a particular population.

For example, recruiting heavily from an artillery unit in order to assess the relative value of indirect fire and close air support would be an example of imbalanced recruitment.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	low risk	high risk	unlikely	unlikely	unlikely	unlikely	low risk
designer / developer	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely
moderator / adjudicator / white cell		unlikely	unlikely	unlikely	unlikely		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		low risk
analyst			unlikely	unlikely	unlikely	unlikely	unlikely
content consumer					unlikely	unlikely	unlikely
observer			unlikely	unlikely	unlikely		unlikely

Figure 9: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Imbalanced Recruitment*. *Imbalanced Recruitment* is when the set of participants and adjudicators are selected from a non-representative pool, thereby influencing the game outcome in a manner that does not reflect the population that appears to have been included.

Gray Areas and Abuses. Complete coverage of potential populations is impossible with a finite number of slots to fill, and even within those bounds it is often impractical given constraints on recruitment, such as classification, expertise, and budget. Rather than judging games harshly on impossible standards, it is more appropriate to expect transparent documentation. If known imbalance in population is reported clearly in the analysis, it is not a problem. In fact, it can often be appropriate to intentionally target a subset of the population to elicit input from a group with a particular bias or knowledge base. The pragmatic standard might best be transparency rather than completeness [18].

An excessive focus on variety of participants can produce a lack of depth in any one area. If the wargame is being run as a campaign of multiple iterations, consider including a few iterations with targeted sub populations. Comparing how players with different specialties, backgrounds, or levels of experience approach the situations can provide interesting analytic findings.

3.8 Interpretive Blindness

Interpretive Blindness is when post-game analysis and interpretation misrepresent the events or logical conclusions of what was actually observed during the game.

This category includes

- one-sided interpretation of ambiguous results, exaggeration of the statistical validity of a quantitative result, and cherry picking quotes and anecdotes in a non-representative manner. An extreme version would be someone proclaiming a result that never occurred in the game.

It does not include

- influencing events of the game itself to validate a preconception (those violations fall under categories such as *Rule Violations* and *Structural Bias*), or
- seeing only part of the data, or drawing a valid conclusion from incomplete data; for those risks, see *Hasty Conclusions*.

For example, cherry picking anecdotes of successes and explaining away instances of failures (whether done intentionally or unconsciously) would be interpretive blindness.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely
designer / developer	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely
moderator / adjudicator / white cell		unlikely	unlikely	unlikely	unlikely		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			low risk		low risk
analyst			unlikely	low risk	low risk	high risk	high risk
content consumer					low risk	low risk	low risk
observer			unlikely	low risk	low risk		low risk

Figure 10: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Interpretive Blindness*. *Interpretive Blindness* is when post-game analysis and interpretation misrepresent the events or logical conclusions of what was actually observed during the game.

Gray Areas and Abuses. Interpretive blindness often emerges from well-intended unconscious cognitive bias [19]. The situation is complicated by the fact that experience is a (positive) case of unconscious cognitive bias, which is certainly appropriate and ethical to apply to wargame creation and interpretation. As a result, it is hard to self identify the harmful cases and experts are likely to be defensive when they are pointed out. As discussed in Section 3.2, a positive organizational culture is important rather than (or in addition to) an externalized punitive incentive.

Concern about unconscious bias can also paralyze experts, limiting the application of their valuable experience. A gut feeling that a result doesn't map to reality is valuable in analysis, even though it also risks losing or distorting the analysis. The obvious bias of experts and low statistical validity can easily be used for ill; they are perfect excuses to ignore an uncomfortable or unexpected finding. Well-intended concern over interpretive blindness can potentially do more harm than good.

Part of the solution lies in the common observation that wargames are often about identifying credible outcomes that had not been considered, not creating high confidence forecasts or predictions. Even if a result is not what experts expect, it can be reported as a finding with an explanation of why the experts are skeptical—see Section A.8.3 for a similar discussion on how to document unusual findings.

Another implication of the emphasis on finding new possibilities (versus forecasting likely outcomes) is to correctly assess the statistical burden of reporting a finding [20]. A non-significant statistical finding can be reported as a discovered behavior; a collection of anecdotes can be an insightful and productive outcome. However, this puts the analyst and content consumer in a tricky situation—they have to be careful to not overstate the predictive value of a game while also understanding the value of reporting non-predictive exploration. The

report has to be crystal clear about what results are interesting anecdotes versus solid statistical results or forecasts.

3.9 Hasty Conclusions

A *Hasty Conclusion* is when a participant draws conclusions and/or publicizes findings based on an incomplete view of the actual events that occurred.

This category includes

- a person only seeing part of the game, part of the analysis, or a subset of participant discussion and assuming that subset is or represents the entire game, analysis, or experience, and
- only considering preliminary findings without bothering to read the completed report's findings.

This category does not include

- a person seeing all of the data and only repeating, remembering, or cherry-picking a subset (for cherry-picking, see the *Interpretive Blindness* category).

For example, an observer assuming that the one cell they observed represented the overall experience of all players would be a hasty conclusion.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	unlikely	unlikely	unlikely	unlikely	low risk	unlikely	high risk
designer / developer	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely
moderator / adjudicator / white cell		unlikely	unlikely	unlikely	unlikely		
red / blue participant			unlikely	unlikely	low risk		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		low risk
analyst			unlikely	unlikely	low risk	low risk	low risk
content consumer					high risk	low risk	low risk
observer			unlikely	high risk	high risk		unlikely

Figure 11: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Hasty Conclusions*. A *Hasty Conclusion* is when a participant draws conclusions and/or publicizes findings based on an incomplete view of the actual events that occurred.

Gray Areas and Abuses. Allowing observers can be a controversial decision in wargame design. Some designers and facilities ban passive observers entirely, on the grounds that they have a narrow view of the game but high influence over perceptions of the outcomes. Under that approach, if you want to see the game, then you have to sign up as a participant and get the full briefing and all the context. However, there are often social, political, and knowledge-sharing benefits to allowing observers without such restrictions. Banning observers may undermine the impact of the game and limit knowledge sharing between wargame designers, so the risks must be balanced against the benefits. Better, perhaps, is to recognize that being a good observer requires some skill and training, much like being an historian or

anthropologist (see Section A.8.3). Observers (and participants) might be put through a short training session to make them aware of this risk.

The fear of drawing hasty conclusions can be a barrier to drawing any conclusion. An unethical practitioner can use fear of drawing hasty conclusions to avoid drawing a conclusion until a decision must be made, at which time they advocate for their desired outcome without being supported by the analysis. Similarly, it can be an excuse to avoid oversight of other potentially unethical practices by removing peer review. The ability to gate-keep access to the event can be abused as a means of preventing counter-interpretations or criticism (or even as a means of limiting who knows about an undesirable conclusion). Wargames are often used to discover possible (and surprising) outcomes rather than to predict the most likely outcome; a hasty conclusion that is clearly presented as an anecdote is nothing to be feared, as long as it is not presented as a high confidence general result [21, 22].

3.10 Role Confusion

Role Confusion is when a single person serving multiple roles takes action that is appropriate under one role but inappropriate under another. This category can include cases where one person serves multiple roles but never at the same time.

For example, a designer also operating as a blue team player would be a concerning case of role confusion.

	sponsor	designer	moderator	red/blue	gray/pink	prior p.	analyst	consumer	observer
sponsor		unlikely	unlikely	low risk	unlikely	unlikely	unlikely	unlikely	unlikely
designer			low risk	high risk	low risk	low risk	low risk	low risk	unlikely
moderator				high risk	low risk	low risk	low risk	low risk	high risk
red / blue				high risk	high risk	high risk	high risk	low risk	high risk
gray / pink						low risk	low risk	low risk	high risk
prior p.							low risk	unlikely	unlikely
analyst								unlikely	unlikely
consumer									unlikely
observer									

Figure 12: Role-Role pairings that are estimated to pose the greatest risks of unethical dynamics due to *Role Confusion*. *Role Confusion* is when a single person serving multiple roles takes action that is appropriate under one role but inappropriate under another.

Gray Areas and Abuses. It is common for one person to play several roles, and doing so is often necessary and reasonable. Some combinations are actually favorable—it is beneficial for the sponsor to also be an invested content consumer, or an analyst to also qualitatively observe the events they are analyzing. Many combinations expose the game to unethical behaviors, either requiring additional vigilance or outright avoidance.

3.11 Absent Ethics

The *Absent Ethics* risk occurs when relevant ethical issues are omitted from the decision space and actions space presented in a game.

This category includes

- failure to account for relevant ethical issues in adjudication, even if the factors are presented.

This category does not include

- omitting factors that can be analyzed separately or which represent only a minor factor on relevant decisions,
- games in which a player is not told about ethical dynamics but they are modeled, analyzed, and adjudicated (cases of altering player perception through in-game fog are covered in *Implicit World View*), or
- intentionally ignoring a game's goal of studying ethical issues (that case would fall under *Goal Distortion*).

For example, a game that studies urban warfare without a model of civilian impacts would be a case of absent ethics.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	low risk	unlikely	unlikely	unlikely	unlikely	unlikely	low risk
designer / developer	unlikely	high risk	unlikely	unlikely	unlikely	unlikely	low risk
moderator / adjudicator / white cell		unlikely	low risk	low risk	unlikely		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	high risk	unlikely
content consumer					unlikely	low risk	unlikely
observer			unlikely	unlikely	unlikely		unlikely

Figure 13: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Absent Ethics*. The *Absent Ethics* risk occurs when relevant ethical issues are omitted from the decision space and actions space presented in a game.

Gray Areas and Abuses. Failing to model ethical elements could be out of negligence, over-caution from the fear of getting it wrong, blatant disregard of the issues, or fear of drawing great criticism for only partially modeling the subject [23, 24]. The intent is separate from the effect, and the effect of ignoring relevant ethical dilemmas is to distort outcomes with real-world consequences.

That said, many games do not warrant ethical dilemmas. Even games about topics involving ethical dilemmas might appropriately omit those dilemmas as separate from the target analysis. An analysis of whether joint fires in an urban EW-setting are viable with current C2 structures can quite reasonably ignore questions of whether civilian casualties are being appropriately considered. When assessing this risk, consider what questions a wargame *doesn't* ask, not just what questions it *does* ask.

3.12 Desensitization

Desensitization is when a participant stops considering ethical factors due to overexposure in game-environments, or when unethical behavior becomes normalized through repetition in fictional environments.

This category includes leaving participants with the

- sense that the issue is fuzzy, contextual, and relative and thus not worth worrying about, and
- the ability to rationalize away any decision, right or wrong.

This category does not include

- temporary burnout from thinking about the topic all day during a wargame (if that desensitization lasts only a few days), or
- rational assessments that might shift someone's view of the line between justified and unjustified acts.

For example, a game that forces a player to repeatedly sacrifice civilians to achieve military goals might be a case of desensitization.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	unlikely	low risk	unlikely	unlikely	unlikely	unlikely	low risk
designer / developer	unlikely	low risk	unlikely	unlikely	unlikely	unlikely	high risk
moderator / adjudicator / white cell		unlikely	unlikely	unlikely	unlikely		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	unlikely	unlikely
content consumer					unlikely	unlikely	unlikely
observer			unlikely	unlikely	unlikely		low risk

Figure 14: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Desensitization*. *Desensitization* is when a participant stops considering ethical factors due to overexposure in game-environments, or when unethical behavior becomes normalized through repetition in fictional environments.

Gray Areas and Abuses. Wargames can be a valuable way to provide 'reps and sets' for decision making [25, 26]. That very feature that makes them effective for building skills also creates a risk for numbing participants to certain actions [27].

However, the concern of desensitization could easily be used to justify either omitting ethical issues or not running a campaign of multiple wargames. Both of those consequences would likely be worse than the risk of desensitization.

3.13 Implicit World View

An *Implicit World View* risk is when irrelevant details of the game convey inaccurate or inappropriate content.

This category includes

- creating a negative view of an entire culture in order to justify the in-game rules governing the armed forces of a nation representing that culture, and
- exaggerating the lack of ethical bounds of an opposing force.

It does not include

- cases where a world view is intentionally conveyed to set the stage for the scenario,
- modeling actual doctrine and historical precedent for what ethical lines a particular adversary might be willing to cross, or
- accurately estimating looser ethical restrictions on an opposing force.

For example, beginning every game with a road to war backdrop that presents middle east adversaries as unprovoked aggressors would be an implicit world view.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely
designer / developer	unlikely	low risk	unlikely	unlikely	unlikely	unlikely	high risk
moderator / adjudicator / white cell		unlikely	high risk	unlikely	low risk		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	unlikely	unlikely
content consumer					unlikely	low risk	unlikely
observer			unlikely	unlikely	unlikely		unlikely

Figure 15: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Implicit World View*. An *Implicit World View* risk is when irrelevant details of the game convey inaccurate or inappropriate content.

Gray Areas and Abuses. The mechanisms that produce violations of *Implicit World View* are very similar to those that produce *Structural Bias*, but the impacts are different. A game that is rigged to drive players to a particular behavior has multiple bad effects—it is bad science (distorting the validity of the resulting analysis, as discussed under *Structural Bias*) and can have a lasting impact on participants (discussed here, under *Implicit World View*).

When a player is given an apparent choice that actually is a forced-choice, they can rationalize that the choice reflects their own beliefs. A game in which players are forced or encouraged to act in a particular way can make players believe they already believed that was the right way, to retroactively justify their in-game actions. The rationalization of illusion of choice can manifest in social issues of entertainment games [28] or ethical issues of professional wargames.

Of course, it is quite literally impossible to create a wargame without conveying a world view. Even very abstract games (such as Hnefatafl [29] or Guerrilla Checkers [30]) imply something about the culture and limits of both sides. It is unrealistic and undesirable to attempt to avoid conveying any worldview in a game. Rather, the designer should aim to (a) be open about the views being conveyed and (b) reflective of whether those views are appropriate.

Assumptions are necessary, but exposed assumptions are much better than implicit, unspoken, or hidden assumptions [31].

3.14 Evading IRB/HRPO

Evading IRB/HRPO (Institutional Review Board / Human Research Protections Office) is when an activity with human participants is classified as 'not a human experiment' in order to avoid providing protections for the participants [32, 33].

This category includes

- getting a game classified as training instead of as an experiment in order to avoid bureaucratic barriers,
- submitting to IRB/HRPO and being declared exempt, and therefore deciding to not perform due diligence to protect participants (ethical obligations can extend beyond organizational approvals),
- fear of judgment—any perception of out-of-game consequences, whether those concerns are real or imagined,
- creating or reinforcing population stigma. (Will running this game with only Marines make the Marines look unethical as a group, thereby causing social backlash against non participating Marines?), and
- psychological stress that participants may experience (for example, if forced to face in-game ethical dilemmas), if those stresses are not already part of their normal professional duties.

It does not include

- psychological stresses that are consistent with what participants already face in their day-to-day jobs.

For example, downplaying the data collection aspect of a game in order to have a game not classified as human experimentation to avoid oversight and paperwork would be a case of evading IRB/HRPO.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	high risk	unlikely	unlikely	unlikely	unlikely	unlikely	low risk
designer / developer	low risk	unlikely	unlikely	unlikely	unlikely	unlikely	low risk
moderator / adjudicator / white cell		unlikely	low risk	unlikely	unlikely		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	low risk	unlikely
content consumer					unlikely	unlikely	unlikely
observer			unlikely	unlikely	unlikely		unlikely

Figure 16: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Evading IRB/HRPO*. *Evading IRB/HRPO* is when an activity with human participants is classified as 'not a human experiment' in order to avoid providing protections for the participants.

Gray Areas and Abuses. IRB and HRPO approval can put schedules at risk and limit what stressful situations a wargame is permitted to include. Those processes can impose very real limits on the theoretical and practical progress of an analysis. That burden can create an incentive for designers and sponsors to push the bounds of what should or should not fall under those protections. Even if a wargame does not fall under IRB formally, considering some of the IRB protections (e.g., informed consent, data anonymization, and justified burden) may be appropriate, with or without formal oversight.

Citing concern of psychological stress could be used as an excuse to not model ethical dilemmas, thus using one risk to avoid another; see *Absent Ethics*.

3.15 Unequal Burden

Unequal Burden is when some members of the recruitment pool are consistently given a much higher burden than others.

This category includes

- when a population is disproportionately asked to bear the burden of participation, and
- when a population is inappropriately denied the benefits of participation.

It does not include

- the potential skew of results from reusing a small pool of participants (for those issues see *Imbalanced Recruitment*), or
- cases where participation has excessive exogenous effects on all participants (those cases are covered under *Evading IRB/HRPO*).

For example, consistently casting minority officers to play foreign adversaries might create an unequal burden within your recruitment pool.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	low risk	unlikely	unlikely	unlikely	unlikely	unlikely	high risk
designer / developer	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely	unlikely
moderator / adjudicator / white cell		unlikely	unlikely	unlikely	unlikely		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	unlikely	unlikely
content consumer					unlikely	unlikely	unlikely
observer			unlikely	unlikely	unlikely		unlikely

Figure 17: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Unequal Burden*. *Unequal Burden* is when some members of the recruitment pool are consistently given a much higher burden than others.

Gray Areas and Abuses. This category can clash with *Evading IRB/HRPO*. IRB policies typically require that participants of an experiment be members of the population that stands to benefit from the research. That requirement can drive participation to over-use a small population, which might create undue burden [34]. On the flip side, an over-emphasis on

avoiding unequal burden could drive designers to recruit inappropriate or inexperienced participants, running afoul of *Imbalanced Recruitment*.

3.16 Disguised Games

A *Disguised Game* is a non-game that is presented as a game, giving participants an inaccurate understanding of the consequences of their actions.

This category includes

- presenting a game as a non-game or presenting a non-game as a game.

It does not include

- hiding true objective from the players (in some types of games, experiments, or exercises, such as "hidden scenario" games, it can be a best practice to hide the true scenario or exact experimental conditions from players), or.
- treating *human experiments* as games (those issues fall under *Evading IRB/HRPO*).

For example, failing to inform players in a game that their game decisions or actions are actually being executed in a real operation would be a disguised game because the players may be exploring possibilities that they know are not suitable for actional operations.

	pre-creation	design	pre-briefing	execution	hotwash	analysis	repetition
sponsor	unlikely	low risk	unlikely	unlikely	unlikely	unlikely	low risk
designer / developer	unlikely	low risk	low risk	unlikely	unlikely	unlikely	low risk
moderator / adjudicator / white cell		unlikely	high risk	unlikely	unlikely		
red / blue participant			unlikely	unlikely	unlikely		
gray / pink participant			unlikely	unlikely	unlikely		
prior participant	unlikely	unlikely			unlikely		unlikely
analyst			unlikely	unlikely	unlikely	unlikely	unlikely
content consumer					unlikely	unlikely	unlikely
observer			unlikely	unlikely	unlikely		unlikely

Figure 18: Role-phase pairings that are estimated to pose the greatest risks of unethical dynamics due to *Disguised Game*. A *Disguised Game* is a non-game that is presented as a game, giving participants an inaccurate understanding of the consequences of their actions.

Gray Areas and Abuses. Adding game-like elements to real activities can be a motivating factor that can be used ethically or unethically [35, 36, 37, 38]. Sports competitions motivate self improvement that also serve the real goals of preparing for the rigors of combat, but sports competitions within armed forces are not unethical. Fear of creating disguised games can undermine the value of gamification. Disclosing the potential uses and purposes of a game-like structure is a better solution than avoiding those structures entirely. A sports competition in an infantry unit is not disguised as anything other than a reason to develop physical skills also valued by that unit.

4 Discussion and Next Steps

4.1 Intent

Ethical discussions often involve assessing intent [44]. We have not used intent to organize the taxonomy. While conscious premeditation feels morally worse than sloppy habits, both can be very damaging. It should be noted that malevolent actors are likely to cause more damage, be more difficult to rectify, and be more difficult to detect since actors are deliberately trying to avoid detection. We include sloppy habits and other unintentional actions as well as since if undetected, the damage may still be significant. Therefore, practitioners should be vigilant against both, especially since intent may be hard to deduce. Any of the categories discussed in this document could arise through premeditated malice, momentary weakness, unconscious bias, ignorance, or gross negligence. As such, while intent is important to assessing a situation, it is not useful to organizing the list of potential transgressions. For the pragmatic purposes of classifying, identifying, and mitigating bad outcomes, we will ignore intent.

4.2 Legality

We omit ethical violations that explicitly violate the law. This taxonomy just covers problems that are not otherwise enforced or monitored. We have thus omitted topics such as leaking classified data, violating IRB once it has been deemed to apply, or ignoring written procedure from your organization.

4.3 Is this Ethics?

You might ask *"Are these really ethical violations, versus just bad practices or sloppy analysis?"*. In regulated professions like Law, Medicine, and (in some countries) Engineering, practitioners have an obligation both to have good intentions and to follow good practices. If something goes wrong, both intention and due diligence are considered; variations in skill, bad luck, and novel situations are not unethical, but sloppy practice in well understood situations would be classified as such [45, 46].

The taxonomy we present errs on the side of inclusiveness when getting into the gray areas of what is or isn't an ethical question. If cheating by players is deemed unethical, then we need to consider how other involved roles might inject a commensurate distortion (whether or not we would recognize such actions as true 'cheating'). For the sake of thoroughness, it is prudent to include non-ethical malpractice along with actions we would deem unethical in a stricter sense. We've taken the position that ethics includes performing due diligence in a professional

activity, providing structures to socially reinforce good behavior, and preventing bias (both conscious and unconscious) from distorting scientific and technical results.

In the end, the real question is not whether the word *ethics* is suitable but whether the resulting taxonomy has pragmatic impact on improving the quality and reputation of wargaming. We hope the utility is there, regardless of the terminology.

4.4 Pitfalls and Shortcomings

Our approach is not without issue. One pitfall would be if wargaming practitioners become overly sensitive to ethical issues. While this can be mitigated, unaddressed it could lead to avoiding wargames, being too risk averse, being unwilling to explore interesting adjudication options, being unwilling to perform preliminary analysis, or not publishing controversial findings. Practitioners should not stifle the creativity that provides wargaming its source of value. In contrast, we hope that more awareness of unethical practices might spur new creative topics and purposes for wargames.

Another pitfall in this approach is the level of organizational buy-in and support required. This approach requires a communal effort and a culture of quality, which may or may not be enforced by a separated external authority with a fear of repercussions. A culture of using ethical violations to one-up a rival within an organization would undermine a healthy culture and not serve the intended purpose of this taxonomy. This taxonomy should support thoroughness and awareness, not simply be used to place blame or add additional bureaucratic hurdles to running a wargame.

Finally, taxonomies and similar tools can provide a false sense of security, that all unethical issues are identified and covered. Such a mentality can fail in any number of ways, and could increase the rate of ethical violations [47]. From complacency to using the taxonomy to justify ignoring a new uncovered unethical behavior, the true value of these tools is the people using them.

4.5 Next Steps

While these tools were developed after observing myriad wargames, applying these tools and improving them across different wargaming communities is paramount to success and adoption. In addition, these tools should be made available in an open source environment so that all can access and propose changes. This would create a living document with insights into how to apply it for different communities and any new areas of unethical behavior identified (or examples of existing) can be shared with the community broadly.

Appendix A. Taxonomy Examples

The examples we give are hypothetical; they are created to illustrate how an ethical violation might occur, not to imply that those particular violations have occurred. The unethical games described in this section are, to our knowledge, fictional.

A.1 Unethical Goals

Unethical Game Motivation: Unethical Goals
Unethical Goals are motivations for creating a game that are themselves fundamentally inappropriate, regardless of the execution and findings of the game itself.

A.1.1 Low Risk Example

Consider a game that is created to assess the most efficient way to commit genocide with the purpose of eventually executing those actions. In this case the player's goal is to assess whether ovens, gas, machetes, or bullets are the most efficient way to exterminate a target civilian population. The sponsor's goal is to use these findings directly. Clearly, such a game is unethical, no matter how well it is conducted and analyzed.

We rate this example as low risk. While it is horrific to think of running a wargame to justify ethnic cleansing or a holocaust, it is unlikely to transpire in current wargaming environments. An example like this is so obviously unethical that it is unlikely it would get very far. However, we include it as a possibility that has likely transpired historically in various parts of the world.

These distinctions are not always easy to determine. Consider the same game described above but now the sponsor's goal is to use the findings to learn how to best disrupt the adversary carrying out the unethical behavior. In this case the players are acting as the adversary in order to help prioritize how to disrupt these activities. Is this unethical? While this case might not be as clear, we argue this category is valid since there are some very clear cases that can be identified. Our recommendation is to err on the side of caution and fully assess the ethical considerations and mitigations when executing a game with questionable goals.

It is important to note that if someone claims the game is for one purpose and they use it for a different purpose, this is an example of Goal Distortion. This highlights some of the overlap in our taxonomy categories and the difficulties of fully partitioning ideas and concepts.

A.1.2 High Risk Example

Consider a game that is created in response to concerns that Weapon X is unethical to use in an urban setting. The game is created with the goal of casting doubt on the idea that Weapon X is unethical, not with the goal of fairly assessing it.

The game is set against an adversary in a region that has low-density urban areas, knowing that an accurate model of that environment will lessen the collateral cost of employing Weapon X. The game is not run as a campaign with a variety of urban densities, nor is it run in a region in which different parts of the map have high or low population density. Even a perfectly executed wargame in that environment has serious inherent biases.

The game is also set against an adversary known to commit ethical violations. A realistic scenario created for the wargame will thus (accurately) model the consequences of failing to employ Weapon X as higher than the collateral damage from the weapon itself. Such an environment is not selected because it is a priority adversary, but rather with the knowledge that an accurate wargame on the topic is likely to justify the use of Weapon X.

Note: the scenarios and adversaries selected in this example may be cases where Weapon X *is* appropriate, justified, and ethical. However, those elements were not selected because they were a prime concern. They were selected with the intent of undermining criticism of Weapon X in general, knowing that it will help get Weapon X approved in a wider setting. The game is not set up to inform *when* Weapon X should be permitted, but rather it aims simply to get it approved.

A perfect, ethical, and unbiased wargame executed on an unethical premise can have an unethical impact on the world. Just as an ethically-executed war can have unethical roots [7], an ethically run wargame can have unethical roots and ultimately be inappropriate.

We rate the risk as high for this example, since it skirts gray areas enough to be very difficult to detect. It might happen unconsciously or as an emergent consequence of organizational pressures. However, while the individuals involved might be innocent, the consequences on policy (and the world) could be significant.

A.2 Goal Distortion

Unethical Game Motivation: Goal Distortion
Goal Distortion is when someone intentionally misinterprets the stated goal of the wargame, pushing the game in a direction other than what was intended and stated.

A.2.1 High Risk Example

Consider a game that was created with the goal of understanding whether a National Guard force is prepared to provide emergency support to civilian authorities during a natural disaster. Such a game might explore different organizational structures and different ways to allocate authority, in order to assess whether they will be effective and whether they are likely to

accidentally violate legal bounds on how military forces can support domestic operations. Those goals are appropriate and ethical.

The designer decides to run several wargames; in each, different legal bounds are assumed for what initiative the military can take in an emergency. Those bounds are known and clearly written in law, and the original goal was to explore how to operate within those bounds, not to explore how those bounds might change. The game looks very similar to what was intended, but it has been subverted. The outcomes now serve to support legal changes to increase military authority in domestic matters, rather than support the best ways for the military to serve civilian authorities.

We rate this type of example as high risk, since it is hard to detect. Designers must recast, formalize, and evolve vague goals and translate them into in-game objectives, scenario creation, and experimental variables. A designer is already dealing with interpreting and reinterpreting the original goal, and thus it is hard to know when that line has been crossed into misinterpretation.

A.2.2 Low Risk Example

Even if the game is designed consistently with the original goals, goal distortion can occur during execution. Consider the same game as before, only assume the designer did not vary legal restrictions; they stuck to the stated goal and created a game about operating within existing legal bounds.

However, consider a moderator who introduces the game to the players as an exploration of appropriate legal bounds, rather than an exploration of how best to serve existing legal bounds. During the pre-brief, they prime the players to ignore legal bounds and to focus on discovering efficient solutions regardless of legal precedent. During the hot-wash, the moderator frames the discussion around whether the legal bounds were a problem. The game design reflected the original goals, but the game execution has fallen prey to goal distortion.

We rate this example as low risk, since it is easier to detect. The designer and analyst are likely to observe the pre-brief and hotwash, and they probably even wrote the material to be presented in those cases. So, significant deviations that distort the original goals are likely to be caught and course-corrected.

A.3 Goal Seeking

Unethical Game Motivation: Goal Seeking
Goal Seeking is a case where the existence of the game produces an inappropriate bias on the resulting analysis, regardless of the game design and execution.

A.3.1 High Risk Example

Consider a game designed to assess whether cutting-edge Weapon X is worth employing instead of a higher volume of the much cheaper Weapon Y. However, rather than commission a series of games to explore different settings, a single scenario is selected with a region, adversary, and tactical scope that are known to favor Weapon Y. The game goals are stated as assessing non-peer adversaries, since Weapon X is known to be thwarted by advanced jamming. The game is set in a rural environment, since Weapon X is difficult to deploy without significant collateral damage. The game is created at a tactical level so that the higher cost of Weapon X is not visible to participants and omitted from the decision space. Even before game design has begun, the outcome has been fixed.

This example is very similar to the high risk example listed for *Unethical Goals*, but without the 'forced' outcome being itself unethical. We rate it high risk for the same reasons—it skirts gray areas enough to be very difficult to detect, and it can easily happen unconsciously by well-intended sponsors. This example could almost be placed under *Structural Bias* or *Goal Distortion* if the scenario decisions were made by the designer during the design phase instead of by the sponsor during the pre-creation phase. However, we want to emphasize that unethical decisions can be made even before a designer gets involved, in whether a wargame is funded on a particular topic and on how the statement of work is written to justify that funding. Sponsors and stakeholders are not passive observers of potential ethical violations [8, 9, 10].

A.4 Rule Violation

Unethical Game Execution: Rule Violation
A *Rule Violation* is when a party intentionally violates the structured rules of the game.

A.4.1 High Risk Example

Consider a case where a red player sneaks over and spies on blue (or white) to see the position of hidden blue forces, even though there are no game mechanics allowing espionage or infiltration. We rate this case as high risk due to both likelihood and severity. Even a brief glance at the truth map can give one side an undue advantage, distorting the effects of the entire exercise. Players, especially Red, are likely to push the limits of the game with the well-intended goal of being a clever active adversary. They may not consider rule violations true cheating if they have experience in games where rules are looser, and indeed those same limit-pushing actions might be encouraged in a different game format. The boundary for how much to allow boundary pushing by red is a fuzzy one, but it is definitely possible to cross it and enter the realm of unethical distortion of the game results. In many venues, it is easy and thus very tempting for players to spy or eavesdrop. Teams might be in different sections of a single room,

walk past each other on the way to the hallway, or participants might chat with friends on the opposing team during a break.

A.4.2 Low Risk Example

Consider a case where blue players are clearly confused by the game rules for camouflage and EW decoys. The moderator reminds them of the rules, giving an example that is quite close to what red is actually attempting. The moderators have (possibly unintentionally) primed blue to recognize red's ploy (and warned them to focus on ploys in the first place), violating the written rules in the game about secrecy.

We rate this case as low risk, since the benefit of correcting rules misunderstanding will often exceed the cost of inadvertently violating those rules by prompting. Also, most moderators will be consciously aware of the risk of conveying information they saw about red's plans when talking to blue. Moderators cheating is just as much cheating as player cheating, but the incentive is weaker.

However, while that particular example may be low risk, we give a high risk rating overall to moderators enabling rule violations during execution. It is easy for small slips or nudges to have significant impact on the game in ways that violate the written rules of the game. Judgment is required to discern when reminding a player of a rule is enabling creativity versus pushing them in a preconceived direction.

A.5 Structural Bias

Unethical Game Execution: Structural Bias
Structural Bias is when the rules, format, or moderation of the game drive players towards certain choices or outcomes independent of underlying models and player inclination.

A.5.1 High Risk Example

During game design, the designer might intentionally or inadvertently make judgments in creating the scoring system that strongly biases players towards one behavior or another, independent from their natural inclinations and likely real-world behavior.

If a player is given 100 success points for completing their mission and loses 1 point per 100 civilian casualties, that is a very different emergent bias than if they lose 50 points per civilian casualty. The judgment of when collateral damage has exceeded the value of the mission objective is not a simple mathematical calculation.

We rate this case as high risk because small rule balance can sway outcomes, and it can be very difficult for designers to separate their experience from their expectation when creating a

set of rules. A common concern with any modeling exercise is whether you are getting out more than you put in, and game rules are a kind of modeling.

A.5.2 High Risk Example

This issue is not just relevant to in-game ethical dilemmas. Any higher level abstraction will require a balancing judgment call.

Consider a game about tactical ground maneuvers. If a player gains a 1.5x attack multiplier for flanking there will be less incentive for maneuver-based tactics than if they get a 3x attack multiplier. If your game is assessing the utility of light versus heavy attack vehicles, your results might be determined by that one abstraction decision, rather than by player creativity or estimates of enemy force structure. If a single design decision is that influential on the outcomes of the game, there are a few options:

- Research the value carefully to justify it. Flanking is a very old military tactic, so there may be a strong justification for how to calibrate it. However, flanking in the presence of modern and future technology might not be well understood, and the designer may still be in the position of making an influential judgment call.
- Model additional influential factors. If the power of flanking is necessarily an influential judgment call, add additional influential factors such as fatigue, terrain, organizational agility, or morale.
- Vary the parameter. If a single influential factor is dominant and requires a judgment call, it can be varied between runs of the game. If the game is short enough or resources are plentiful enough to run a campaign of games, the balance factor can be varied to observe how much it influences player behavior. Repeated games are not always feasible: such repetition may be too costly, the number of repetitions required grows exponentially with the number of unknown factors, and repetitions might be better spent on other experimental variables. However, when viable, it is a powerful tool to consider the game as an experiment and to systematically vary key parameters.
- Use a different methodology. If one factor is so influential that it overpowers player-to-player variation, and if high repetitions are required to study the parameter, a wargame might be the wrong methodology. A modeling or simulation effort might be appropriate to sort out that factor either instead of a wargame or to lay the groundwork for a subsequent wargame.

While there are several best-practice mitigations of this situation, we rate it as a high risk because such situations are common, influential, and easy to overlook. Once a designer fills in some placeholder values for game balance and sees the behaviors they expect, they are likely to

stop tweaking those numbers and stop looking for counter-evidence that they might be wrong. Design is an iterative process of revising a system until it behaves realistically, which fundamentally risks becoming the process of iterating the system until it produces the answers you expect.

Whether malicious or accidental, there is a constant concern in modeling and game design that design judgments are predetermining outcomes [48]. Wargames are especially vulnerable because they are often run once or only a few times, making it hard to know if an observed result is a fluke, a trend, or an outcome forced by the mechanics.

A.5.3 Low Risk Example

During rules creation and explanation, the complexity of a rule can influence player behavior independent of the rational incentives induced by that rule. Merely by modeling a dynamic of the world as an in-game mechanic draws attention, and it draws further attention if the rules are long or complex. Complex rules draw the attention of some participants ("they talked about it for 20 minutes so it must be important") and discourage others ("I didn't follow all 20 minutes of the explanation so I'll just avoid that part of the game"). Either risks unnatural distortion of player behavior.

The principle of minimalism is often touted in game design as an artistic and aesthetic ideal or as a practical goal for making a game accessible and marketable. Neither of those goals concern us here, but we still advocate being cautious when deviating from minimalist designs. Avoiding unnecessary mechanics is not just a matter of being clever, elegant, or accessible; it is about being transparent and sending the right message to new players about what to focus on.

Including a game rule that is not necessary may send a message the game does not intend to send and/or bias in-game behaviors. All rules introduce some complexity, so this risk alone is not a reason to omit a rule, but it does demand some reflection about whether the game's complexity matches its intended focus. Complex rules shift player attention, and shifted attention changes player behavior.

We rate this as low risk. The effects are likely subtle and less influential than other factors. Wargames tend to be good at creating an atmosphere in which players are free to ask rule questions or get help translating intent into in-game action. However, one should be aware of how the style and duration of the rules explanation can exert undue influence.

A.6 Imbalanced Recruitment

Unethical Game Execution: Imbalanced Recruitment
Imbalanced Recruitment is when the set of participants and adjudicators are selected from a non-representative pool, thereby influencing the game outcome in a manner that does not reflect the population that appears to have been included.

A.6.1 Low Risk Example

Consider a game that involves modern peer conflict, in which the population of participants and adjudicators are drawn primarily from the pool of artillery experts. Their tactics will likely rely on use of artillery rather than, e.g., air support, information operations, or electronic warfare. The effectiveness of artillery may be overstated, and the role of other domains de-emphasized. Furthermore, unconventional use of artillery may be overlooked, as the experts are more likely to have existing doctrine ingrained in their thinking. Those inclinations in no way invalidate the outcomes of the wargame, but they need to be acknowledged, documented, and taken into consideration when interpreting findings.

We rate this risk as low. A simple demographic survey plus good transparency of the participant list mitigates much of the risk. If you know that your participants are of a particular background, it is easier to put the results in context and present them as just one possible outcome. The risk remains, but it is easily mitigated.

A.6.2 High Risk Example

Consider the same situation, but now the focus is on urban conflict and managing civilian collateral damage. The recruitment of artillery experts is much more concerning, not because of their emphasis on artillery but on the lack of organizational diversity.

Participants from a civilian advocacy group or military public affairs might bring a very different perspective in how to model the impacts on the civilians. For example, the artillery officers might focus on avoiding direct civilian casualties, but fail to assess the longer term impacts of destroying infrastructure for water, power, economic prosperity, or local loyalties. A game that may appear to be just a bit biased towards artillery might actually be missing important dynamics about short- versus long-term collateral impacts.

We rate this risk as high. Recruitment pools are often very limited, and they are often focused on experts of kinetic engagement, either by design or by necessity. Omitting other perspectives can have significant invisible influence on the findings. Furthermore, while many wargame designers acknowledge this risk, mitigation options are often very limited. Even being

transparent about a lopsided participant pool does not fully mitigate the loss of perspective, especially for a methodology that prides itself on revealing a variety of credible outcomes.

A.7 Interpretive Blindness

Unethical Game Execution: Interpretive Blindness
Interpretive Blindness is when post-game analysis and interpretation misrepresent the events or logical conclusions of what was actually observed during the game.

A.7.1 High Risk Example

Consider a game that is run with a mix of players with different experience levels, and the analyst plots how different players made different micro-decisions throughout the game. There is a lot of noise in the data, because contextual factors and chance also affect the decisions made, but they still fit a line to the data plot. Someone viewing the data will see the line and draw the obvious conclusion—experience leads to a certain type of preference. Actually, the data might be all noise and non-reproducible. The trend matched the analyst's intuition and the data plot conveys more confidence than it was intended to convey. Despite good intentions, the damage is done.

We rate this category as high risk. It is a well-understood problem in academic research [49, 50, 51], and we see no reason why wargame practitioners would be any more immune. In fact, with wargames, the situation is likely worse since the data tends to be more sparse and anecdotal. The noisier the data, the easier it is to (accidentally or maliciously) see patterns that are not present [52].

Wargames thrive in noisy environments that are too complex to systematically sweep, serving more as exploration than scientific validation. By their nature, wargames require more judgment in their interpretation and thus are especially vulnerable to interpretative biases [21].

A.7.2 Low Risk Example

Consider an observer who wants to write a balanced article about a wargame assessing a novel CONOPS. The observer has no particular bias and is concerned about unduly emphasizing either side of the debate. They quote an equal number of pro and con quotes from participants, even though actual opinions were overwhelmingly critical. Content consumers see the article and assume that CONOPS is still under discussion, when it should actually be viewed much more critically.

We assess this as low (but non-zero) risk both in terms of its severity and likelihood. Any reasonably responsible observer will reflect both sides *and* which side is dominant. The cost of presenting both sides rather than an authentically lopsided report might extend the life of a bad

idea but probably not by much. The strength of the individual anecdotes on both sides will likely convey the true outcome of the wargame, not just the volume of those comments.

A.8 Hasty Conclusions

Unethical Game Execution: Hasty Conclusions
A *Hasty Conclusion* is when a participant draws conclusions and/or publicizes findings based on an incomplete view of the actual events that occurred.

A.8.1 High Risk Example

Consider a case where several Red teams act as disconnected terror cells in distributed opposition to Blue's efforts. An observer or participant only sees the most successful of those cells, and hears blue discuss their frustrations during the hotwash. They write up and publish an article about how challenging the situation will be for Blue and how the proposed CONOPS is ineffective. However, a more thorough analysis might reveal that most red cells were ineffective and that blue was highly successful despite feeling frustrated. That analysis will come out in the official report, likely unread by anyone who read the quick article. Even if the article is later corrected or retracted, the damage has been done.

We rate this case as high risk due to high likelihood and potentially high severity. Many people (both observers and participants) observe only a narrow slice of the game, and humans tend to settle on their first impressions even when later presented with more nuanced analysis. Reporters trying to publicize the results are rarely trained analysts, and participants are highly biased by their individual experience. While the sponsor might read the full report, a much wider audience has formed an opinion based on partial perspectives.

A.8.2 Low Risk Example

Consider a case where an analyst finds that data is mostly missing from one day of a multi-day game. They analyze the data they do have, and assume that the trends for the missing day were similar. Data collection problems are a reality, as is making the best of incomplete data. However, in this case, suppose the missing data includes a key deviation from the patterns observed on other days, and masking that deviation distorts the analysis.

We rate this as low risk. Despite partial data being a common occurrence, there are usually multiple sources to corroborate a result. Even if the official logs of that day are missing, informal reports from participants can guide the analyst to determine if the missing data was likely to be anomalous. The analyst is also likely to follow best practice and note the missing data, putting the final conclusion in context. This is the type of judgment analysts are trained to make and trained to present responsibly.

A.8.3 An Aside on Herodotus

Let us take a brief aside into ancient Greek history and the history of history itself. It sheds some insight onto what good observation and good reporting look like, striking a pragmatic balance between fact and interpretation.

Herodotus, around 450 BCE, documented Phoenician claims that they circumnavigated Africa circa 600 BCE. They reported that when they turned west to go around the southern tip of Africa, the noontime sun was on their right side. Herodotus says that such a fantastical claim is evidence that they were lying about the entire journey, since traveling westward at noon will always put the sun on your left (given his experience living north of the equator, and the early state of Greek astronomy). We now know that the anecdote told by the Phoenician about the sun is strong evidence that they really *did* circumnavigate Africa, as no other Greek at the time would have understood the effects of being in the southern hemisphere on the noontime sun [53]. Herodotus did us a service by documenting the Phoenician claims despite his skepticism.

Herodotus is often cited as the first true historian because of his dedication to the facts and context, in contrast to prior chroniclers whose job was to glorify their sponsoring king or track accounting details. The case with the Phoenicians is an example of ethical reporting on his part; he reported what *was* believed, not just what *he* believed. He did not omit claims that contradicted his own beliefs, even when he supplied editorial commentary on whether or not he found those claims credible. The combination of the factual reporting plus his own editorial notes together are more valuable than either would be individually.

This anecdote is not an anomaly, and there are many similar cases where we benefit from Herodotus's professionalism. Many find his discussion of theories for why the Nile floods comical, but he is owed great credit for clearly separating facts, general opinion, and personal opinion even when discussing absurd claims. The (well distinguished) combination of those three gives us greater insight into the time period that any one would alone.

Anyone writing up the findings from a wargame should strive to be an historian in Herodotus's tradition; document what happened while feeling free to also supply editorial commentary on whether it sounds credible. Commentary is not itself unethical; it provides valuable context and helps non-expert readers understand and interpret the wargame results. Indeed, it is partly the job of the report writer to interpret the findings, so long as that interpretation does not involve altering, omitting, or obfuscating the facts or opposing opinions from other participating roles.

Practical Advice to Observers and Reporters: Document all observed opinions and findings, regardless of their credibility. Provide your own commentary on credibility, realism, and

interpretation, but keep that commentary clearly differentiated from the documented facts of the event. Do not pick and choose what to omit based on your conclusions, but do not hesitate to draw conclusions.

A.9 Role Confusion

Unethical Game Execution: Role Confusion
Role Confusion is when a single person serving multiple roles takes action that is appropriate under one role but inappropriate under another.

A.9.1 Low Risk Example

A 'Pink cell' that combines roles of White and Red cells is low risk. Both are active roles, and their incentives and restrictions are very similar. Indeed, the goal of having a pink cell is to have red and white incentives aligned, so that red behaviors elicit certain situations of interest rather than simply maximizing red success.

A.9.2 Low Risk Example

There is low risk for a content consumer to also be an observer. Both are passive in most phases, so there is little opportunity for misaligned incentives or restrictions.

A.9.3 High Risk Example

Combining the designer and analyst roles can be quite risky. During pre-briefing and hotwash, the designer would have an opportunity to influence participants in a way that clashes with a passive analyst role. During the analysis phase, the problem reverses—the analyst should be impartially evaluating the findings, but the designer's incentive for the game to have worked as intended can influence the analysis.

A.10 Absent Ethics

Unethical Indirect Effects: Absent Ethics
The *Absent Ethics* risk occurs when relevant ethical issues are omitted from the decision space and actions space presented in a game.

A.10.1 High Risk Examples

Consider a game about urban warfare that does due diligence at modeling civilian deaths from kinetic action. However, the game only models immediate impact on the civilian population. If players evacuate a portion of the city, they can reduce civilian deaths to a justifiable level for the military objectives being achieved. However, no model of civilian

infrastructure is included. Players who destroy a water pumping station, power plant, or evacuated hospital see no in-game penalty for the long-term impacts of those lost facilities. As a result, players are left with a skewed perception of which urban strikes are low-impact and which are high impact.

We rate this risk as high, since it is difficult and complex to model long-term effects. Even when it can be done accurately, it may unduly distract from other objectives of the wargame. As such, there are a lot of pressures to only model immediate forms of collateral damage, increasing the risk of conveying the wrong message to participants.

A.10.2 Low Risk Example

Consider that same game, modified to properly account for long-term impacts on civilian infrastructure. The game leaves the participants with an appropriately balanced understanding of the opposing pressures and dilemmas of urban warfare. Players spend time during the game discussing those issues and their impact on the pragmatics of accomplishing the stated mission. However, the analysis omits the ethical elements of the game, reporting only on the success rates of various tactics at accomplishing the tactical military objectives. Content consumers are left with a skewed perception of the issues involved.

We rate this risk as low, since the analyst report would likely be reviewed by observers, moderators, and designers so such an omission would likely be caught. However, if the report were widely read, the impact of such an omission on doctrine and policy could have significant long-term ethical consequences.

A.10.3 Low Risk Example

Consider a game in which the blue team must deal with a misinformation campaign conducted against a host nation by a peer adversary. Blue is given an open ended option to counter the opposing influence operations with their own information operations. If they violate existing policy and ethical guidelines (such as using the public affairs office to spread misinformation or targeting their own citizens) the adjudicators are instructed to provide in-game penalties. That game appears to be providing ethically sound lessons and exercises. However, the game never models cases where allied messaging can get out of hand. If the players follow the rules, there are never any consequences, creating the impression that the rules are perfect.

Consider, instead, a potential scenario in which the players conduct a completely legal information operation including an outreach campaign with locals. Local social media takes up the cause and spreads misinformation about fictional atrocities committed by your adversary. That campaign was not intended by blue's actions, but it now serves blue's aims. Should blue

discredit the campaign? Allow it to continue without any direct support? Reassert their own carefully crafted message?

We rate this example as low risk because the game is likely still teaching a lot of good ethical lessons. The absence of gray areas limits its value, but it likely still has a net positive value. Overly criticizing partial efforts to address ethics can have the unintended effect of discouraging ethical considerations rather than strengthening them.

A.11 Desensitization

Unethical Indirect Effects: Desensitization
Desensitization is when a participants stops considering ethical factors due to overexposure in game-environments, or when unethical behavior becomes normalized through repetition in fictional environments.

A.11.1 Low Risk Example

Consider a game that models supporting emergency response efforts in an allied nation. Players must allocate limited resources between search-and-rescue, medical attention, evacuation, and infrastructure restoration in an effort to minimize loss of life and maintaining social order. No matter how resources are allocated, civilians will die and confidence in the system of government will be shaken. For example, if players allocate resources to search-and-rescue, a few lives are saved and a lot of confidence is built, but if players allocate resources to evacuation, many lives are saved but confidence is lost.

The hope is that players learn how their actions make tradeoffs in an ethical gray area. However, the risk is that they will simply become comfortable with that gray area. Once a player has spent an entire wargame deciding which populations will live and die, they may lose some of their normal hesitation to take actions that will cost lives. They may become comfortable with sacrificing lives in service of a mission. Rather than improve the quality of decision making in ethical gray areas, the game might have made decision makers callous.

We rate this risk low for a one-off event. Even if a three-day intensive game involves constant ethical dilemmas, we would expect any resulting desensitization to be temporary. We also consider a certain amount of desensitization to be helpful; part of making good decisions under pressure is being unemotional and calculating. So, as long as the lessons implicit in the wargame were appropriate, associated emotional numbness may not cause any harm.

A.11.2 High Risk Example

Consider a campaign of games that address a family of related ethical dilemmas officers might face in a mission. In each case, players are forced into uncomfortable situations without

clear best actions and where written policy and doctrine leave room for individual judgment. The game aims to study and improve how judgments are made in ambiguous situations. Players are encouraged to participate in multiple runs of the game to experience a broad range of scenarios and to study how those participants react differently in different contexts.

In this case, participants are left with the impression that the ethical gray areas are mostly no-win situations where all actions are bad, and they are too complex for their level of authority. They adopt the attitude of leaving the decisions to others, rather than taking responsibility or initiative. Later, in deployment, they are passive observers of ethical gray areas rather than active assessors, finding it easy to assume that any course of action is justifiable and thus not their place to voice concerns.

We rate this risk as high. Showing the subtleties, complexities, and multifaceted nature of ethical situations can produce complacency, even among echelons who bear responsibility for assessing the ethics of an order. While it can be good to force players to confront unpleasant dilemmas, it is also important to show that (in retrospect) their actions matter in material ways. Even if all actions cost lives, and no action is perfect in all situations, there are none-the-less clearly better and clearly worse options within the gray area.

A.12 Implicit World View

Unethical Indirect Effects: Implicit World View
An *Implicit World View* risk is when irrelevant details of the game convey inaccurate or inappropriate content.

A.12.1 High Risk Example

Consider a game modeling asymmetric warfare against a decentralized insurgency. The blue team is given strict limitations on kinetic strikes within city limits. The red team, representing a particular foreign nation, is given no restrictions on urban strikes and no penalty for collateral damage.

That asymmetry creates realistic dynamics during the game, in which superior blue firepower is limited by political and ethical bounds, while the outclassed red team uses the flexibility afforded by their looser bounds to make up for their inability to win direct confrontations.

However, that is likely a gross oversimplification of the situation, conveying that particular insurgency as comic book villains who lack all ethics. It stereotypes the enemy in a manner that can reflect poorly on their entire culture outside of the game (e.g. "all Muslims are terrorists") and undermine the goal of discovering realistic outcomes (e.g. the enemy has no limitations). In fact, an insurgent force likely has strong aversion to certain types of collateral damage (or, at

least, of having such damage attributed to them). A more nuanced model of the types of limitations they face will be both a more valuable analysis and less likely to reinforce out-of-game stigmas of the entire ethnic group.

Even if you are modeling an opponent who really does cross many ethical bounds, there is reason for moderation. Excessively demonizing an immoral enemy just casts doubt on real criticisms of their transgressions. Exaggerated propaganda about German war crimes in WWI lead to a delay in acceptance of legitimate reports of Nazi war crimes in WWII [54].

We rate this risk as high because we have observed similar cases to be common, and because it directly influences a population (warfighters) who might need to amiably interact with members of the cultures being stereotyped.

A.12.2 Low Risk Example

Consider a game studying naval conflicts with a near-peer 'Z'. A campaign of games are run about different strategic situations that might arise against Z's armed forces. The same 'road to war' text is used for all the games, and they only differ in the terrain and objectives. That backdrop material always involves Z invading an allied nation to expand its access to natural resources.

In this case, assume that it really is an accurate characterization that Z would invade a nation to expand territory. However, Z would not necessarily use military invasion to improve access to economic resources. There are also many other reasons Z might invade a nation besides pursuit of raw material. The wargame in this example is not studying the causes of war, only the ability to win the war once started, so such omissions do not directly impact its analytic value.

While that adversary's motives have not been unfairly characterized, they have been over-simplified by repetition. This campaign of games might leave repeat participants with the impression that Z feels no cultural or ethical hesitation to invade a nation for economic gain. In fact, that example was just included as an excuse to justify the wargame, and was not the intended message or even an important part of the analysis. None-the-less, damage has been done; participants are left with an exaggerated (if not wholly inaccurate) view of Z's national character.

We rate this risk as low for two reasons. First, many players (to the dismay of wargame designers) pay little attention to the 'road to war' backdrop. Second, this risk occurs after lots of repetition, not after a single game, and many wargame participants only participate in wargames on occasion. The repetition necessary to imbue a world view is often absent, lowering the potential impact.

A.13 Evading IRB/HRPO

Unethical Indirect Effects: Evading IRB/HRPO
Evading IRB/HRPO (Institutional Review Board / Human Research Protections Office) is when an activity with human participants is classified as 'not a human experiment' in order to avoid providing protections for the participants [32, 33].

A.13.1 High Risk Example

Consider a game that recruits from a population of war college students by having instructors make it a mandated activity. The wargame designers are not going to formally publish the results, so it does not qualify as a human experiment. IRB/HRPO declares the game to not be an experiment and not require their oversight. However, student participants are then judged by their instructors on their in-game performance, even though the game is not attempting to be an accurate predictor of real-world performance. It was a brainstorming activity that was held against participants' careers as if it were a formal evaluation.

Even if no harm or judgment is imposed on participants, fear of judgment can undermine the value of a wargame. Participants who feel they are being assessed will follow current procedures and avoid trying anything clever or new. That is a significant problem if the wargame was run to elicit new ideas, explore unconventional responses, and assess the merits of changing doctrine.

We rate this risk as high because bureaucratic pressures make it tempting, and it is easy to overlook potential impacts on participants. The designer may know that a game is an abstract model of future events meant to brainstorm concepts not assess skill, but the participants may suffer backlash if observers do not understand that subtlety.

A.13.2 Low Risk Example

Consider a game that is about performing kinetic operations in an urban environment. The game emphasizes collateral damage and the risk of civilian casualties. Participants are recruited from civilian positions not used to making life-or-death value judgments about innocent lives. Other participants are military officers comfortable with the violence of war, but unfamiliar with viewing the civilian fallout. After the game, players are shown detailed stats about civilian deaths estimated from their actions, including photos of children in the aftermath of past urban clashes. Both groups of participants feel guilt and distress for having made in-game decisions and being confronted with graphic depictions of their side-effects.

We rate this category as low risk, since most participants in wargames will be well aware of the potential impacts of warfare (especially urban warfare) on civilian populations. In general,

IRB and HRPO restrictions permit stress in an experiment if that stress is consistent with what the participants already face in their normal duties. It is certainly possible to unduly induce stress in the participants, but the population likely to be recruited into a military wargame is likely to already have grappled with such issues.

A.14 Unequal Burden

Unethical Indirect Effects: Unequal Burden
Unequal Burden is when some members of the recruitment pool are consistently given a much higher burden than others.

A.14.1 High Risk Example

Consider a campaign of wargames being conducted to study combined arms operations and policy. In the available recruitment pool, there is only one artillery officer who is repeatedly recruited to participate in the wargames. In contrast, there are several available participants to play the role of infantry and logistics. The one artillery officer is expected to participate in a larger number of wargames, in addition to normal duties.

We rate this case as high risk in the situation where a game is being run many times. Recruitment pools are often limited, and thus it is likely that a single person with a particular background might incur disproportionate burden. Military hierarchies can make it hard to say no to participation requests, so a junior officer with a specific background might get saddled with undue and unequal burden.

A.14.2 Low Risk Example

Consider a pol-mil game studying political factors in escalation and deterrence. A common concern in such games is whether participants who play as foreign powers will adhere to the cultural and historical pressures on those powers. To mitigate that concern, participants who play China in the game are recruited from personnel with Asian backgrounds. In this particular case, the one Asian-American officer in the recruitment pool is pressured to participate due to their ancestry.

We rate this risk as low, since the described situation is a single game, not a campaign—the potential burden is low. Additionally, there is generally an understanding among creators of pol-mil wargames that ethnic ancestry is not the same as cultural awareness—they are less likely to over-burden one population over another, and they are more likely to seek out someone who has studied US-Chinese relations. However, if the wargame creators were less experienced and running a series of related wargames, the burden and likelihood would increase and could rate as high risk.

Of course, frequently recruiting the same participant for the same role across several instances of a game has other concerns. Even if the burden is appropriate and fair, a single participant represents a single perspective, and a wargame is often meant to identify a wide range of perspectives and possible outcomes. See *Imbalanced Recruitment* for related concerns about how recruitment can undermine the validity of post game analysis.

A.15 Disguised Games

Unethical Indirect Effects: Disguised Games
A *Disguised Game* is a non-game that is presented as a game, giving participants an inaccurate understanding of the consequences of their actions.

A.15.1 Low Risk Example

The classic example of this situation is described in the science fiction novel *Ender's Game* [55]. Participants believed they were playing a virtual game-like training exercise, but it was actually an interface to commanding a real force engaged in kinetic combat. Choices apparently made in a simulation had real-world genocidal consequences.

While that example is pure fiction, similar analogues are much more realistic—consider a case where a game is presented as a theoretical exercise but is actually used to plan a real operation. A player might experiment with strategies they would not actually want employed in reality. They might push the bounds on risk and ethics to explore the situation, but would be aghast to know that their explorations were being taken as serious suggestions.

Acknowledgements

We are grateful for the input from the Unethical Wargaming Working Group from Connections 2021 [6], especially Stephen Downes-Martin for organizing the working group and Daniel Tyler Brooks for some excellent references.

References

[1] Christopher A. Weuve; Peter P. Perla; Michael C. Markowitz; Robert Rubel; Stephen Downes-Martin; Michael Martin; Paul V. Vebber. Wargame pathologies. September 2004 https://www.cna.org/CNA_files/PDF/D0010866.A1.pdf

[2] Daniel Brooks. Propaganda games. *Youtube*. Oct 2020. https://www.youtube.com/watch?v=UP4_bMhZ4gA

[3] Ed McGrady. How to Build Unethical Wargames. In *Unethical Professional Wargaming, Connections US 2021 Working Group Report*, 2021

[4] Owen, W. How to succeed in unethical wargaming: A user guide. In *Unethical Professional Wargaming, Connections US 2021 Working Group Report*, 2021

[5] Kacper Kwiatkowski; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*, chapter Civilian Casualties: shifting perspective in *This War of Mine*, pages 691–701. Game Histories. MIT Press, 2016.

[6] Stephen Downes-Martin (ed.) Unethical Professional Wargaming, Connections US 2021 Working Group Report, *PAXsims* 2021

[7] Michael Walzer. Just and unjust wars: a moral argument with historical illustrations. Allen Lane, London, 1978.

[8] Stephen Downes-Martin. Your boss, players, and sponsor: The three witches of war gaming. *Naval War College Review*, Vol. 67(No. 1, Article 5), 2014. https://digital-commons.usnwc.edu/nwc-review/vol67/iss1/5/

[9] Yuna Huh Wong; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*, chapter Irregular Warfare: The *Kobayashi Maru* of the Wargaming World, pages 531–537. Game Histories. MIT Press, 2016.

[10] Hiroyasu Akutsu. The Layers of Ethics Surrounding Wargaming: How and Why Some Wargames Were "Unethical" in Japan. In *Unethical Professional Wargaming, Connections US 2021 Working Group Report*, 2021

[11] John Curry; Tim Price. Matrix Games for Modern Wargaming: Developments in Professional and Educational Wargames, volume 2, 2014. https://lulu.com/shop/product-21760442.html

[12] James Caroland; Greg Conti. Lessons of the Kobayashi Maru: Cheating is Fundamental. *ShmooCon*, 2012. https://www.youtube.com/watch?v=7902m6MTu0E

[13] Stephen Downes-Martin; Korene Phillips; Ken Shogren; Sorrel Stetson; Rosemary Tropeano. How can we avoid risky and dishonesty shifts in seminar wargames? In *Connections: US Wargaming Conference: Game Lab*, National Defense University, July 2018. https://paxsims.wordpress.com/2018/08/13/how-can-we-avoid-risky-and-dishonesty-shifts-in-seminar-wargames/

[14] Micah Zenko. Millennium challenge: The Real Story of a Corrupted Military Exercise and its Legacy. *War on the Rocks*, November 2015. warontherocks.com/2015/11/millennium-challenge-the-real-story-of-a-corrupted-military-exercise-and-its-legacy/

[15] Kyle Mizokami. The U.S. Lost a (Fictional) War With Iran 18 Years Ago. *Popular Mechanics*, January 2020. https://www.popularmechanics.com/military/a30392654/millennium-challenge-qassem-soleimani/

[16] Russell Vane; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*, chapter Hypergaming, pages 241–249. Game Histories. MIT Press, 2016.

[17] Zoltán Harangi-Tóth. Unethical From the Beginning – Wargaming and the Culture of Winning in the Former Warsaw Pact Countries. In *Unethical Professional Wargaming, Connections US 2021 Working Group Report*, 2021

[18] Stephen Downes-Martin. Exploit Group Dynamics to Corrupt a Professional Wargame. In *Unethical Professional Wargaming, Connections US 2021 Working Group Report*, 2021

[19] M.G. Haselton; D. Nettle; P.W. Andrews; David M. Buss (ed.). *The Handbook of Evolutionary Psychology*, chapter The Evolution of Cognitive Bias, pages 724–746 . John Wiley & Sons Inc., 2005.

[20] Robert MacDougall; Lisa Faden; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*. Game Histories. MIT Press, 2016.

[21] Brien J. Miller; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*, chapter The Application of Statistical and Forensics Validation to Simulation Modeling in Wargames, pages 183–200. Game Histories. MIT Press, 2016.

[22] Peter P. Perla; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*, chapter Operations Research, Systems Analysis, and Wargaming: Riding the Cycle of Research, pages 159–182. Game Histories. MIT Press, 2016.

[23] Brian Train; Volko Ruhnke; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*, chapter Chess, Go, and Vietnam: Gaming Modern Insurgency, pages 513–529. Game Histories. MIT Press, 2016.

[24] Ed Beach; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*, chapter A Mighty Fortress Is Our God: When Military Action Meets Religious Strife, pages 539–544. Game Histories. MIT Press, 2016.

[25] Scott Bond. Wargaming at the Academy: Why Invest in NHWL and Building a Culture of Wargaming. National Security Innovation Network: Defense Innovation Network, January 2021. https://cimsec.org/wargaming-at-the-academy-why-invest-in-nhwl-and-building-a-culture-of-wargaming/

[26] Phillip Reiman. The Case For a Legal Exercise Range. *Unpublished notes.*

[27] Soraya Murray; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*, chapter Upending Militarized Masculinity in *Spec Ops: The Line*, pages 319–328. Game Histories. MIT Press, 2016.

[28] Jacob Geller. Does Call of Duty Believe in Anything? 2020 https://watchnebula.com/videos/jacob-geller-does-call-of-duty-believe-in-anything https://www.youtube.com/watch?v=FtCV421T52s

[29] Unknown Author. Hnefatafl. Traditional tabletop board game circa 400. https://boardgamegeek.com/boardgame/2932/hnefatafl

[30] Brian Train. Guerrilla Checkers. 2010. https://boardgamegeek.com/boardgame/71035/guerrilla-checkers

[31] Jim Wallman; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*, chapter Cultural Wargaming: Understanding Cross-Cultural Communications Using Wargames, pages 545–551. Game Histories. MIT Press, 2016.

[32] U.S. Army Medical Research and Development Command. Human Research Protection Office (HRPO). 2020 https://mrdc.amedd.army.mil/index.cfm/collaborate/research_protections/hrpo

[33] Massachusetts Institute of Technology. Committee on the Use of Humans as Experimental Subjects (COUHES). 2021 https://couhes.mit.edu/

[34] Riverside University of California. Unequal demands on women for university service harm careers. August 2014 https://www.sciencedaily.com/releases/2014/08/140822083626.htm

[35] Sohini Podder. Robinhood to scrap confetti animation on app after 'gamification' criticism. March 2021 https://www.nasdaq.com/articles/robinhood-to-scrap-confetti-animation-on-app-after-gamification-criticism-2021-03-31

[36] Julia Ebner. The gamification of terrorism. March 2021 https://www.youtube.com/watch?v=rPC_Zf16Nts

[37] Pew Research Center: Internet & Technology. The Future of Gamification. May 2021 https://www.pewresearch.org/internet/2012/05/18/the-future-of-gamification/

[38] Reed Berkowitz. A Game Designer's Analysis of QAnon. September 2020 https://medium.com/curiouserinstitute/a-game-designers-analysis-of-qanon-580972548be5

[39] Charles Perrow. *Normal Accidents: Living with High-Risk Technologies - Updated Edition*. Princeton University Press, rev - revised edition, September 1999 (orig. 1984).

[40] Diane Vaughan. *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. University of Chicago Press, enlarged edition, January 2016 (orig. 1996).

[41] Nancy G. Leveson. *Engineering a Safer World: Systems Thinking Applied to Safety*. The MIT Press, reprint edition, December 2016 (orig. 2012).

[42] Edward Tenner. *Why Things Bite Back: Technology and the Revenge of Unintended Consequences*. Vintage, 2nd edition edition, September 1997 (orig. 1996).

[43] Stephen Downes-Martin. Wargaming to Deceive the Sponsor: Why and How? In *Connections UK 2016 Conference*, King's College, London, September 2016. https://www.professionalwargaming.co.uk/WargamingToDeceivePaper.pdf

[44] Kieran Setiya. Intention. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 edition, 2018. https://plato.stanford.edu/entries/intention/

[45] Wikipedians. Regulation and licensure in engineering. 2021 https://en.wikipedia.org/wiki/Regulation_and_licensure_in_engineering

[46] William F. Owen. Unethical Wargaming. Let Us Be Incompetent! *Proceedings of CONNECTIONS'21*, 2021.

[47] Carmen Morris. Don't Jump Into Unconscious Bias Training Until You Read This. July 2020 https://www.forbes.com/sites/carmenmorris/2020/07/21/dont-jump-into-unconscious-bias-training-until-you-read-this/?sh=15d132983f41

[48] Rachel Simmons; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *Zones of Control: Perspectives on Wargaming*, chapter Goal-Driven Design and *Napoleon's Triumph*, pages 201–207. Game Histories. MIT Press, 2016.

[49] Fallacy Man. Most scientific studies are wrong, but that doesn't mean what you think it means. May 2016 https://thelogicofscience.com/2016/05/10/most-scientific-studies-are-wrong-but-that-doesnt-mean what-you-think-it-means/

[50] Chris Chambers. The Seven Deadly Sins of Psychology: A Manifesto for Reforming the Culture of Scientific Practice. Princeton University Press, April 2017.

[51] Michael Shermer. When scientists sin: Fraud, deception and lies in research reveal how science is (mostly) self-correcting. *Scientific American*, July 2010. https://www.scientificamerican.com/article/when-scientists-sin/

[52] Wikipedians. Pareidolia. 2021 https://en.wikipedia.org/wiki/Pareidolia

[53] Garrett G. Fagan. Great battles of the ancient world. Lecture Series in The Great Courses by The Teaching Company, 2005. https://www.thegreatcourses.com/courses/great-battles-of-the-ancient-world

[54] Vejas Gabriel Liulevicius. World War I: The "great" war. Lecture Series in The Great Courses by The Teaching Company, 2006.

[55] Orson Scott Card. *Ender's Game*. Tom Doherty Associates, 1991.

[56] Katie Salen and Eric Zimmerman. *Rules of Play: Game Design Fundamentals*. Cambridge, Mass: MIT Press, 2003. Print.

About the Authors

Robert Seater is a member of the technical staff at MIT Lincoln Laboratory. He uses interactive models, serious games, and human-machine teaming as data-driven analysis tools to study expert decision making, training & education, mission-level impacts of emerging technologies, AI integration into workflows, and system-level requirements analysis. He has worked across a wide range of application domains, including emergency response management, chem-bio defense, drug smuggling interdiction, military logistics, distributed command and control, urban warfare, countering human trafficking, hypersonic missile defense, and naval missile defense. He received a BS from Haverford College in mathematics and computer science in 2002 and a PhD from MIT in computer science and requirements engineering in 2009. He has been at MIT Lincoln Laboratory since 2009.

Joel Kurucar is an Associate Staff Member at MIT Lincoln Laboratory in the AI Software Architectures and Algorithms group. He has been at the lab for six years and was previously in the Humanitarian Assistance and Disaster Relief Systems group and the Informatics and Decision Support group. Before joining MIT Lincoln Laboratory, Joel was permanently medically retired from the United States Coast Guard at the rank of Lieutenant. Joel earned a B.S. in Operations Research and Computer Analysis from the U.S. Coast Guard Academy, a M.S. in Information Technology: Database Systems Technology from the University of Maryland, University College, and a M.A. in National Security and Strategic Studies with an area of study in Information Operations from the U.S. Naval War College. Before coming to the lab, he worked for the USCG for six years doing maritime law enforcement and operational planning, the department of interior, and a startup focused on training and development. His expertise is in data analysis, AI/ML, serious gaming, algorithm development, and decision support systems.

How to Build Unethical Wargames[2]

Unethical Wargaming Working Group
Connections US 2021 Wargaming Conference

ED McGrady, PhD
Principal, Monk's Hood Media LLC
Adjunct Senior Fellow, Center for New American Security
703 395 5340(c)
mcgradye@gmail.com

Abstract

In this paper I consider the internal and external factors that affect the ethical environment for a game. I'm concerned about things like whether the players are treated fairly, is the game objective, and whether the game is used by others in an ethical way. I examine each of these challenges in turn. Fairness is defined as taking away agency from the players, either directly or indirectly, either by other players or the game itself. If players do not have the full and accurate set of options available to them, perhaps because something has been excluded or perhaps because game Control adjudicates incorrectly, then their play in the game is essentially a lie, and a waste of time. Objective games are the most complex and difficult to understand, because there are simply so many ways to throw a game or make it "slant." In the paper I describe, in detail, a game design that looks to change players' perceptions, without explaining to them what the game's objective is. Finally, we conclude with the most common way that games are manipulated unethically, by outside forces simply stating incorrect things about the game. Games can be a powerful tool, they can affect players and people's perception of reality. Many people realize this, and use games and game results in ways that forward their agenda, regardless of the actual outcome of the games.

[2] The contents of this paper represent the opinion solely of the author.

Introduction

> *"They used to call the devil the father of lies. But for someone whose sin is meant to be pride, you'd think that lying would leave something of a sour taste. So, my theory is that when the devil wants to get something out of you, he doesn't lie at all. He tells you the exact, literal truth. And he lets you find your own way to hell."*
> *- Mike Carey*

My first thought on the topic of building unethical wargames is to find something unethical that people do, like cheat poor widows, and build a game to help them do that. But the focus of this working group seems to be more along the lines of: how do we build wargames that trick people into believing one thing when another is the truth? How to build a wargame that, in my parlance, is "slant" or "thrown." Note that this is an inclusive definition of "unethical" as it can include games not intentionally thrown, but also thrown by bad design practice or through fully believed but incorrect assumptions.

My idea of what "slant" means is also more expansive than simply trying to get the game to say what the designer wants to say. Slanting a game can mean the game has elements that move it in a direction, but the whole game is not totally thrown in favor of a particular outcome. You can denigrate a particular new fighter aircraft in a game, but still have a reasonable overall outcome. You can also misrepresent a totally fine game. All these fit into what I'm talking about here.

That also seems like a pretty simple proposition: just assert whatever it is you believe, for example that lizard people really rule the world, and keep doing so until it becomes a fact.[3] No need to put yourself out doing a game. This is already simply part of doing business in 2021. The con is the truth, and all truth is a con.

In fact, it would seem that a lot of conventional wisdom and policy is built on whoever has the longest and loudest argument. Or at least whose argument is believed by the right people. Games need not apply. In fact if you look at the history of games being able to predict things, they do a reasonably good job of warning you of some of the dangers ahead. Even going back to the Russians and their game on the eastern front with Germany which predicted General Samsomov and General Rennenkampf had to get along if they were to succeed. They didn't get

3 https://www.newschannel5.com/news/newschannel-5-investigates/nashville-bombers-bizarre-writings-reveal-belief-in-aliens-and-lizard-people

along in the game, and they didn't get along in the war, and they ended up at Tannenberg.[4] So including the irrational is important, but believing the game is what really matters.[5]

This all goes to my first point: that we can think about games from two points of view

- Inside the game.
- Outside the game.[6]

While what happens inside the game is obvious, what happens outside the game can be expansive. It goes beyond the simple "sponsor – objective – scenario – mechanics" elements of a game to the reality that the game is embedded in. The social construct within which the game is created is part of the externality of the game. Whether it is a military (war) game, whether it is done by people who are greedy or slimy, whether it advances a political agenda that is horrid. It covers beliefs, intentions, context, social perspective, and philosophical systems.

My first two examples are of this broader externality. They are beyond the game itself and question the purpose and underlying social structure within which the game is created.

And, if we are to design an unethical game, this externality will provide the fuel with which we will fire our success.

Ethics is, of course, situational and changes with the perspective and position of the observer.[7] As a game player it could be that the game is "unfair" in the sense that I am not treated the same way that other players are being treated. As the designer it could be that the game is slanted toward one point of view or another, in particular the "wrong" point of view. And as the consumer of a professional game, it could be that the results of the game are used to support something that the game had absolutely nothing to do with.

---

[4] The best reference for this event seems to be Perla, Peter P (2012) *Peter Perla's The Art of Wargaming*. History of Wargaming Project. Pages 92-93.

[5] In fact the list of games that warned about disaster and were ignored is pretty long: https://web.archive.org/web/20060426004030/http://www.fema.gov/news/newsrelease.fema?id=13051, https://www.npr.org/2020/05/19/853816473/years-before-the-pandemic-war-games-predicted-a-global-tempest , https://www.wbur.org/onpoint/2020/07/28/election-war-games-trump-scenario, https://cneos.jpl.nasa.gov/pd/cs/pdc19/ (wait, that last one hasn't happened, yet).

[6] Past writers on wargame pathologies have described an "outer" and "inner" game similar to what we are discussing here. However, in our construct the "outside" elements are universal, not focused on the particular culture or organization within which the game is embedded. We include a critique of the whole society, its foundational assumptions and points of view, in our construct. Which broadens the overall discussion. In other words, you might consider the idea of a wargame to be unethical. Or you might consider the power structure of the society within which the game exists as invalid. This deals with things well beyond any direct association with the game. For a discussion of inside and outside points of view see, for example, *"Pathologies of War Gaming and What to Do About It"* Stephen Downes-Martin et al. Naval War College Briefing Feb 2004

[7] Of course, it's not "obvious" but we are making provisions so that we don't end up judging our gentle readers who may have beliefs that others would accuse of being unethical. See, for example, Mosteller, Timothy. *Relativism: A Guide for the Perplexed.* 2008.

Using those categories: players, designers, and consumers, I think there are three ways to view an "unethical" or what I would call "slant" or "thrown" game:

- Are the players treated fairly? Do the players retain their individual agency in the game?
- Is the game objective regarding the topic and objectives?
- Are the results of the game used in an accurate, appropriate, and logical way?

I will therefore break this essay into three parts. In the first part I'll discuss how to treat players unfairly. There are, of course, several ways to do this but in general it all involves either the other players or the game designer doing ill to the player. I think of this in a rather unusual way: the ill things that other people do in a game end up taking away agency from players. Whether it's cheating, lack of fairness, or other manipulation, these remove players' ability to act in the game in one fashion or another.

I'll then talk about how to build a game that is slant, one where what it seems to be is something other than what it is. There are, of course, hundreds of ways to do this so I'm going to be very specific in my description: I'm going to do an outline of a slant game.

The game will focus on the Reptilian Conspiracy[8] as an example of how to build unethical games. In doing so I'm going to be coming at the problem from the perspective that I really want to do this, not as some sort of Sci-Fi or toy design. It will mostly factor into the second two sections. My reasoning is that if I can build a game to legitimately raise questions along those lines, I'm really pushing the envelope of objectiveness and results (with apologies to those who actually believe in this line of reasoning).[9]

I'll also say a bit about the current and very pernicious fashion of using games as props for whatever idiotic idea the leadership has in mind. Many of which have nothing to do with the game itself.

Finally, I'll disclose my real agenda: explaining that external factors that have nothing to do with the game are the most powerful ways to create unethical games. They are the easiest and most lucrative to construct an unethical game. If you happen to be the evil one you may not even be aware that you are doing it. Powers of belief, agenda, and nature are a lot more powerful than simple tricks to throw a game slant. Working with them is quite powerful.

[8] See, for example, Icke, David. *The David Icke Guide to the Global Conspiracy (and how to end it)*. David Icke Books Ltd. October 2007. Or if you'd rather the short version: https://en.wikipedia.org/wiki/Reptilian_conspiracy_theory

[9] Sorry. Not sorry. The Reptilian Conspiracy is simply thinly disguised anti-Semitism by another name.

Players

Cheating is not the same as an unfair game.[10] But both abuse the players, in the first case it's the players who are involved, while in the second it's the game designer. Both target the players as victims of unethical behaviors.

There are as many ways to cheat as there are games. In professional games you have to also ask why someone is cheating, because professional games introduce more than just the usual "winning" or "profit" motives. In professional games you can also cheat to further an agenda, and it is a relatively common thing seen in games. If gone unnoticed by the controllers it can significantly affect the outcome of the game.

In my way of thinking about cheating there are two different types of cheating:

- Hard: where you deliberately do something that violates the assumed or defined rules of play
- Soft: where you shade your play in such a way as to achieve an outcome.

Hard cheating is classical cheating: holding out a card, using loaded dice, moving a piece when no one is looking, conspiring with the other side to throw the game, and other such activities. I would contend that these sorts of actions are rare in professional games, unless the controller has lost control of the game and players have begun not taking the game seriously. Which happens.

A gray area for hard cheating is rules-lawyering or playing against the game system and not playing the game itself. This is where players find some sort of loophole in the rules and exploit it for their benefit. In professional games, which are mostly free kriegsspiel anyway, this happens by the players rolling control on a decision or identifying a weakness that exists within control's knowledge or experience base.

For example, an inexperienced controller is playing with the CO of the Reserve Air Group and 10 of his subordinate fighter pilots in an air game. The CO objects to a decision made by control, and control gives in. Now any time in the future that control tries to make a decision that the players don't like, they all pile on control and push them to reverse or change their decision. Because control doesn't know the system and is not experienced enough to stand up to the CO, they give in, essentially letting the players run the game.

[10] Salen and Zimmerman in *Rules of Play* (MIT Press, 2003) define cheating as "to break the rules, to have a relationship to the formal system that is different from the relationship that the formal system itself presupposes and endorses." They are focused on player cheating, but in my construct Control can also cheat by violanting the explicitly or implicitly specified rules of play or outside the game behavior that construct the game. When players do it it's called cheating, when Control or the designer does it I call it an "unfair" game. This reflects the fact that cheating by control almost inevitably unbalances the ideas in the game in favor of one agenda or another (not necessarily one "side" or another). Players are good at detecting this. And they don't like it.

While that is a bit of an extreme example, players can and do routinely attempt to bully game controllers on certain decisions and situations. Often they are successful.

Soft cheating is far more common, though certainly not common enough in my experience to be a major problem in game results. Soft cheating occurs when a player decides to throw the game by not doing something they should do or doing something they should not. This tends to occur more on the "blue" (US) side than on the adversary side, probably because adversary players are less tied to particular game outcomes. Though I will mention ways it can occur.

So, for example, a blue player's flag officer has made it perfectly clear that a new bomb is not what the service wants to buy. Instead, they want to buy a missile. The lower ranks are well aware of this, and so when those individuals play in a game they find few reasons to use the bomber, making the missiles the hero of the game. While this is a rather ham-handed way of doing things and easily caught by objective analysts, it is a surprisingly common tactic.

A more subtle way of doing the same thing would be to use the bombs only in situations where they were destined to fail, a low probability target or a target out of parameters that places the launch platform at risk. Then the bomb missions would fail often, while the missile missions would succeed. Players would claim that that was due to some inherent advantage of the missile, which is why they played it as they did. This more subtle approach can be harder to detect without a careful analysis of game play and outcomes.

Truthfully you should probably be doing an analysis to figure this out, not a game. But that's another story.

Cheating by controllers involves manipulating the game in ways that are unfair to players within the context of the agreed upon world of the game.[11] It is possible, even likely, that a professional game could be unbalanced, where one side or another is at a disadvantage toward "winning". But within that unbalanced context, within the accepted structure and rules of that game, if the designer or control then manipulates that unbalanced system, they are being unfair.

I categorize this more in line with cheating than design problems, because the focus is on the players and not on the overall game design. Here control sets up a particular set of players to succeed or fail at a task regardless of the common sense of the situation. For example, Red's missiles miss all of the time because Blue's countermeasures are very effective. Red continually

[11] Why is being unfair to players unethical? Because it violates the magical circle within the game. Now by "unfair" I don't mean some sort of game required balance issue, rather that the Control for the game deliberately slants one or more factors in the game without regard for how these things work in the real world, and without an underlying game requirement to do so. So, one side's aircraft have maintenance and logistics factored into sortie rates, but the other side's don't simply because the designers didn't care or notice.

sends its missiles in, and they continually miss. Red has no chance of either understanding why or changing the situation.

This is unfair not because it's unrealistic (it may or may not be realistic) but because it removes the agency from the players. If players are required to make a decision, or do a task, in the game, it should be a task or decision that has some expectation of success associated with it. To remove player agency in those tasks is unfair to the players, kills their game play experience, and will lead to player resentment.

But is this cheating? I contend it is because it is cheating the players out of their game experience, making them into nothing more than tools that control uses to further its ends in the game. Which is exactly what player on player cheating does.

When a player stacks the deck, or loads the dice, or does something else to manipulate the situation in a game it's cheating. But what cheating does is not just bias the outcomes, it takes away the decisions that the other players are making in the game. Even though they do not realize it, their decision space has been diminished by the cheating that is occurring. This diminishment can happen because of hard cheating, or because control is doing something to diminish the role of the player.

The players are being treated as something less than fully human, as objects and not subjects. Which is the ultimate problem with cheating, that you are diminishing the other participants in the game and their freedom of action.

But there is a more "meta" interpretation of this problem of treating the players as less than human. It reinforces the overall Department of Defense's (and whole of US Government) tendency to treat humans as objects in their thought and analyses. The whole idea of "human terrain"[12] has been used in attempts at models and simulations, including those used in games, to try and characterize the complex and variable relationships that humans construct. "If only we can manipulate people like objects, our information campaigns would be so much better." This idea permeates how people are characterized in population models, and this idea of the reified player[13]

In role playing games, which are very similar to professional games (though everyone is very reluctant to admit it) the controller has to be careful not to show bias towards one or more players. Since a lot of what the controller does is not observable by the players the players have to be confident that the controller is not up to something behind the scenes. Giving one player advantages in experience or potential combat opponents, or, worse, doing hard cheating on die

[12] https://en.wikipedia.org/wiki/Human_Terrain_System

[13] Treating players as objects. See, for example, Honneth, Axel (ed.). *Reification: A New Look at an Old Idea*. Oxford. 2008.

rolls for that player, will lead to resentment on the part of the other players. Even though this may help the party as a whole in the game, the sense of unfairness is present in many species, not just humans and it matters even if everyone comes out ahead in the end.[14]

Objective Games

I am not building a game on the Reptilian Conspiracy. I have changed my mind.

> *[No, I haven't. Throughout this section I'm going to interleave commentary on what is really going on, with the game proposal sort of language that I'd use to tell the players and sponsors what is going on. My goal in both of these games is to lure the players into becoming the lizard people, into becoming the conspiracy, and seeing globe-spanning conspiracies as the only logical end-state for modern economies and power structures.*
>
> *The idea of doing something different than what you tell the players is both a common mechanic[15] in wargaming and fraught with danger. Implicitly it means that you, the arbiter of what is real, are lying to the players. If they discover the "big lie" then everything else you have told them is also suspect and the game can collapse in a puddle of mistrust. Here we are doubling down on this trope, by never revealing to the players what our agenda is. Even in our analysis and post-game build out we will continue with the charade and not disclose our true intentions.]*

Instead of lizard people, I will describe two different games that, if executed, will focus on power structures, power sharing, and information in the global sphere. One game will allow players to explore the intricacies of how power and finances are managed in the 2020s. Another game will focus on information and how information can be used to frame social and political issues across the globe. In their roles, players will represent the key leaders in government and industry and make decisions consistent with existing and stated corporate goals.

Players will either represent the political, national security, economic, or diplomatic elements of their nation's power. While corporate players represent the primary owners and financial drivers behind corporations (not just the CEOs). This includes the banking system, as well as the financial markets.

---

[14] See, for example: Bekoff, Marc and Jessica Pierce. "Wild Justice: Honor and Fairness among Beasts at Play," *American Journal of Play*. Spring 2009.

[15] It is often called the "hidden scenario" technique and has been used to great effect in games dealing, in fact, with unethical or questionable behaviors. See, for example, Price, Timothy McCoy (May 1985) *Home Front 86*, and the art game “Train” (https://venturebeat.com/2013/05/11/brenda-romero-train-board-game-holocaust/)

[I have placed the players in the roles where agency exists for the kinds of decisions I expect them to make. I will not assign cabals or groups ahead of time, I'll let the players realize that collaborating in groups will work out better for them than not collaborating. Also, I'm going to build in mechanics that reward international cooperation over national-only cooperation. I'm not going into a lot of detail here because I'm not actually designing the game.]

For adjudication the games will rely on existing economic and social models derived from the literature. But because of time pressures and the level of play expected the models will be abstracted by control to represent bulk population attitudes and large-scale financial flows. I'll be managing those models as Control.

[Here is a bunch of complete bullshit. I'm hedging so I can make up models that will do what I want them to do. Of course, I will have tons of references that no one will check, and I'll do my best to get the structure and lingo right. But the models will be designed to do what I want them to do: reward conspiracy and punish good behavior.]

Mechanically there will be provisions to allow individuals to play competitively, but there is nothing stopping them from cooperating. In fact, the mechanics will encourage cooperation as market manipulation will give larger profits than "normal" markets. Likewise, cooperation internationally will allow both sides to stay in power through conflicts that really are not conflicts, but which produce lots of byproduct contracts and funding they can give to their business co-conspirators.[16]

The goal is to create an "insider" set of rules that the players themselves establish in order to both "win" as well as internalize the ideas in the game. The players imagine themselves coming about these rules naturally but in fact the game design is constructed to lead them there. Breaking the implied rules should not only punish the players in terms of victory, but also in terms of insider/outsider group dynamics.[17]

[In this game mechanics will be essential. In particular I will need a clever and subtle way of rewarding international collaboration and punishing other types of

[16] Here we use player cooperation as a tool to achieve our corrupt objectives. But players can also cooperate on their own in games, causing what Hank Brightman calls "corruption between players." This kind of corruption can occur with opposing sides cooperating to throw the game. Here we are using that technique, but to throw the game as Control and not the players. Hank Brightman in "Nash in Najaf: Game Theory and Its Applicability to the Iraqi Conflict", *Air & Space Power Journal*, Fall 2007, Vol. XX1, No. 3, Pp 35-41 (https://www.files.ethz.ch/isn/120990/fal07.pdf)

[17] This is similar to the idea of Kayfabe in wrestling (https://www.edge.org/response-detail/11783) or a migration toward corruption in a game theoretic model (Brightman, Hank J. Nash in Najaf: (Fall 2007) *Game Theory and its Applicability to the Iraq Conflict*. Air and Space Power Journal.

actions. For both games I will need some way to represent the information environment that players will believe. I may have people actually playing some of those roles, as media titans or the general public. One mechanic that is always used in these sorts of games is a player role card where the players are told what they are supposed to do. I can't do this here because the whole point of the game is to get the players to believe in their roles, so their roles have to emerge organically from the scenario and mechanics of the game.

I am using the most powerful tool of gaming, participation and suspension of disbelief, to accomplish my objectives. I want the players to learn from the game that there are international conspiracies and look at all of the news and other data through that lens.]

Of course, each time anyone cooperatives or collaborates there will be the risk of exposure. Something that players will need to deal with through any number of suggested dirty tricks ranging from outright force (assassinations) to monetary bribes to ruined reputations through hack and release operations. Joint cover-ups will be more effective than individual cover-ups.

[Here is where the two games will diverge. In the economic game it will focus mostly on power with information being only one additional element of power. For the information game I want the players' reputations to be at stake, as well as their cash or power. This will require a subtle and complex approach toward information in the game. My preferred method is to have public players representing the "Greek Chorus" of social media and public preference. In the game I will give these players charters and roles designed to drive the other players toward my objectives. I won't worry about the social media or public players learning the ideas from the game, instead I'll simply use them as instruments against the other players.]

As play goes along those with more resources will be able to acquire more resources in a positive spiral. However, they will also need to cover up the fact they have resources, work with national leaders to manipulate the system to allow them to keep their resources, and so forth. Nationally elections will occur every now and then, but the players will be able to manipulate the elections through their resources.

[This goes back to the mechanics and models I'll use. How I'll represent the government is crucial. The government players will have to appear to be on the side of the people, while all of the actual incentives will focus on serving those within their conspiratorial groups. This will be a tricky but necessary mechanic. It

will be made even more difficult by the requirement that the players can't figure out what I'm trying to do here.]

The overall goal of the game is to put the players on a positive resource spiral, but one that requires increasing actions of collusion, manipulation, and dirty tricks in order to stay on the spiral and not lose power or money. This will not be disclosed to the players at the beginning of the game, rather the goal is to have them encounter these challenges, and solutions, on their own. The stated goal of the game will be to examine economic and social conflict through the 20th and 21st centuries. How it evolves and how systems evolve to deal with it.

[Remember, players never know where the game is going the first time they play it. Which is usually the case with professional games. So as the designer I have a huge advantage over them. I can lead them organically to a place where they realize what I want them to realize, even if they would have resisted if I had told them in advance.]

At the end, the players will have become the lizard people, manipulating the world economy and governments for their own ends. But they won't know that.

[Actually, making them realize they are lizards is probably a bridge too far. Instead, I'll scale down my ambitions during the design phase and settle for just making them into international conspirators who are manipulating the world to get what they want. What I will do is steadily increase the nastiness of the actions they are required to do.]

The problem with this game is that it needs a lot of players. It's hard to have one, or more, conspiracies without conspirators. What if I only have two players and want to do a boardgame?

That is trickier. I'd prefer 5-6 players and to make it a Euro-game with some mechanics like worker placement and random tasks through cards, but you could also do it as a two player. Each player would be responsible for building a network of influential people in order to manage their wealth and power. In the information game you would be accumulating fame and influence, while in the economic game it would be money and power. You would need to gather people like media moguls, Instagram influencers, and others to your cause. You would use them to manipulate either the global economy, or the global information environment. I'd lose the whole security aspect of things because, short of major power war, wars are just part of the environment (from which you can profit).

Players might have cards that allowed them to control events, recruit key people, and engage in various kinds of positive and negative actions. Since this is a more rigid, and open, game than I'd like I'll need to have a mechanic that gives the players a choice: do the right thing

or do what they need to do. Again, I will set the mechanics so that conspiratorial choices are presented as better or more effective than free and fair ones.

I am lying to the players from the very start of the game. I am deliberately keeping the topic and intent of the game from the players because I know they will resist my point of view if I share it with them. This is a power of professional, free, closed kriegsspiel. Everyone but the players know where they are going. Build trust with the players and they are easily manipulated.

I am also using the most decisive power of gaming: immersion and ownership of decisions, to manipulate the players. By experiencing those decisions for themselves as "inevitable" I'm manipulating the players' experiences (games provide synthetic experience) and in doing so I am implanting in their minds the idea that conspiracies are a reasonable way for the powerful to deal with the world. If I came right out and said "we're playing a lizard person game" the players would either a) quit and ridicule me or b) treat it as a lark.

> *[So, the first unethical thing that happens is the big lie. The obscuring of the intent and objective of the game from the players (and possibly others) in order to deliberately manipulate the players toward your own purposes.]*

I am also obscuring and fudging the models and adjudication in the game in order to push things in the direction I want them to go in. I don't necessarily have to be either obvious, or even deviate from the literature to do it. Most economic and social projection models have error bars. I'm going to use them. Many disagree, I'll pick the ones that benefit me. Even more commonly the models don't exist, so I'll make some assumptions and make them up.

I'm lying through my choices, in a way that is exceedingly hard to call me out on.

> *[The second unethical thing is the technical lie. I will use the gaps, uncertainty, and obscurity of the material to my advantage. I certainly will not tell anyone I'm doing this. Perhaps including myself.]*

I have set a serious design challenge for myself in this project. I need mechanics that are subtle and allow the players to believe they have arrived at the answer through their own actions. But if I can do it the payoff will be substantial because then the players will really believe my lie. Not because I convinced them, but because they convinced themselves.[18]

[18] There is evidence that this sort of big lie works in the real world. See, for example, https://medium.com/curiouserinstitute/a-game-designers-analysis-of-qanon-580972548be5 and Papasavva, A. et al. (20 Oct 2020) "'Is it a Qoincidence?': A First Step Towards Understanding and Characterizing the QAnon Movement on Voat.co," Computers and Society. arXiv:2009.04885v3 as well as Zeeuw, D. et al. (2 Nov 2020) "Tracing normiefication: A cross-platform analysis of the QAnon conspiracy theory" *First Monday*, Volume 25, Number 11 https://firstmonday.org/ojs/index.php/fm/article/download/10643/9998 doi: https://dx.doi.org/10.5210/fm.v25i11.10643

This is not impossible, but it's too difficult to do as part of this paper. Think of the mechanics of the boardgame Diplomacy. They are simple, extremely simple. You have armies and fleets. You can support, break support, attack, and move. Fleets can do amphibious movement of armies. You write your orders down and execute them simultaneously. There are seven players. That is about all there is to it. But the behavior that emerges is extremely complex and moves the players toward a balance between agreement and betrayal.

I need mechanics like Diplomacy mechanics. Seemingly simple but ones that create complex behaviors that point players in the direction I want them to go.

> *[The upshot is that you should not forget that mechanics, as much as scenario, player roles, and adjudication, can influence the trajectory of a game.]*

Finally, since this is a game designed to focus on the players, I don't want post-game analyses and reports influencing the players perception of the game. But what I do want is a hotwashup designed to cement my ideas in the players 'minds. I will facilitate the hotwash in such a way that I amplify the points I want to make, once they are made by the players, and downplay those players or ideas that don't agree with me. Of course, I need to seem fair, I cannot be obvious about this, but with a good facilitator and the right set of questions I can cement my lessons.[19]

> *[While this game focused on the players, not the information coming out of the game, the intent is for the players to influence others and go write their own reports about the game. How much more acceptable will be their conclusions if they make them, as opposed to either the designer or the analysts].*

Accurate and Logical

Most of the time when I hear someone complaining about a game it's about a circle of worthies declaiming a game as demonstrating one important thing or another. In other words, someone did a game, it got briefed to SECDEF, now SECDEF thinks that the game showed he (or she?) should make a certain decision. Of course, the game in question never actually showed that, was terrible in design, construction, and adjudication, and actually barely covered the topic of interest to the SECDEF.

Inevitably the discussion circles around to how we can fix this.

Based on my 30 years of being an analyst in DoD I have to say that you cannot fix this.

[19] For a discussion of how to do this see: Downes-Martin, S. (15 Feb 2020) "Preference Reversal Effects and Wargaming." Connections North 2020 Wargaming Conference. https://paxsims.files.wordpress.com/2020/02/2020-preference-reversal-effects-and-wargaming.pdf

Most game designers and players do their best to create games that are accurate, logical, consistent, and focused on the topic. (Playability is another matter completely). What happens is their best efforts are often not good enough. And many games (and analyses) are misinterpreted after they are completed. After the game ends and is reported out, elements of the report or players' impressions from the game are given life as rumors, briefings, and "facts" that go on to take on a life of their own.

Stopping this process is like stopping gossip. A good idea in principle, but difficult to execute in reality.

In the conspiracy game I described above the players are the seeds of discontent I want to sow into the general population after the game ends. They now will have some experience building and using conspiracies in the game and should be looking for them in real life. I will, of course, write a report about my game that emphasizes that we have played many rounds and repetitions of the game, and every single time we devolved to a world that was manipulated by the rich, famous, and powerful.

> *[The players are the ambassadors for the game. What they see and experience in the game will come out in their work as they go about their day jobs after the game ends. Suppose one is sitting in on a meeting where the topic of the game comes up. Their opinion about what happened in the game will be shared. Because it is an experience, an experience had by a colleague, in a quasi-technical (game) environment it will have credibility. So, it's not just enough to write the correct report, we also want the players to be manipulated (or not manipulated) as well.*
>
> *But my report is also cleverly framed to have impact. I know that the DoD worships at the altar of technology and statistics. Running multiple games was to my advantage since it exposed a lot of players to my point of view. But it also means I can reframe the game in the final report as a "statistically relevant" experiment designed to show that, sociologically, people converge to conspiracy. Now I'm not going to actually expose my report to a real journal (though it would likely get through), instead I'm going to publish and distribute it informally so that many people see it, but few people see it who can do an adequate review.]*

"But wait," you say, "what about the lizard people? You have only demonstrated that you can do a game on conspiracies. Not on getting people to believe in lizard people specifically. That is probably a bridge too far."

Not necessarily. Now we have to turn to the external frame of the game, how the game is viewed outside of the context of the game. I need to get my game results paper, and some of the players, in front of those who already believe in the lizard people conspiracy.[20] They will re-interpret the results to include the farther elements of the conspiracy. I will have shown that their explanation is likely to be correct.

> *[This is what happens in the Pentagon. The game does not convince people. Rather people who are already convinced interpret the outcome of the game to reinforce their position. If things have to be ignored, modified, or rearranged to do so then that is just fine. I will note that this also happens with analyses and policy statements. It's part of bureaucratic and organizational dynamics. Just be glad that what you do is part of the conversation!]*

The Subtle Side of Pathologies

The point of all this is that you do not have to have obvious indicators of deception in a game. Subtle design choices, manipulation of player and observer biases that already exist (especially toward technology, statistics, and "validity"), and clever insertion of game outcomes into the conversation and narrative around an issue can accomplish a lot. Of course, you can always just bulldoze your way toward an outcome, as I said at the beginning simply stating something as fact over and over again seems to work just fine. No game needed.

But a well-designed game has the advantage of providing synthetic experience to the players. Players will take away that experience as "negative training," making them multiple advocates for whatever position you might take. This then becomes further evidence that your position is correct, reinforced by a report that sounds technical and "sciency."[21]

It's a lot easier exploiting external factors that have absolutely nothing to do with the game in order to shape the conversation to your will. It becomes a downhill ride as you reinforce beliefs and trends that already exist and are waiting to be exploited.

[20] Which won't be really that hard: https://www.reddit.com/domain/forum.davidicke.com/

[21] Typically, with references to repetition, hypothesis, experiments, and data. Regardless of whether anything objective, thoughtful, or innovative was actually accomplished.

What Is To Be Done?

I think that we can come to several principles that inform how games are treated, and how games treat topics:

- Don't be evil. In other words, as a designer consider your own ethical limits and boundaries and be strong enough to avoid crossing them. This can be easier said than done in many cases. For example, your political or social outlook may in itself be considered criminal or unethical by others. In this case an ethical game from your perspective is perfectly unethical from someone else's. Likewise, employment is always a consideration, and the employment of subordinates usually matters to those responsible for them. So, choices must be made. But within those limitations have some respect for what a good game is, and work toward it.
- Design games that are fair and accurate. Fair to the players in that they treat all players with equal respect and give players agency within the scope of their roles. And make sure the game is accurate to the real world as best you can, given all the constraints and trade-offs that must be made.
- Speak clearly and accurately about what your game shows, and what it does not show. You cannot change what other people say, but you can be clear about what you say. Lack of clarity, long-windedness, and poor writing can be used by others against you as they "interpret" your game results.
- Always get in front of the right people. This means get the right players for your game, game for good sponsors, and push to be in, or doing, the final briefing. Shaping the people around you will help shape how your results are interpreted.
- Stop trying to apply a science model to something that is not objective. Games are a subjective experience, an event, an occasion which cannot be replicated no matter how hard you try.[22] The more we perpetuate the myth that games can be used in some sort of validity exercise, the more we risk an unethical outcome. And, most importantly, the more we dilute the actual good that games can do.

For the purposes of this working group, you would just reverse the logic in the above and satisfy the requirement to be unethical.

[22] There are exceptions to everything, abstract or toy games can and are used in research, but they have very specific and limited mechanics and scenarios.

About the author

Dr. McGrady writes, speaks, and teaches on the design of professional games. He is an adjunct senior fellow in gaming at CNAS, teaches and manages game design courses for MORS/Virginia Tech, and runs a business devoted to using games and game techniques to bring innovative experiences in new areas. In the past Dr. McGrady built and directed a team of 10-20 analysts at CNA devoted to the design and execution of professional games. Dr. McGrady has written, taught and presented on the topic of games and their use in organizational and individual learning. He has designed and run games for many different clients ranging from the White House to the Department of Agriculture. Dr. McGrady has also built a team at CNA devoted to chemical and biological response operations, including domestic response operations. Dr . McGrady has deployed as an analyst with US Forces in Haiti during operation *Uphold Democracy*, onboard *USS Nimitz* for Desert Storm and with operational E-2C squadrons. Dr. McGrady holds a Ph.D. in Chemical Engineering from the University of Michigan. He has published extensively in the Chemical Engineering, physics, and national security literature and is widely cited for his work on the mathematics of aggregation and fragmentation.

Exploit Group Dynamics to Corrupt a Professional Wargame[23]

Unethical Wargaming Working Group
Connections US 2021 Wargaming Conference

Stephen Downes-Martin, PhD
Research Fellow, Naval War College
401-935-4808
stephen.downesmartin@gmail.com
https://sites.google.com/site/stephendownesmartin/

Abstract

This paper describes some methods of corrupting a wargame by exploiting the dynamics of small group discussions that are a common feature of wargames designed for senior officers. This corruption can be deliberate or an inadvertent side effect of incompetence. The purpose of the thought experiment is to discover wargame design principles that might be missed if one were to consider lessons learned only from well designed wargames, to inoculate wargaming against deliberate and inadvertent manipulation of wargame design, and to protect against self-deception. Design principles that protect against malign manipulation of group discussions during pre and post game engagement with the sponsor and the sponsor's command, game design and execution are proposed to protect against the wargame being invisibly corrupted.

The Malign Wargaming Thought Experiment

If we define well-designed and executed professional wargames that successfully support the national security objectives of their sponsors as benign, then in contrast malign games are those that deceive intelligent, rational and honest sponsors and the decision makers up their chain of command to make decisions that do not best support their objectives or are actively damaging to national security in order to better support the personal proclivities of the malign practitioner. Malign wargaming is the set of behaviors and practices engaged in by any of the stakeholders that produce a malign game.

[23] The contents of this paper represent the opinion solely of the author.

Consider the thought experiment of deriving lessons learned and mistakes to avoid if you were to deliberately design a malign wargame with the objective of deceiving the sponsor. This thought experiment is useful in three ways:

1. discovering wargame design principles that go beyond the lessons learned that follow benign wargaming (Downes-Martin 2016, 2017, 2019a, ch.16 Longley-Brown 2019),
2. inoculating wargaming against deliberate and inadvertent manipulation of wargame design by senior stakeholders (Downes-Martin 2014) and
3. protecting ourselves from self-deception by inadvertent malign practice.

The normal sources for best practices and lessons learned are well-run benign games that provide "what we did well" and "what went wrong" in the just completed game, which in turn should then provide "how to do things right" and "what mistakes to avoid" lessons learned for future games. They provide characteristics to seek and behaviors that interfere with those characteristics to avoid, i.e. best practices and lessons learned from benign games. However, it is not good enough simply to do lessons learned reviews following benign games. The opposite of a benign game is not a poorly designed or executed game, it is in fact a malign game as defined above. Lessons learned will not surface or identify malign practices since malign practices can also be used to deceive the lessons learned and review processes themselves.

Senior Stakeholders make Malign Wargames Probable

One cannot assume that malign wargames do not occur. Three risk factors have been identified as present in nearly all cases of scientific fraud.[24] The perpetrators

> *"knew, or thought they knew, what the answer to the problem they were considering would turn out to be if they went to all the trouble of doing the work properly; were under career pressure; and were working in a field where individual experiments are not expected to be precisely reproducible*"
> (Goodstein 2010, Shermer 2010).

One must accept the possibility that all three factors are present among the senior stakeholders of any wargame that addresses important national-security issues (Downes-Martin 2014). The first risk factor is usually present since research indicates that senior people tend to be overconfident in their ability to control events (Gladwell 2009; Chapter 10 in Wills 1994; Langer 1975). They may be under career pressure from their organizations and communities to

[24] The word "fraud" is used by the scientific community when discussing intellectual impropriety by scientists. However, deliberate malfeasance is not required for the presence of deliberate intellectual distortion by senior people and so deliberate malfeasance should not be assumed. A better term for our purposes might be "deliberate intellectual distortion for non-cynical purposes". See (Bailey 2021) for an excellent review of science fraud.

support funded programs of record and preferred concepts. Finally professional wargames are inherently unreproducible.

Therefore senior stakeholders are vulnerable to intellectual distortion and are highly motivated, perhaps with the best of intentions, to nudge the design, execution, analysis and reporting down pathways they believe to be best for national security independent of the wargame outcomes (Downes-Martin 2013, 2014). This does not guarantee that intellectual distortion takes place, however since it occurs in the wider national security world it would be naive at best and irresponsible at worst to assume that it does not occur within wargaming (Downes-Martin 2014).[25]

As H. G. Wells observed over a century ago (Wells 1913)

> *"... it is remarkable how elastic the measurements of quite honest and honourable men can become."*

Professional Wargaming

It is worth distinguishing the target of malign practice, the professional wargame, from hobby and entertainment games by their purpose. Professional Wargaming supports and informs decision making about the deployment and employment of people and resources in some future real-world situation, and where the gain or loss in the real world situation being gamed have serious consequences. Professional wargaming is normally thought of as applying to national security decision-making, but it is also applicable to the civilian world, subject to suitable and obvious modifications, including business investment and hospital pandemic response decisions for example. Although corporations that develop and sell entertainment games can engage in professional wargaming to investigate their marketplace and competitors, the games that they develop for purely entertainment reasons are not themselves professional games.

First note that all wargames used to inform professional decision-making must be able to distinguish between two or more alternatives, *one of which may be the status quo*. Ideally this is done in such a way that the sponsor of the wargame can rationally select one of the alternatives as better than the others, or even preference rank them all. The (possibly ranked) alternatives along with the wargame narrative are then added to other information available to the decision-maker to assist in making a decision about the alternatives. This general and abstract description of professional wargaming can be best understood by examining a few of the

[25] Consider the "Fat Leonard" (US) and "SNC Lavalin" (Canada) scandals. If this level of deliberate corruption occurs among senior military and civilian leaders, then well-intentioned intellectual distortion is a certainty. The idea that somehow "wargaming" by its nature is immune from the darker side of human nature is risible.

organizational purposes (Wong, Bae, Bartels & Smith 2019) to which wargames are put, for example:[26]

Purpose	Example alternatives: what is being gamed	Narrative based evidence: advantages & disadvantages	Example decisions
Operational plans and decisions	Courses of Action (COAs)	Advantages and disadvantages	Which COA to implement; decision to replan
Concept and technology development	New concepts and technologies	Advantages and disadvantages	Which concept(s) are worth further investigation and development
Training	What is being trained	Who improved and by how much	Who, if any, is promoted or moved to where.
Education	What is being taught	Who learned what and how well	Who, if any, is placed where based on what they have learned
Experiential and socialization	What is to be shared by the experience	Who learned what and what new information was generated	Task force construction; Readdress concepts being experienced or socialized

Professional wargaming tends to use discussion-based seminar game designs featuring small group discussions by stakeholders and players before they decide their moves. They do this for three main reasons. First, the players at professional wargames are usually senior military officers or civilian executives who are rarely hobby gamers. They lack patience with abstract games that bear little resemblance to their day to day responsibilities and thus lack motivation to learn the rules of a game -- no matter how well the game rules and abstractions *represent* the real world from which the players come. The game design needs to engage senior leaders in the wargame while requiring minimal personal gaming experience by them. Second, senior leaders are comfortable with small group discussions since much of their work involves such discussions. Finally for games whose objective is to explore novel future situations, seminar games support discussion among experts about topics that are not clear cut or well understood and for which there are few clear cut rules for adjudication. The result is that discussion based seminar gaming is often the preferred approach to professional wargaming.

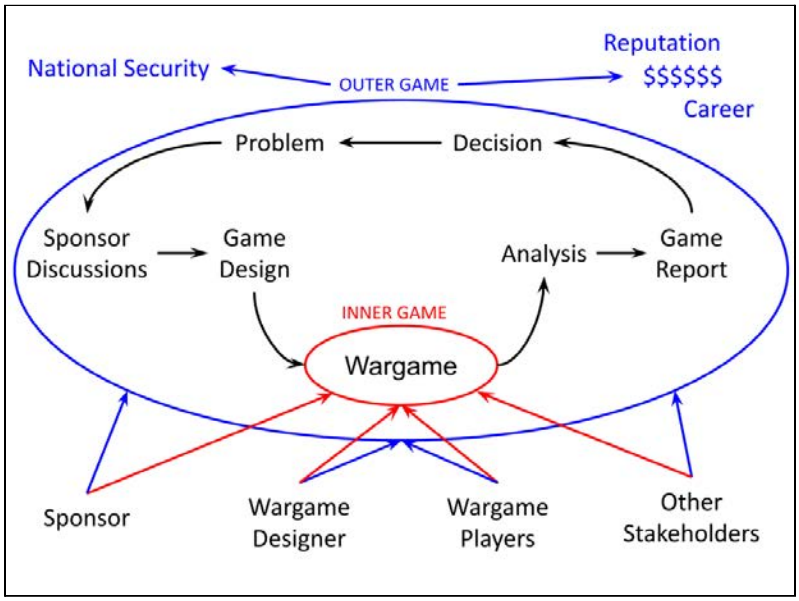
No matter what the proximate organizational purpose, the foundational purpose of a professional wargame is to assist in deterring war or winning current or future wars. Based on these observations it makes sense for a benign game to structure the discussion format and the

[26] It is possible the game narrative indicates some abstract notion of the status quo is the best decision. For example in concept development games it is possible that the narrative implies that none of the new concepts are better than the status quo. In planning games this might occur if none of the COAs are considered acceptable, and if the status quo of "do nothing" is not acceptable, the decision then is either to replan or accept the least bad COA.

information presented to discussants in ways that best support the sponsor's game objectives, and this provides one possible departure point for the malign wargame thought experiment. How can we manipulate the content and format of the information and the discussion process for malign purposes? It turns out to be surprisingly easy.

The Inner and Outer Game

It is not just during game play and among players in a game cell that discussions occur. Even for non-discussion based game designs, there are many opportunities for malign manipulation of game related discussions to influence decision making. To understand this consider the interaction between the Inner and Outer games of a wargame.

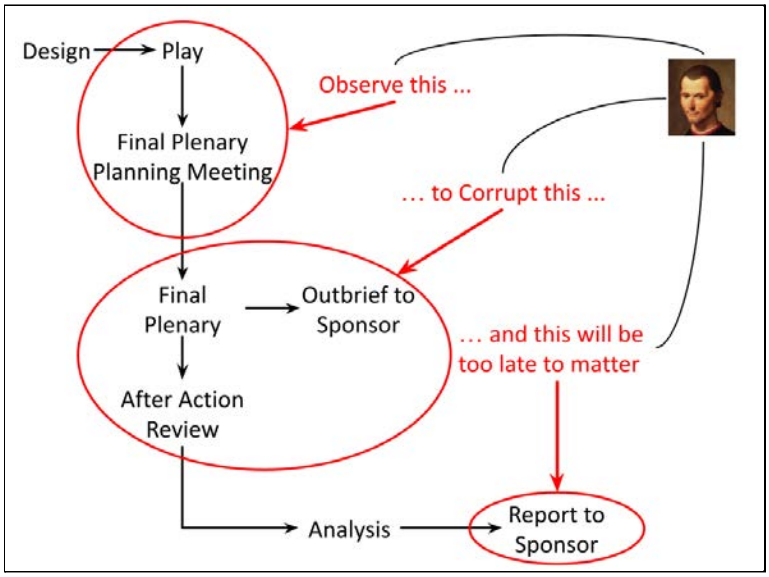


During the outer game, small group discussions occur at every stage of the game starting with the initial meeting with the sponsor, various design meetings prior to the game, post game analysis meetings and report briefings to the sponsor about what the game meant for the sponsor's objectives. During the inner game discussions occur during play, a final plenary, and post game the after action review, internal analysis, and about what to put in the report and how to say it. By the very act of engaging with the sponsor and other stakeholders you are playing the outer game. Therefore design and play, internal to your organization, an "outer game" tailored to the details of your inner game (what you know about your sponsor, other stakeholders, etc), either for each game or sporadically for a long-term repeat sponsor.

The inner game is what we normally consider to be the wargame. The primary players in the outer game are the sponsor (and the sponsor's organization), the game designer (and the designer's organization), and any stakeholders with an interest in the decision that the game is supposed to inform. The stated outer game objectives address some aspect of national security. However, there are unstated objectives dealing with careers, budgets and reputations. As argued earlier, it would be irresponsible and naive to ignore these unstated objectives, and experience one advises not to do so. One of the moves of an outer game player is to insert

themselves as players in the inner game. Hence we sometimes see sponsors playing in a game in a player cell or the adjudication team. Subject Matter Experts (SMEs) invited to play in the inner game are almost by definition outer game stakeholders. Note that wargame designers and their organizations usually provide supporting functions during an inner game that also provides them with opportunities for malign practice. If you "lose" the outer game, the inner game (a.k.a. "the wargame") becomes irrelevant no matter how brilliantly designed or competently executed.

The final plenary at the end of the wargame is a key opportunity for malign practice. The final plenary is a facilitated discussion between players from all player cells to go deeper into questions about the sponsor's objectives. It is best practice for wargame departments to hold a planning meeting to identify questions to ask at the final plenary, how to present those questions, and how they will facilitate the final plenary. The malign practitioner will be present during the game, observe the adjudication cell, each end-of-day planning meeting by the wargame department and each end-of-day player plenary (if held). This allows the malign practitioner to assess how each alternative is faring during the game, how the narrative produced by the game supports or damages their preferred alternative compared to the others, observe who in the game prefers which alternatives, and note the preferences held by each of the communities represented by players, wargame team members and other stakeholders involved in the game.



Push for the executive briefing to the game sponsor as soon as possible, preferably on the last day of the game, to exploit the vividness criterion (p. 116 Heuer 1999) and thus inoculate the sponsor against the later game report that might contradict your carefully crafted malign conclusions. Any wargame analysis report delivered weeks to months later is likely to be ignored on the grounds that the sponsor already has received the wargame briefing and thus knows the answer. In any case the sponsor most likely has moved on to other problems. The implied or explicit conclusions delivered in the outbrief or executive briefing becomes the de facto game result and the de juro result from the analysis is likely to go unread.

Malign success is not guaranteed, but it is made more likely by careful observation of the players and judicious use of the various methods of manipulating small group discussions during both the outer and inner game as described below.

Moves to Create Malign Professional Wargames -- and what to do about them

"Rule of Eight"

Research into the effect of group size on discussion (for example see Fay et al 2000) indicates that when the group size reaches and exceeds about ten people then the conversation shifts from interactive dialog to serial monolog, possibly due to the "many minds" problem (see for example Cooney et al 2020). Recommendations abound in the management and facilitation practice literature for restricting group size to about five people.

My own observations over several decades and a large number of meetings that involve group discussion is that only about eight people voluntarily make substantial contributions, i.e. after about eight people have spoken the remaining group members tend to remain mostly silent. A few additional people might make a comment or two. I have not observed any correlation between who contributes and the characteristics of the contributors other than who speaks early. If the discussion goes on long enough and the topic changes then I have observed that the contributors sometimes shift to a different (possibly overlapping) group of approximately eight people.

Malign Move: Be "inclusive" and invite everyone to each discussion throughout the outer and inner game, even people who you know disagree with your preferences (this will protect you against accusations of bias or of having a preconceived agenda). Then recruit your own team of six or more people and prepare them to seize the "discussion initiative" by speaking early and often in support of your agenda. This raises the probability that only the topics you are interested in are discussed and in a way that supports your agenda. The others will likely not notice what you are doing since they have been subconsciously acclimated to the "rule of 8" by past participation in poorly designed meetings.

Counter Move: Split large groups into groups of about 5 to 8 people. The groups either discuss the same topic in parallel and you integrate their results later, or you expand the breadth of the discussion by having the groups discuss different topics.

Brainstorming and BOGSATs

Immediately following its introduction in the 1950s brainstorming has routinely been debunked as an effective mechanism (Lehrer 2012) compared to easily implemented normative processes. This is quite separate from the observation that most DoD brainstorms do not even follow the primary rule for brainstorming, which is that there should be no criticism (positive or negative) of any idea that surfaces during the brainstorm. My own observation over the last few decades is that within two minutes of the start of a brainstorm a senior member has grunted or otherwise indicated approval or disapproval of some junior's idea – and the brainstorm is over. The senior member's implied approval or disapproval sets the direction of the discussion from that point onward.

Furthermore, it has long been proven that a disciplined three-stage normative approach gives superior results than those obtained from ill-disciplined methods such as brainstorming (even when run properly) or BOGSATS (Mullen et al 1991).

Malign Move: Prepare your "team of eight" within the larger group to raise pre-selected ideas you approve or disapprove of and to use body and verbal language to indicate approval or disapproval of them.

If possible, recruit a senior officer or civilian who agrees with your pre-selected ideas to participate. Then casually mention the possibility that the group will oppose those ideas and support others. You have primed that senior person to visibly react positively to anything that is close to the preferred ideas and negatively to anything that is close to the opposite. That senior person is innocent of malign intent, simply ignorant of how Brainstorming works (which is common among senior, highly self-confident, people).

Counter Move: Make sure groups have a maximum of eight people and use the proven normative approach for each. Group members first silently write down their ideas, proposals, responses or courses of action (depending on the task) at the start of the game move and without discussion. These are then posted to the wall for all to review and discuss. During the discussion additional ideas will surface and irrelevant responses removed. The third phase occurs after move submission and is in preparation for receiving the adjudication results. All three phases can be fitted into a short discussion group so long as it is kept to 8 or fewer disciplined people and is rigorously facilitated.[27]

[27] See (Goal/QPC 1996) for a very detailed description of one method of implementing a normative process, and (Downes-Martin et al 2107) for an example of successful use of the normative method applied to a working group.

Wisdom of Crowds or the Madness of Mobs?

The work on Superforecasting (Tetlock 2005, Tetlock & Gardner 2015) and on Wisdom of Crowds (Surowiecki 2004) is all too often grossly oversimplified into the claim that "aggregate group forecasts and decisions are much better than individual ones." This is clearly nonsense; a group of idiots is unlikely to make a better forecast or decision than a single expert, and a single expert is more than likely to make a better decision about their area of expertise than a group of experts in another field. No patient in their right mind would ask a group of expert sanitation engineers about their cancer treatment instead of the single expert oncologist. The wisdom of crowds and superforecasting research is much more nuanced, with interesting areas to manipulate, than the popular understanding of them.

A more accurate summary of the research might be "A group of experts satisfying four requirements who make a numerical decision or forecast about their area of expertise is more likely to make a better decision or forecast than a randomly selected individual from that group". Furthermore research indicates that a group brought together and using BOGSAT and Brainstorming consistently underperforms the same group using normative methods in which individuals first work independently, then in a group, then review and refine as individuals (Nijstad et al 2006; Lehrer 2012; Mullen et al 1991). The four requirements for a group to exhibit the "wisdom of crowds" rather than the "madness of mobs"[28] are (Surowiecki 2004, page 10):

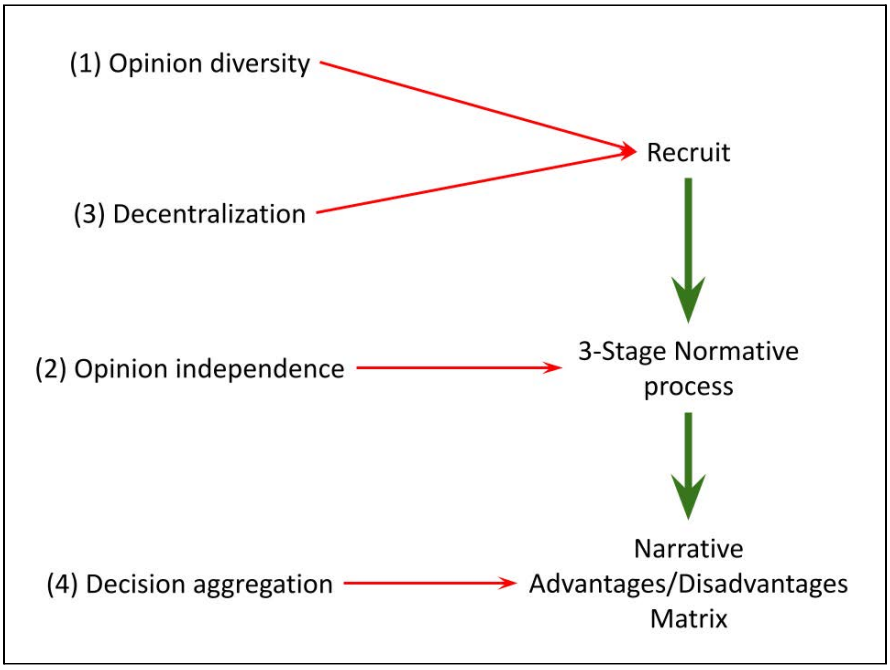
> ***"Diversity of opinion:*** Each person should have private information even if it's just an eccentric interpretation of the known facts.
>
> ***Independence:*** People's opinions aren't determined by the opinions of those around them (Lorenz et al 2011).
>
> ***Decentralization:*** People are able to specialize and draw on local knowledge.
>
> ***Aggregation:*** Some mechanism exists for turning private judgments into a collective decision."

While the DoD groups I have observed often satisfy the Decentralization requirement, they mostly fail on the other three for Wisdom of Crowds. They frequently consist of subject matter experts from the same communities of practice or Service with peer pressure to conform to doctrine. Opinions, in the form of statements or votes, are often collected sequentially and publically. Aggregation is often based on flawed voting schemes using junk arithmetic.[29]

[28] My thanks to Paul Vebber for the term "Madness of Mobs" in this context.
[29] For details of this kind of problem and how to deal with it see the briefings and reports on the Puppet Mastery web page at https://sites.google.com/site/stephendownesmartin/puppet-mastery.

Malign Move: Break the four requirements. For example, select a few senior (high ego) people who you know have opinions you prefer, add a larger group of junior people who are not expert in the topic under consideration, allow undisciplined brainstorming, then ask for decisions or answers or points sequentially starting with the most senior person. Then claim "wisdom of crowds". Use junk arithmetic (such as rank ordered decision matrices) to aggregate group opinions, and then manipulate those numbers using Puppet Mastery techniques[30].

Counter Move: Use the first and third requirement to recruit discussion group members. Use a normative approach to deal with the second requirement. If aggregating opinions use qualitative pro and con decision matrices, not popular junk arithmetic approaches (Puppet Mastery web page).



[30] https://sites.google.com/site/stephendownesmartin/puppet-mastery

Dishonesty Shift

Research indicates "that there is a stronger inclination to behave immorally in groups than individually" (Kocher et al 2016), resulting in group decisions that are less honest than the individuals would tolerate on their own. "Dishonest" in the context of the research means the group decisions break or skirt the ethical rules of the organization and societal norms, involve cheating and lying. Furthermore, the group discussions tend to shift the individuals' post-discussion norms of honest behavior towards dishonest. First the discussion tends to challenge the honesty norm, then inattention to one's own moral standards (during the actual discussion) and categorization malleability (the range in which dishonesty can occur without triggering self-assessment and self-examination) create the effect that "people can cheat, but their behaviors, which they would usually consider dishonest do not bear negatively on their self-concept (they are not forced to update their self-concept)" (Mazar et al 2008).

The research indicates that it is the small group communication that causes the shift towards dishonesty that enables group members "to coordinate on dishonest actions and change their beliefs about moral behavior". The group members establish "a new norm regarding (dis-)honest behavior" (Mazar et al 2008). Appeals to ethics standards seem to be effective in the short term (Mazar et al 2008) but there is little evidence for long term effectiveness (Kocher et al 2016).[31]

Malign Move: Decide if the decision you want is outside the moral norm of the group. If so, introduce what you want by small increments, nudging the group slowly down the immoral path by overemphasizing the benefits. If the decision goes towards what you do not want then make credible arguments about how that decision can be viewed as immoral based on possible knock-on effects of the decision you do not want.

Counter Move: Formally remind the group at the start of every discussion about the ethical standards required of the group and explicitly set these standards within the context of the wargame. For each "subgroup of eight" have a single named individual in charge who is accountable for the subgroup decision (Downes-Martin et al 2018).

[31] One must treat the results of research into dishonesty with caution since the three established risk factors for intellectual fraud (Goodstein 2010, Shermer 2010) discussed above also apply to researchers examining dishonesty. There is significant evidence of intellectual fraud by some of the researchers into dishonesty (Uri 2021, Goldman 2021).

Risky Shift

Research into risky or cautious shifts during group discussion looks at whether and when a group decision shifts to be riskier or more cautious than the decision that the individuals would have made on their own (Batteux et al 2017; Dodoiu et al 2017). One element driving the shift appears to be who bears the consequences of the decision – the group members, people the group members know (colleagues, friends, family), or people the group members do not know. There is evidence that individuals tend to be myopically risk averse when making decisions for themselves (Thaler et al 1997). Research indicates however that "risk preferences are attenuated when making decisions for other people: risk-averse participants take more risk for others whereas risk seeking participants take less" (Edelson et al 2018). Whether the group shows a risky or cautious shift depends on the culture from which the group is drawn and the size of the shift seems to depend on the degree of empathy the group feels for those who will bear the consequences and risks of the decision.

Research into leadership shows that "responsibility aversion" is driven by a desire for more "certainty about what constitutes the best choice when others' welfare is affected", that individuals "who are less responsibility averse have higher questionnaire-based and real-life leadership scores" and do not seek more certainty when making decisions that are risky for others than they seek when making decisions that are risky for themselves alone (Greenfieldboyce 2018; Edelson et al 2018). However, this research says nothing about the starting risk-seeking or risk-avoiding preference of the decision-making leader. Exploiting this shift is difficult if one does not have deep insight into the participants' characters. Nevertheless some suggested exploits and counters are provided as a starting point.

Malign Move: Understand the risk seeking or averseness of individuals in the group and give speaking preference to whichever best supports your objectives. Ignore the others. Emphasize the uncertainty of the situation if you want to delay decision making and trigger a panic decision based on prior beliefs (Dorner 1996). Since research indicates that good leaders tend to be risk neutral when making decisions for others, exclude them from the discussion by putting them in a separate room of "graybeards discussing important topics". This latter is a technique I have observed to be successfully used.

Counter Move: Identify good leaders for each "subgroup of eight" and hold them personally accountable. Design the discussion to include the probabilities, consequences, the need for risk neutrality and present each situation in terms of losses and gains. During post game analysis report on the risk seeking and aversion profiles of the members and caveat the decisions accordingly (Downes-Martin et al 2018).

Gender Ratios

Anecdotal evidence from private conversations with professional facilitators indicate that, ceteris paribus, the gender ratio in a small group undertaking a professional discussion will have a significant effect on the performance of that group. Recently a female professor who researches gender issues and consults for the US Special Forces (name withheld by request) told me that when women make up:

- less than a third of the group they tend not to speak up in support of their ideas,
- more than a third but less than half they tend to speak up in support of their ideas and the men present tend not to listen,
- more than half the group they tend to speak up in support of their ideas and the men present tend to listen.

Research indicates that the proportion of time women speak in small group discussions seems to depend on the gender ratio and the decision rule, i.e. whether the group is seeking a majority or consensus decision (Karpowitz & Mendelberg 2014, p 121). Very loosely summarised, men tend to discount womens' contributions during a small group discussion except when they are in the clear majority.

If these numbers are valid, then, ceteris paribus, we can identify a number of malign moves when breaking up a large group into small subgroups. Each of these depends on the actual numbers of male and female participants.

Malign Moves:

- Put the males with whose opinions you do not agree into subgroups with women making up more than one third but less than one half the group -- this will tend to disrupt the subgroup. However, you will lose the effect of women who do agree with you.
- So, if you have enough female participants, put the women with whose opinions you do agree into subgroups with males with whose opinions you do not agree such that the women are in the majority -- this will tend to reinforce your preferred opinion and dilute your opposition.
- On the other hand, if you want to disenfranchise women in the group completely while *appearing* to support diversity, spread them around all the subgroups but keep them below one third of each subgroup. This malign move might also be made by a benign but naive designer who is ignorant of gender ratio dynamics, and thus is an inadvertently malign move.

Counter Move:

- Having identified possible malign moves, the counter move becomes simple. Create as many subgroups as possible with the women in the majority. Avoid naively distributing women to every subgroup if this puts them into a minority in every subgroup.

Collusion Rings and Peer Review

Peer review is one of the more effective methods of assuring scientific integrity.[32] "It effectively subjects an author's work to the scrutiny of other experts in the field" (Kelly, Sadeghieh & Adeli 2014). Although the peer review process is applied to written products, by implication it is applied to the scientific processes used in the research being reported since those processes should also be part of the scientific report. Hobby game designers face rigorous scrutiny by the people who buy their products -- a form of peer review perhaps. Professional wargaming (dealing with real and serious national security problems) does not use peer review -- perhaps it should (recommended in Downes-Martin et al 2017). Such a peer review would ideally be carried out by sister wargaming organizations and wargame experts outside the organization that is producing the peer reviewed wargame. The precise implementation of the peer review would not exactly mimic that used in scientific publication (anonymity of the wargame producers for example is unlikely to be possible!), but the goal would be the same -- to assure the wargame stakeholders that the best possible practices had been employed and the results can and should be taken seriously.

We must now ask the question: "Are the three risk factors for intellectual fraud present among the reviewers?" The troubling answer is not only "yes", but that evidence of collusion rings among peer reviewers in several scientific fields has surfaced (Littman 2021).

Malign Moves:

How would a collusion ring work among wargamers, sponsors, senior players and other stakeholders? Examining how collusion rings work in the scientific fields provides some ideas:

- Hide conflicts of interest
- Collude before the game about moves to make and moves to avoid
- Non-playing senior officers *order* subordinates who are players to collude
- Non-playing senior officers *imply* career damage if subordinate players do not collude.
- Peer review chains of command play soft with each other

Counter Move -- Wargame Directors:

- must have the expertise to identify collusion, and the intestinal fortitude to face down their superiors when collusion occurs (Downes-Martin 2014)
- build into the game report a section titled "Wargame Validity" that addresses how well the game was designed to address the sponsors' objectives, how well the game was executed according to the design, and whether the players had the required expertise
- Shine a spotlight on collusion

[32] https://www.elsevier.com/reviewers/what-is-peer-review

What is to be Done?

Prepare yourself

1. Professional wargaming is a social activity, therefore you must be familiar with the social science of groups engaged in play and conflict. This includes the psychology of small group dynamics, preference reversals, prospect theory, decision analysis, negotiation analysis, game theory, and voting theory.
2. Wargaming has a purpose. Just "playing the game" is not a good enough objective, and "the sponsor liked the game" is not a good enough measure of success. Understand the sponsor's objective, and the objective of the sponsor's boss. Then design and play the outer game, because if you lose the outer game the inner game does not matter.
3. Use pre-game surveys to collect participant (players and other stakeholders) demographics and map out in advance which participants have biases for and against which alternatives.
4. Be familiar with the range of malign moves and be prepared to counter them, even if the perpetrator is your own boss.

Design Every Group Discussion

5. Break large discussion groups (including player cells) into subgroups of no more than 8 people. Put one person in charge of each subgroup and hold them accountable for the subgroup's product.
6. Use the standard three stage normative process for any discussion.
7. Recruit players based on opinion diversity and decentralization, design the three stage normative processes around opinion independence, and if aggregating opinions within or between discussion subgroups use a narrative based advantages & disadvantages matrix
8. Establish the ethical standards or norms that are relevant to the wargame, and remind everyone of these at the start of each discussion.

Finally

9. Simply reading hobby game rules and playing hobby games is wildly insufficient.
10. It is not enough to be good at implementing the bureaucratic wargaming processes of the organization to which you belong.
11. Go back and take 1, 2, 3, and 4 seriously!
12. Identify malign moves by designing and playing "outer games" inside your organization between red (unethical) and blue (ethical) stakeholders, where Blue's objective is to support the best possible decision for national security and Red's objective is to support a decision preferred by Red irrespective of whether that is best for national security.

References

Bailey, R. (2021). How Much Scientific Research Is Actually Fraudulent? It may be more than you think. Reason, 7.9.2021 https://reason.com/2021/07/09/how-much-scientific-research-is-actually-fraudulent/

Batteux, E., Ferguson, E. & Tunney, R. (2017). "Risk Preferences in Surrogate Decision Making", *Experimental Psychology*, 64(4), 290-297, https://doi.org/10.1027/1618-3169/a000371

Cooney, G., Mastroianni, A., Abi-Esber, N., & Brooks, A. (2020). "The many minds problem: disclosure in dyadic versus group conversation", *Current Opinion in Psychology*, Volume 31, February, Pages 22-27, https://doi.org/10.1016/j.copsyc.2019.06.032

Dodoiu, G., Leenders, R. & Dijk, H. (2017). "A meta-analysis of whether groups make more risky or more cautious decisions than individuals", *Academy of Management Proceedings*, Vol 2016, No 1, 30 November, https://doi.org/10.5465/ambpp.2016.194

Dorner, D. (1996). "The Logic of Failure: Recognizing and Avoiding Error in Complex Environments", *Addison-Wesley*.

Downes-Martin, S. (2013). "Adjudication: The Diabolus in Machina of War Gaming", *Naval War College Review*, Vol 66, No. 3, pp 67 – 80, digital-commons.usnwc.edu/nwc-review/vol66/iss3/6/

Downes-Martin, S. (2014). "Your Boss, Players and Sponsor: The Three Witches of Wargaming," Naval War College Review Vol 67, No 1, pp 31-40, digital-commons.usnwc.edu/nwc-review/vol67/iss1/5/

Downes-Martin, S. (2016). "Wargaming to Deceive the Sponsor: Why and How?" *Connections UK Wargaming Conference*, London UK Sep, www.professionalwargaming.co.uk/WargamingToDeceivePaper.pdf

Downes-Martin, S. et al. (2017). "Validity and Utility of Wargaming", *Working Group 2 report MORS Wargaming Special Meeting*, Washington DC Oct, paxsims.wordpress.com/2017/12/09/mors-validity-and-utility-of-wargaming/

Downes-Martin, S. (2019). "Malign Wargaming I: Group Dynamics in Wargames and How to Exploit Them", *Connections North Wargaming Conference*, McGill University Feb. paxsims.files.wordpress.com/2019/02/group-dynamics-in-wargames-sdm-20190216.pdf

Edelson, M., Polania, R., Ruff, C., Fehr, E. & Hare, T. (2018). "Computational and neurobiological foundations of leadership decisions", *Science*, 03 Aug, Vol. 361, Issue 6401, https://www.zora.uzh.ch/id/eprint/153022/19/Computional_Edelson_akz.pdf

Fay, N., Garrod, S. & , Carletta, J. (2000). "Group Discussion as Interactive Dialogue or as Serial Monologue: The Influence of Group Size", *Psychological Science*, November 1, https://doi.org/10.1111/1467-9280.00292

Gladwell, M. (2009). "Cocksure: Banks, Battles, and the Psychology of Overconfidence," *New Yorker*, Jul. www.newyorker.com/magazine/2009/07/27/cocksure

Goldman, S. (2021). "The problem with academic funding, in 1 ironic scandal", *The Week*. Aug 23. https://theweek.com/talking-points/1004044/why-scholars-of-cheating-faked-their-own-results

Goodstein, D. (2010). "On Fact and Fraud: Cautionary Tales from the Front Lines of Science", *Princeton University Press*

Greenfieldboyce, N. (2018). "What makes a leader", *NPR* August 2, https://www.npr.org/sections/health-shots/2018/08/02/634639437/what-makes-a-leader

Heuer, R. (1999). "Psychology of Intelligence Analysis", *CIA*, https://www.cia.gov/resources/csi/books-monographs/psychology-of-intelligence-analysis-2

Karpowitz, C., Mendelberg, T. (2014). The Silent Sex. *Princeton University Press.*

Kelly J., Sadeghieh T., & Adeli K. (2014). "Peer Review in Scientific Publications: Benefits, Critiques, & A Survival Guide". *EJIFCC*. 2014;25(3):227-243. Oct 24. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4975196/

Langer, E. (1975). "The Illusion of Control", *Journal of Personality and Social Psychology* 32, no. 2 (August), pp. 311–28. https://content.apa.org/record/1977-03333-001

Lehrer, J. (2012). "Groupthink: The brainstorming myth", *The New Yorker*, January 20, https://www.newyorker.com/magazine/2012/01/30/groupthink

Littman. Michael L. (2021). "Collusion Rings Threaten the Integrity of Computer Science Research". *Communications of the ACM*, Vol. 64 No. 6, pp 43-44, June, https://dl.acm.org/doi/pdf/10.1145/3429776

Kocher, M., Schudy, S. & Spantig, L. (2016). "I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups", *University of Munich Discussion Paper* 2016-8, https://doi.org/10.5282/ubm/epub.28966

Longley-Brown, G. (2019). "Successful Professional Wargames: A Practitioner's Handbook", *History of Wargaming Project*, www.wargaming.co/professional/details/professionalhandbook.htm

Lorenz, J., Rauhut, H., Schweitzer, F. & Helbing, D. (2011). "How social influence can undermine the wisdom of crowd effect", *Proceedings of the National Academy of Sciences*, May, https://www.pnas.org/content/pnas/108/22/9020.full.pdf

Mazar, N., Amir, O. & Ariely, D. (2008). "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance", *Journal of Marketing Research*: December 2008, Vol. 45, No. 6, pp. 633-644, https://journals.sagepub.com/doi/pdf/10.1509/jmkr.45.6.633

Mullen, B., Johnson, C. & Salas, E. (1991). "Productivity Loss in Brainstorming Groups: A Meta-Analytic Integration", *Journal of Basic and Applied Social Psychology*, Volume 12, 1991 - Issue 1, pp. 3-23, 7 June, https://doi.org/10.1207/s15324834basp1201_1

Nijstad, B., Stroebe, W. & Lodesijkx, H. (2006). "The illusion of group productivity: A reduction of failures explanation", *European Journal of Social Psychology*, 36, 31-48, https://doi.org/10.1002/ejsp.295

Shermer, M. (2010). "When Scientists Sin," *Scientific American* 303, no. 1 (July), p. 34, www.scientificamerican.com/

Surowiecki, J. (2004). "The Wisdom of Crowds", *Anchor Books*

Tetlock, P. (2005). "Expert Political Judgment: How Good Is It? How Can We Know?", P*rinceton University Press*

Tetlock, P. & Gardner, D. (2015). "Superforecasting: The Art and Science of Prediction", *Crown Publishing*

Thaler, R., Tversky, A., Kahneman, D. & Schwartz, A. (1997). "The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test", *The Quarterly Journal of Economics* Vol. 112, No. 2, May, pp. 647-661, https://doi.org/10.1162/003355397555226

Uri, S., Nelson, L., Simmons, J. (2021). "Evidence of Fraud in an Influential Field Experiment About Dishonesty", *Data Colada*. August 17. https://datacolada.org/98

Wells, H. G. (1913). "Little Wars: With an Appendix on Kriegspiel", http://www.gutenberg.org/files/3691/3691-h/3691-h.htm

Wills, G. (1994). "Certain Trumpets: The Nature of Leadership", *Simon & Schuster*

Wong, Y., Bae, S., Bartels, E. & Smith B. (2019). "Next-Generation Wargaming for the U.S. Marine Corps", *RAND*. www.rand.org/pubs/research_reports/RR2227.html

About the Author

Dr Stephen Downes-Martin is a Research Fellow at the US Naval War College and an independent scholar analysing the use and abuse of decision support tools at the strategic, operational and tactical levels of warfare and business. His current research focus is how to manipulate decision support tools such as wargaming to deceive decision makers, how decision makers deceive themselves and allow themselves to be deceived, how to detect such manipulation and how to protect decision makers from them. He works with and for a wide variety of government, military, aerospace, academic and commercial organizations in the US and internationally. His formal education includes a PhD in Relativistic Quantum Field Theory from King's College London University, a Masters with Distinction in National Security and Strategic Studies from the US Naval War College, and a Master of Advanced Studies in Mathematics from Cambridge University. His full bio is at: https://sites.google.com/site/stephendownesmartin/

Unethical Wargaming: Let us be Incompetent![33]

Unethical Wargaming Working Group
Connections US 2021 Wargaming Conference

William F. Owen (MRES)
Editor and Co-founder of Military Strategy Magazine
william@infinityjournal.com

Abstract

This essay argues that bad professional wargaming is one and the same as unethical wargaming. Professional Wargaming is a safety critical activity, so the pursuit of insights must be pursued over and above those ideas commonly associated with games. Thus the professional community must be a lot more rigorous and critical of gaming methods and designs, and of their ability to provide insights.

Introduction

The premise of what follows is simply that the most significant unethical practice in professional Wargaming is both a lack of competence and professional comprehension.

It is extremely important to understand that this is about professional Wargaming, with the two operative parts of that description being "professional" and "war." Games with no military content are games, not wargames. Professional Wargaming is done by and with military and civilian professionals to produce insights to inform decisions usually associated with training, experimentation and force development. It can reasonably be suggested that much of what follows might be read across as "gaming," but the negative consequences would be marginal. This makes a distinction between professional wargamers and professional gamers.

Professional Wargaming is very distinct from gaming or hobby wargaming which we will now call Hobby Games. Hobby Games are produced to be fun as well as having social and intellectual benefits to those that play them. Generally, they are not created with the intention of being a model or approximation of reality that can be used to inform professional insights, despite what

[33] The contents of this paper represent the opinion solely of the author.

many game designers claim. The idea that you can learn about the conduct of real operations solely from the perspective of Wargaming is deeply contestable, to say the least.

Hobby Games and Hobby Gamers have a limited place in the professional arena, but important exceptions do exist. Hobby games cover a large number of formats, from desktop computer simulation to boxed and figure-based tabletop games. Hobby games can and have produced useful professional insights, but those insights usually require a vast amount of qualification and need to be validated by other professional methods to be of any use. No one who really understands Wargaming can be totally dismissive of hobby games or hobby gamers because both hobby and professional Wargaming can trace their origins back to similar events in time and space, be that the advent of Kriegspiel or something similar. The critical point for this thesis is those hobby wargames, and hobby wargamers are a different breed and not a separate species from the professionals. There are no ethical problems in Hobby Wargames, mostly because they are part of the entertainment industry and correctly understood, should not impact the professional decision making, which is where the ethical implications lie. That is not to say that Professional Wargaming is free from the negative impact of the entertainment factor or even entertainers posing as professionals, especially when hobby wargames are touted as having a professional application or are stated as being able to produce valid professional insights.

Again, that is not to say that some modelling and simulation tools within hobby/computer wargaming do not have useful or even outstanding professional applications. They do, but most require a good deal of modification to be usefully employed, and that requires professional input. Thus, anyone who wants to promote a hobby wargame to a professional community needs to understand the ethics of doing so. In professional Wargaming, "being wrong" comes with a cost to someone.

The distinction between hobby and professional should be irrelevant because "Wargaming" is a tool, not an identity; thus, it should be in the application of that tool that the distinction lies. Those that see themselves as "professional wargamers" should be wary of such qualifications. No one comes to Wargaming out of the blue. Most people professionally involved in Wargaming have another professional qualification or background that supersedes that of "wargamer." Army Officer/Soldier, Mathematician, Analyst or Engineer, for example. Valid exceptions may exist, but the wider point here is professional Wargaming is about the quality of the insights they produce, not about the egos, reputations or even profits of those involved.

Wargaming is or should be a "show by doing" activity as in it is a practical skill, much like carpentry. Today, there is a minor industry organising wargaming conferences, talking and writing about Wargaming, researching the history, advocating for Wargaming, or even "being a

wargamer", and while valid and useful to varying degrees, none of the things does what professional Wargaming does, as in that which contributes to the apprehending of and deep understanding as to the how and why of military operations. Wargaming is a tool applied for a broader purpose that requires additional skills and knowledge. It is with those wider skills and experience where the real advantage lies.

Ethics

Thus "Wargaming" does not have an ethics problem. Wargamers have an ethics problem in that a lot of them do not know what they are doing and are producing unsafe insights. If you were a bad doctor or a bad airline pilot, people would develop a somewhat negative opinion of you based on the fact you were prepared to trade your ego and identity for their safety or lives. Is Wargaming that different? To paraphrase Mark Baum, the problem with fraud is not that it is not nice or mean, but that it never works. The most significant unethical behaviour in Wargaming is incompetence and/or allowing incompetence to thrive. Bad wargamers should be challenged in the same way you would challenge a bad welder, engineer, Lawyer or Medical Doctor.

Competence matters. Warships and/or armoured brigades are commanded by men who know how to do it for real, so they are not "playing a game." In a professional wargame, they should be giving the same orders and instructions as they would in reality, and they should be getting the same inputs to support their decisions, or those inputs are modelled in a way they and everyone else can understand. This is a fairly obvious yet frequently transgressed fundamental of Wargaming. With military professionals, it is fairly definable because experience counts and can be demonstrated. Time, money and real-world conditions often mean that enemy armoured brigades or warships are not commanded by people with the practical experience that matters, but that does not mean this condition should be accepted. Practical experience most often highlights where the game system is falling short, so to advocate against this is to drift into an Ender's Game fantasy where real fighter pilots can be bested by real McDonald's employees, but the daily rate for fast-food employees is less than that of fighter pilots.

This is important because Wargaming should not be about make-believe, pretending, or even role-playing, mostly if the enemy commander's role is being filled by a 27-year-old PhD candidate with no military experience, but clearly, exceptions do exist. Training would be an obvious example. It may be that the person concerned needs to understand something about command and operations, which can be achieved by exposing him to those decisions via Wargaming. Some could obviously suggest that the same 27-year-olds lack of military experience means they will be an opponent that acts unpredictably and imaginatively, but that

usually also risks working in an unrealistic and unlikely way. That said, training audiences can use those skills and in-game mechanics to provide real training value because the real-world experience is largely irrelevant or can be input by those monitoring the process.

It is undoubtedly true that personal growth, understanding, and real insights can occur as a result of being exposed to Wargaming or the discussions inherent to their conduct, but if you cannot tell the difference between a bad game designed and conducted by the less skilled from something worthwhile, all that growth, understanding and insight may be worthless. The often heard defence of wargaming involvement that "it made me think" may actually be the provision of false facts and misleading evidence.

In and of itself, a military background is not a qualification in professional Wargaming, but military understanding and knowledge is non-discretionary and is clearly required. This extends beyond the "war" in "wargaming." Hobby gamers just have to understand how to play the game and win. Professional Wargamers have to understand the reality of what they are doing in terms of what it means in real-world practice. Star Ship Captains and Level 18 Elf Spell Casters don't need the knowledge to derive from actual experience. In real wargames, someone has to know how long it takes to prepare a 40-meter single-span 4-lane road bridge for demolition in the real world. This does not mean it is legitimate to quiz game participants about their knowledge or experience to show gaps in their learning. Even a highly experienced commander may not know how long it takes to prepare the bridge for demolition, but they would know what questions to ask, and that is where the critical difference lies.

People

Wargaming is about people making decisions and the consequences of those decisions compared to someone else in an imagined lethal competition where someone loses and someone wins. Generally speaking, Professional wargames require professionals to make those decisions, and those decisions should be associated with breaking the will of his opponent to continue in combat. Wargaming is thus competitive and thus distinct from other formats of games which some conflate with Wargaming. With Wargaming, someone loses and suffers the negative effects of it, even if it is just a bruised ego or a round of drinks. If winning and losing does not matter, then nor does the good or bad decisions made during the game. Unless you understand what is causing someone to win and what is causing someone to lose, your Wargame is not doing anything useful either for training, the experiment or for developing the force. Wargames have to be real competitions. If they are not, then you will either teach the wrong lesson or develop the wrong force. How is being knowingly wrong ethical?

Based on some standard practices, it is entirely fair to ask if some people in professional Wargaming really understand the process they are involved in as regards decision making? For example, does any professional make decisions based on terrain as represented by a Hex-map? The answer is never, and it can only mislead, yet hex maps are not unheard of in professional wargames, and yet people have claimed professional insights from hex-map games. Does any commander base their decisions on their experience of a dice-role modifier (DRM)? No, they do not, nor have they based their understanding of manoeuvre on what the "rules" tell them. If real-life military experience and knowledge is not the primary enabler for someone to compete in a professional wargame, then there is a problem. If your Wargame does not enable the professional skill set your participants have, then how ethical is it?

Thus, there is a clear difference in perception and utility if your Wargame is pitting two players against each other in "a game," with rules, dice and gameplay mechanics, and two commanders and their staffs in a fight that reflects their training and experience.

Even dice are actually more problematic than many suppose. In hobby gaming, they are used to represent chance and uncertainty. The same reasoning and justification are constantly used in professional Wargaming, but usually, dice should be used in a far more discerning way than many suppose. Dice work well when the outcome is statistical or based on actual probability. The dispersion of long-range gunfire is a good example of where dice work well because you are using probability to simulate the outcome of a single measurable event or action, such as a probability of hit or probability of kill. In sharp contrast, if you are using dice rolls to judge the outcome of a company attack, then something is very wrong for the simple reason that a bad-dice roll contains no insights. Nor does a good one. If a platoon attack is being resolved by a dice roll, then there is such a vast amount of abstraction inherent to the game system that any insight would be hard to come by. The biggest problem is that chance, uncertainty and/or "friction" is inherent to everything in warfare. The vast majority of wargame systems tend to use dice exclusively in relation to combating resolution tables, yet chance and uncertainty affect every aspect of operations, including movement, signals, logistics, planning and much more. How often are dice rolls routinely used for movement? If dice are making better insights less likely but just making for "better gameplay", then how ethical are dice? Many gamers fail to understand that the Clausewitzian analogy or friction distinguishes between those things causing friction and interaction with the enemy. They are not the same.

It is noteworthy that even the most cursory Google search will indicate a body of writing which is mostly advocacy for Wargaming or the reputational writing of wargame advocates with the opposing view of wargame sceptics being almost entirely absent. That opposing view would be that some professional wargaming is being done by the wrong people, asking the wrong questions for the wrong reason and doing it badly. For those that find this an intolerable and

evidence-free slur upon the sacred art of Wargaming, then you are probably part of the problem.

Has Wargaming proven useful historically? Yes, it has. However, have wargames produced utterly misleading insights which have been dressed up as evidence? Yes, they have. Simply put, there are more Millennium Challenge 2002 failures than there are the successes of the Western Approaches Tactical Unit (WATU). Those two examples are chosen deliberately because both examples are widely and commonly misunderstood.

Millennium Challenge 02 (MC02) failed because it asked the wrong questions in the wrong way and arguably may have used the wrong people to ask it. WATU worked because a highly competent Royal Navy Officer used Wargaming to address a definable time and space problem, which he knew a wargame type process could produce insights for. What WATU did wasn't ground-breaking or even that innovative. The officer concerned was a wargamer before he formed WATU, and he had been introduced to Wargaming by the Royal Navy. Wargaming was well understood by most navies well before 1942. WATU was simply doing what most knew certain formats of Naval Wargames could do and were proven to do. WATU was successful because it addressed a very discrete limited problem with a defined objective. Of course, WATUs real success was nothing to do with Wargaming. Its success was that it helped sink U-boats and kill their crews.

The prominence of both MC02 and WATU in Wargamer mythology is also symptomatic of many not understanding the organisational risks associated with professional wargames. Those enthusiastic or evangelical about Wargaming tend to focus on the "game" as the event or the experience. They tend not to understand wargames as being a tool that relates to a wider decision-making process where they should be highlighting the very real limits of wargames. Wargaming has real limits, and ethical wargamers should probably talk more about what Wargaming cannot do rather than what they think it can. The "think it can" qualification is critical because lots of wargame advocates have a vastly inflated opinion of what Wargaming might achieve.

Hobby Wargamers are often prone to suggesting that wargames can refight historical battles or tell you the outcome of future ones, or even ones that never occurred. Such claims are usually and most often fundamentally flawed, but that does not make the claims any less likely or less believable in the eyes of some convinced of a wargame's ability to unearth the truth.

The problem here is the "so what" of Wargaming. The value in professional Wargaming is understanding how and why things happened, which may (or may not) have relevance to the real world. Thus, professional Wargames exist as or come to have value, as reports. Reports are what inform decisions, not the wargames themselves. Sometimes the reality of what happened

in the Wargame is not that which is reflected in the report. It may be substantially different. The truth may lie in the eye of the beholder. Why would that be the case?

Suppose a game report strongly indicates that new main battle tanks are not worth the investment compared to the effectiveness of other vehicle designs also tested using the same process. This insight is not going to be welcomed by those who work in a tank-centric organisation convinced of a tanks merit and effectiveness, so a report may not suggest that those insights are as certain as others aware of what the outcome might suggest based on the same game. This is because the organisational risk is a real thing in the world of professional Wargaming. The right answer can be the wrong solution. In the real world, if a wargame sponsor asks "which tank", then telling the sponsor the solution indicated is "no tank", it means you don't understand Wargaming from an organisational perspective. You're just a gamer.

However, if the problematic game report was the result of one game, executed once, then it can be safely ignored. Conversely, if that report was the result of many games, run multiple time, using multiple wargame methods, across a wide range of conditions using other means to validate or better understand the insights that were most commonly apparent, then it would be foolish to ignore that report, and it would be harder to obscure or disguise the insights it generated. Wargames cannot give you YES/NO answers, and one Wargame done once will tell you nothing useful, but a long, expensive and rigorous series of wargames will indicate where the trade-offs may lie and/or confirm or deny their relevance/existence of the issues under examination. Ethically, anyone promoting or advocating for wargames as a method to support decisions is honour bound to point this out.

Conclusion

Excellent and useful wargames are long, expensive processes that require a great deal of hard work by a significant number of people and are probably not very exciting and even less "fun." This is not a popular view or one many of those who want to use Wargaming professional want to hear, but there is a simple fact that rigour and integrity are never easy. Good Wargaming is not a quick and easy option. It requires lots of time and not a small number of people.

"Unethical wargaming" is really just lousy Wargaming poorly done by people who don't really understand Wargaming. The solution probably lies in educating those that use Wargaming as a tool as to what the bad practices look like and how those can manifest themselves in associated "red flags" in terms of those selling wargaming expertise or advice.

If you believe in professional Wargaming, then you think it might help save lives, time and money. These are all important. If you are content to do it badly or have not taken steps to test

your own understanding and competence in relation to Wargaming, you are risking lives, time and money, and sooner or later, you will be found out.

About the Author

William F. (Wilf) Owen joined the British Army in 1981 and served in both Regular and Territorial units. He has worked on security and defence advisory contracts in the Middle East, Far East and West Africa. He has also been the Editor of three periodicals focussing on Asian and Middle Eastern defence issues and is currently the co-founder and Editor of the Editor of Military Strategy Magazine (formerly Infinity Journal) a peer-review academic journal focussing on strategic theory and applications. He is also widely published in UK and US service journals. He has worked as a contractor for the British Army on a wide range of doctrine, command and force development issues.

The Layers of Ethics Surrounding Wargaming

How and Why Some Wargames Were "Unethical" in Japan[34]

Unethical Wargaming Working Group
Connections US 2021 Wargaming Conference

Hiroyasu Akutsu, Ph.D
National Institute for Defense Studies, Tokyo
akutsu-h@air.ocn.ne.jp

Abstract

In addressing the question of "how to use unethical practices to make your wargame say what you want it to say," this paper argues that one can do so by exploiting the various levels of ethics in wargames to justify his or her decisions and suggests that a general solution is to articulate from the beginning the "wargame ethics" that the participants are obliged to follow. The paper makes those points by employing an analytical model called "the Layers of Ethics Model" and briefly examining some representative cases in Japan.

Key Words: layers of ethics; Midway Naval Wargame; Total War Research Institute; Mitsuya Study; collective self-defense

Introduction

The question of "how to use unethical practices to make your wargame say what you want it to say" is hard to tackle. This is because it is hard to define what is unethical in wargaming. One can be both ethical and unethical depending on the environment in which the wargame is played. Instead of discussing as many definitions as possible of what unethical practice in wargaming means, I will briefly discuss the environments where unethical behavior could occur in wargaming, set an analytical framework to represent such an environment, set a historical context in Japan, and then analyze a few representative cases to show how and why they can be seen as unethical within the context relevant to the ongoing discussion. Finally, I will return to the question of "how to use unethical practices to make your wargame say what you want it to say" and propose a general solution.

[34] The contents of this paper represent the opinion solely of the author.

1. The Layers of Ethics and How Wargaming Becomes "Unethical"

First, to set up an analytical framework for this paper, I will argue that one becomes unethical if he or she behaves against the ethics that define or surround the wargaming situation. The ethics of/about a particular wargame may generally depend on the rules that have been explicitly or implicitly defined. Of course, such rules may come from broader and often tacit common beliefs in the society where the wargame is conducted. Those rules may also be against the social ethics if the wargame is originally designed to encourage any kind of red team conduct that goes against the social ethics. Deception or even using nuclear weapons is usually unethical in many human societies, but they are recognized as "accepted" wargame decisions. If red team thinking is unacceptable in a particular wargame as well as a society that surrounds the wargame, then red team thinking is unethical. There is a case in which wargaming was once unethical in a society, or wargaming in the military, in particular, was thought unethical in a country. In most Western societies and countries, wargaming is accepted, if not as popular as computer wargames, so wargaming per se is not necessarily unethical.

The above discussion suggests that the unethicality of wargames comes from the gap between ethics within and outside the wargames. If particular rules are only vaguely defined or if there is no clear code of conduct in a particular wargame, then one can intentionally exploit the vagueness and absence of rules to manipulate the wargame toward a particular direction and achieve one's own agenda, which is against the ethics that would have defined the rules more clearly otherwise. Then the other participants of the wargame would see this conduct as unethical if they naturally comply with the shared ethics that they believe also apply to the wargame. Thus, if one's behavior in a particular wargame is seen opposed to or at least inappropriate to the dominant ethics in the wargame, that behavior is unethical. If that behavior is seen as complying with the dominant ethics that defines the direction and even the rules of the wargame, then that behavior is ethical.

This brings me to a view that there may be overlapping ethics within and outside wargames and that unethicality occurs when those divergent ethics become frictional to one another. If we simply replace "ethics" with "agenda" in this picture, the whole picture resembles what I would call the inter-agenda conflict model which I believe was first introduced by Stephen Downes-Martin.[35] In particular, the layers of ethics model can be used to further analyze the dynamics between and among the "inner game" and the "outer game" in wargames.[36] This view

[35] Stephen Downes-Martin, "Your Boss, Players, and Sponsor: The Three Witches of War Gaming," *Naval War College Review*, Volume 67, Number 1, Winter 2014, Article 5, pp. 31-40.
[36] Stephen Downes-Martin, "Group Dynamics in Wargames and How to Exploit Them," a paper presented at Connections North 2019 Wargaming Conference, February 16, 2019.

might be called the ethics versus ethics model or the layers of ethics in wargaming. This view can be illustrated as in Figure 1 below.

Figure 1: Definitions

Social ethics:	Shared ethical codes in a society or nation.
Wargaming ethics:	Shared ethical codes in a wargaming society.
Wargame ethics:	Shared ethical codes in a specific wargame.
Designers' ethics:	Ethical codes shared between wargame designers.
Players' ethics:	Ethical codes shared between particular players.

In Figure 1, in a particular nation or society, certain norms define social ethics and they surround the other lower and smaller layers including the common ethics in a particular wargaming community. Those ethics define what is right and wrong about wargaming and all those involved in a particular wargame are expected to comply with them. Certain behavior in wargaming would be recognized as unethical if it conflicts with the upper layer ethics. Of course, if the ethics of all the layers are in agreement and compatible, then there should be no ethical problem.

It should be noted, however, that each layer of ethics may involve multiple internal "codes," whether implicit or explicit, that may be mutually frictional or contradictory in the first place. This situation can be illustrated in Figure 2 below and can be called dual ethics or dual ethical structure.[37] While some ethical code is shared in a particular wargame that is not necessarily compatible with the normal wargaming ethics but that is based on the higher social ethics. In the controversial Millennium Challenge 2002 wargame (MC02), for example, the red team behavior of Lt. Gen. Paul Van Riper might have been in compliance with the common ethics of the American wargaming community about the role of Red team, but it might also have been

[37] The focus of Figure 2 is the layers below "Social Ethics" in Figure 1. In Figure 2, A and B represent specific ethical codes. It should be noted that there may be more than two such codes.

"embarrassing" to the sponsor whose ethics was to save the official face of the military community and the unspoken objective which was for Blue to win no matter what.[38] From the sponsor's perspective, Lt. Gen. Van Riper's action might appear "unethical," or at least problematic. But the sponsor's decision to change the rules of engagement in the wargame and restart it after the General's action can also be considered as "unethical". This is because such a change was against the higher layer ethics of the wargaming community which prioritizes the players' innovative ideas over intervention by the Controller or the sponsors. It can also be argued that the General's criticism of the sponsor and exit from the wargame was unethical because it was against the highest social ethical code (including the higher military organizational code) that military personnel must obey their superiors' orders and support their decisions.

Figure 2

However, if the higher layer ethics of the wargaming community had another ethical code that the sponsor's agenda should override the lower layer of designers' and players' ethics, then it would be inappropriate to criticize the sponsor of MC02. This is an example of the problem stemming from the dual ethical structure of the wargaming community in a particular society. To make clear, it also could happen that while contiguous ethical layers are mutually frictional, the highest layer and the lowest layer are in agreement. The wargame players could utilize the highest social ethical codes to justify their decision or action against the sponsors' words and deeds in a particular wargame.

[38] Micha Zenko, "Millennium Challenge: The Real Story of a Corrupted Military Exercise and its Legacy," War on the Rocks, November 5, 2015 https://warontherocks.com/2015/11/millennium-challenge-the-real-story-of-a-corrupted-military-exercise-and-its-legacy/

Thus, the proposed layers of ethics model is vertically and horizontally interactive and dynamic. Of course, the real structure and process of ethical dynamics are more complicated and the model is only used as a framework for argument and analysis in this paper.

2. The Layers of Ethics Model and Historical Context: A Japanese Case

2-1 A Brief History of Wargaming in Japan after the Meiji Restoration

I now take the case of wargaming in Japan from the beginning of the Imperial Japanese Army (IJA) through World War II to the present era and put it in a brief historical context. Since the Meiji Restoration in 1868, Japan's national slogan or goal was to "modernize", rebuilding itself into a "rich nation with a strong army (富國強兵)." This means that there was a nation-wide consensus favoring a strong military.

To focus on the military aspect, the ultimate goal was, of course, to "win the war." Before that, however, Japan had to concentrate on enhancing its police force to deal with the various kinds of political elements of division within the nation. National unification was actually completed in 1877 when a major civil war (the Satsuma Rebellion) was finally terminated. Then Japan was able to concentrate on building a strong army, and in this very process Japan encountered western-style wargaming in the 1880's. The IJA first translated American Army wargame rules and other related materials into Japanese in the early 1880's but had not introduced wargaming as an educational program until the IJA War College invited Klemens Wilhelm Jacob Meckel from Germany and learned from him the German way of wargaming in 1885.

In the Imperial Japanese Navy (IJN), on the other hand, the Royal Navy (RN) wargame methods were first introduced when the well-known Lieutenant AKIYAMA Saneyuki sent from the US a full set of Fred Jane's "Naval War Game" to Captain YAMAO Tanin at the Imperial Japanese Naval War College (IJNWC) in 1889. YAMAO then began teaching both tactical (兵棊演習) and operational (圖上演習) wargaming at the college. AKIYAMA, upon returning from the US and joining the college, and the other teaching staff members, further developed naval wargaming methods for their naval education as well as preparation for the war with Russia from 1904 to 1905.

The IJA, on the other hand, did not seem to have a systemic wargaming teaching or operational program equivalent to that of the IJN. Due to the lack of historical evidence, it requires intellectual restraints on confident judgement, but this is presumably because the IJA tended to prioritize field exercises rather than wargames in which there was no actual mobilization and movement of troops. It only conducted occasional wargames to validate its operational plans and exercise the implementation of such plans. Documented examples would

include the Northern Manchuria Wargame in the 1930's to prepare for war with USSR in Manchuria and the Torago (虎號) Wargame to plan on operations in the Pacific and China in 1943.[39]

After the Japanese-Russo war, the IJNWC continued teaching both tactical and operational wargaming and conducted both types of wargames until the Pacific War. Historical records of the IJN wargames following the Japanese-Russo War are still sparse, but one of the well-known wargames was the ten day-long naval operations in Hawaii in 1941 which facilitated the IJN leadership's decision to go to war with the US in the Pacific.[40]

2-2 Controversial War-Time Wargames

From the perspective of unethical wargaming, the most controversial wargame among those known in the West is probably the Midway Naval Wargame conducted on the deck of Yamato in 1942.[41] There are several issues with this wargame. The first is the fact that the controller told the adjudicator, Rear Adm.Matome Ugaki, to intervene in a gamed engagement during which two Blue (Japanese) aircraft carriers were destroyed, specifically to reduce the number of rounds fired by Red (US) from nine to three and to reduce the number of destroyed Blue aircraft carriers to one. As a result, the Blue team defeated the Red team as planned, but the Blue suffered a week-long delay of operations, lack of fuel, and some destroyers were deadlocked in the wargame. A second issue is the possibility that the intervention by the Adjudicator might have been aimed to save the face of Adm. YAMAMOTO Isoroku, Commander-in-Chief of the Combined Fleet, who was firmly determined to carry out the plan despite any negative game results. The third issue is the effectiveness of the wargame measured against the actual outcome of the operations in which the IJN lost four aircraft carriers and more than 3,000 lives.

The above summary of the issues may seem simplistic, but it usefully highlights the point of the current discussion. Regarding the first issue, from the viewpoint of both wargaming ethics and wargame ethics in Figure 1, if both layers of ethics were to allow interventions by the Adjudicator, UGAKI's action was not unethical. However, if both layers of ethics had not allowed such an action, then it would have been unethical. The fact remains that there had been naval

[39] HATTORI Takushiro, On the Torago Wargame of 1943, July 21, 1949; and The IJA, A Wargame Study on Northern Manchuria, dates unknown. These documents are held at the National Institute for Defense Studies Military History Archive.

[40] The National Defense College (the current National Institute for Defense Studies), Hawai Sakusen (Naval Operations in Hawaii) (Tokyo: Asagumo Shinbun, 1967)
http://www.nids.mod.go.jp/military_history_search/SoshoView?kanno=010

[41] The National Defense College, Middowei Kaisen (Battle of Midway) (Tokyo: Asagumo Shinbun, 1971)
http://www.nids.mod.go.jp/military_history_search/SoshoView?kanno=043#

wargame rules in the IJN since the Meiji era and they did allow the Adjudicators' intervention.[42] However, available evidence is sparse in terms of what judging criteria UGAKI referred to when he modified the damage to the Blue done by the Red. Therefore, it is now almost impossible to determine the unethicality of the wargame. As for the second issue, given the national, social, and military ethics during the war-time era in Japan, Rear Adm. UGAKI's action should deserve credit because it saved the face of Adm. YAMAMOTO at that time. Regarding the third issue, however, it is not about the ethicality of the wargame per se, but the operational effectiveness of the wargame during an actual war. Therefore, it would be unfair to criticize UGAKI in hindsight for consequently allowing YAMAMOTO to push his own agenda through and causing a huge loss in the actual battle. In this kind of situation, the Commander-in-Chief should usually take the blame. It can be said that the results of the IJN failed to meet the social ethical code of that time which was to prioritize war-winning and that therefore the IJN, not the wargame per se, was unethical.

Another example of controversial wargames in this era was the one conducted at The Total War Research Institute (TWRI) (總力戦研究所) in 1941. The wargame became widely known to the world when it was first raised as a suspicious "conspiracy" war planning at the International Military Tribunal for the Far East in 1946. The wargame was the first-ever pol-mil strategic wargame at the cabinet-level in Japan.[43] The formal sponsor was the Cabinet Office, not the IJA, but the latter was politically more powerful at that time. There are two versions of interpretation of the wargame. One is the more popular version that can be termed the "impossibility theory" argued for by INOSE Naoki, a famous writer who later became Vice Governor of Tokyo.[44] According to this version, the players confirmed the impossibility of Japan's victory over the US, but the result was not welcomed by Lt. Gen. TOJO Hideki, the Minister of the Army, as "not confirmed," "almost just on paper," and "not proved until we actually fight." INOSE's book reveals many statements that he obtained by individual interviews and further journalistic research.

Within the framework of the layers of ethics model, the players suffered conflicts with both the sponsor (the Minister of the Army as part of the Cabinet) and the highest layer of social ethics. From the standpoint of the sponsor and the higher layer of society, the wargame was definitely unethical and therefore deserved "rejection."

[42] The naval wargame rules were updated at the IJN Naval College at least three times from the Meiji to Showa eras before the Pacific War. Those documents are held at the NIDS Military Archives.
[43] MORIMATSU Toshio, Souryokusen Kenkyusho (The Total War Research Institute) (Tokyo: Hakuteisha, 1983)
[44] INOSE Naoki, Showa 16-nen Natsu no Haisen (Defeated in War in the Summer of 1941) (Tokyo: Chuko Bunko, 1983)

The second version of interpretation of the wargame was suggested by a graduate of IJA Academy and research fellow at the National Institute for Defense Studies.[45] The theory can be called the "difficulty theory." This version reveals many difficulties and vulnerabilities that Japan would face in promoting total war operations in Asia and against the US and the UK based on the wargame records, but it remains objective about the lack of clear evidence on TOJO's rejection of the wargame and has not provided any judgement about the official response from the sponsor to the wargame. However, the author of this version suggests that TOJO and the IJA might have drawn some useful lessons from the wargame and tried to utilize them in future operations. If so, then the wargame was totally ethical in the sense that it had achieved its role as "clarification of problems" and "providing lessons to the sponsor."

3 Post-WWII Wargaming and Ethical Dynamics in Japan

Regarding post-WWII wargaming in Japan, evidence is extremely sparse. On one hand, just as in the western world, official defense, diplomatic, and security-related policy simulations or games have usually been conducted quietly in Japan. On the other hand, however, the war-time experience and the post-war unique political situations affected the way such simulations and games were conducted in the process of rebuilding defense capabilities since the establishment of the Self-Defense Forces (SDF). In other words, the kind of restraints on defense activities including wargaming and simulations are often quite different from those on defense activities in the West.[46]

Having said that, however, it is possible to provide a rough sketch of the historical flow of events concerning wargaming in Japan and to choose some notable wargame and policy simulation short of war that can be recorded in relation to unethical wargaming.

3-1 Changing Ethics in Wargaming in the Post-WWII Japan

The flow of ethics concerning wargaming in Japan after the WWII is sketched as follows in the form of hypothetical statements that reflect the basic evolution towards a more accepting attitude to wargaming and policy simulation in Japan:

After 1945 and during the Cold War:

- *Social ethic*s
  "War is terrible and should never be accepted, and the alliance with the US is NOT a military alliance but just political and economic cooperation, therefore the right of collective self-defense is not accepted and exercised by any means."

[45] MORIMATSU, Souryokusen Kenkyusho, 1983.
[46] Perhaps, Germany's post-WWII situation was similar to Japan's in terms of military-related restraints.

- *Wargaming ethics*
  "No wargame should be conducted forever, and the very word "wargaming or wargame" is a taboo."
- *Hidden wargaming ethics (1)*
  "Wargames should not be revealed to the public but should be conducted carefully in secret and in a highly limited environment."
- *Hidden wargaming ethics (2)*
  "Blue players MUST NOT break the tacit rule of not deciding to accept or exercise the right of collective self-defense even when such a realistic and pragmatic decision is necessary to make alliance cooperation with the Blue ally easier and more effective."
  (So it became more "innovative" for the players to make alliance cooperation easier and effective without exercising the right.)

After the 1993-1994 Nuclear Crisis on the Korean Peninsula:

- *Social ethics*
  "War is bad, but contingencies could happen anytime anywhere near or in Japan, so some "simulation" is needed to get better prepared against the rainy day."
- *Wargaming ethics*
  "Yes, simulations should be conducted whenever necessary, but these are not wargames as generally accepted in the west."
- *Hidden wargaming ethics*
  "The existing tacit rule about the right of collective self-defense should basically remain the same, but more flexible thinking may be necessary."

After the rise of China in the 2000s,
the Fukushima disaster in 2011, and
the official and partial acceptance of the right of and also exercising collective self-defense in 2014:

- *Social ethics*
  "Not just war but any contingencies could happen anytime anywhere and affect Japan, so more simulation and preparation are needed. The change in the existing interpretation that the Constitution prohibits exercising the right of collective self-defense is no longer realistic."
- *Wargaming ethics*
  "Any critical contingencies should be simulated and should be the themes of TTX and other forms of gaming. Not just realistic but more flexible thinking is needed.
- *Hidden wargaming ethics*: N.A.

After the 2016 US Presidential Election:

- *Social ethics*
  "The world is becoming even more uncertain, so more simulations are needed."

- *Wargaming ethics*
  "Wargaming should be studied more and Japan should not miss the bus when other parts of the world are more enthusiastically studying it."
- *Hidden wargaming ethics*: N.A.

3-2 SDF's Mitsuya Study and a Think Tank Wargame about Japan-Korea Naval Cooperation

There are some records of wargames that have raised ethical issues. One well-known case is the so-called Mitsuya Study, or Study on Three Arrows (Ground, Maritime and Air Services), which was an internal and Joint wargame in the Self-Defense Forces (SDF) to see how the three services could cooperate in contingencies on the Korean Peninsula (the 2nd Korean War). The wargame was conducted as part of a more comprehensive research on such contingencies in 1963. The misfortune of this wargame was that it was politicized at the Diet. As a result, the Director General of Japan's Defense Agency (now the Ministry of Defense) resigned and many other senior staff members took the responsibility, and wargaming became a real political "taboo." Considering the layers of the ethics model, the wargame was unethical in the highest social ethical layer that restricted SDF wargaming at that time.

Another episode is the one I was personally involved in.[47] In the late 1990's, there was a non-governmental research program between a Japanese think tank and a South Korean think tank. The two think tanks had several wargames to see how Japan and South Korea could cooperate in dealing with North Korea's provocations. One of those wargames involved a situation in which North Korean submarines infiltrated into Japanese territorial waters. South Korean naval ships chased them and approached Japanese territorial waters. Japan, then, had to decide if it should allow the ships to enter and if Japan should directly cooperate with them by exercising the right of collective self-defense. The objective of the sponsor from the Japanese side was for the main player, who played Prime Minister of Japan, to say, "No," and end the wargame there so that the problem of the right of collective self-defense could be discussed among the participants. However, against the sponsor's expectations, the main player initially said, "Yes," instead. The statement was reversed and the wargame continued as if the prime minister did not actually decide to exercise the right.

The above episode indicates that in this wargame, the main player was unethical because he was against the sponsor's ethics that reflected the social ethics of that time. From the players' ethical standpoint, the sponsor may have appeared unethical or even unfair because he was

[47] This section is the simplest summary of an actual wargame that I can show in this paper and lacks some other important details, but should suffice to highlight the core issue of ethicality in wargaming. For an "official" account of this wargame, see Tae-Hyo Kim, "A Simulation: Possibilities and Limits of ROK-Japanese Naval Cooperation," in eds., Sang-Woo Rhee and Tae-Hyo Kim, *Korea-Japan Security Relations: Prescriptive Studies* (Seoul, South Korea: Oruem Publishing House, 2000), pp. 239-258.

against the wargaming ethics that encourage "innovative" ideas. The sponsor, however, was totally ethical in light of the highest social ethical code in the 1990's which restricted the exercise of the right of collective self-defense.

Conclusion

In this paper, I first set up a proposed layered ethics model as an analytical framework to roughly distinguish unethical wargames from ethical ones in Japan by utilizing some historical records from the Meiji Restoration in 1868 to the present. This is an initial attempt to make a distinction between ethical and unethical in terms of wargaming, and is a work in progress. However, this paper was able to highlight at least how some past wargames and practices in wargaming were unethical in Japan by using the layers of the ethics mode.

The analysis in this paper suggests that one can utilize ambiguities in the multiple layers of ethics to manipulate the outcome of wargames. To do so effectively the unethical practitioner should know what is ethical, what is not, and what is ambiguous at each layer of the model in which the wargame occurs, and be prepared to deal with others' attempts to stop them by exploiting the ambiguities.

Given that social ethics, including international and national ones, cannot be changed artificially, it would be even more important to clarify in the beginning what ethics the sponsors, designers, players, and other participants of wargames are expected to comply with. Regarding wargaming ethics, as the wargaming community becomes more and more international and multicultural, some efforts will be necessary to establish common and shared ethical codes as well.

Acknowledgements

I thank Dr. Stephen Downes-Martin for organizing and leading the Unethical Wargaming Working Group as well as providing valuable comments on the earlier drafts.

About the Author

Hiroyasu Akutsu is Professor and Chief of the Policy Simulation Division at the National Institute for Defense Studies (NIDS) in Tokyo. His current studies include scenario planning, policy simulation/strategic gaming and operations research (OR). He was a visiting fellow at the Royal United Services Institute (RUSI) in London (2010-2011), a Japan Chair visiting scholar at the Centre for Strategic and International Studies (CSIS) in Washington D.C. (2014), and a visiting exchange scholar at UK MOD's Development, Concepts, and Doctrine Centre (DCDC) in Shrivenham (2018). He received his Bachelor and Master of Law from Keio University and his Ph.D. in political science and international relations from the Australian National University (ANU). <http://www.nids.mod.go.jp/english/research/profile/anzen/02-akutsu.html>

Unethical from the Beginning

Wargaming and the Culture of Winning in the former Warsaw Pact Countries[48]

Unethical Wargaming Working Group
Connections US 2021 Wargaming Conference

Zoltán Harangi-Tóth, MAJ
Faculty of Military Sciences and Officer Training
National University of Public Service, Budapest, Hungary
harangi-toth.zoltan@uni-nke.hu

Abstract

This article is written by an European wargamer coming from a culture that historically defines a lot of things differently than most of my fellow researchers are used to. I think the different points of view of different cultures might illuminate the topics in ways that are missing when examined from a more closed, monocultural perspective. The goal is to improve communications and the sharing of wargaming expertise and experience between wargamers from Central and Eastern Europe (CEE) and those from a Western European tradition. I also hope that the ongoing discussion will help the late comers from the CEE region to join the international community more easily since all participants engaged in such an activity benefit from it.

Historical Background

As Hungary's history was built on the western Christian traditions, it's fundamental ethics were also based on these since the Middle Ages.[49] However there were some periods of time,

[48] The contents of this paper represent the opinion solely of the author.
[49] BODA Mihály: The Hungarian Theory of Just War Based on the Idea of the Holy Crown. In: *Journal of Military Ethics*, 2020/4. http://real.mtak.hu/112918/ (Accessed 13. February 2021.)

when foreign powers' occupation completely changed the basic structures of society, and re-framed how ethics were handled. This feature is not unique, because during the 20th century in Central and Eastern Europe almost all countries suffered border changes, loss of territory, and changes in their ethnicity, not to mention the regime changes that followed the many wars. While this paper's topic is based on mostly Hungarian experiences, most of the mentioned features affect the whole region.

Wargaming became a part of the General Staff inventory after the 1866 summer campaign against Prussia, when the Austrian Northern Army lost badly at the Battle of Königgraetz. After the 1868 reforms the new officer corps was distanced from society and established a very strict professional ethical code. It was a rigid, heavily authoritarian and hierarchical one which included the elite General Staff officer system. From then until the end of the Second World War, Kriegspiel type wargames with seminars and field gaming (exercises) were an unchanging part of the General Staff officer training.[50]

However after the war Hungary and the whole CEE region remained occupied by the Red Army and following the 1948 communist takeover a whole new, deeply politicized, officer heavy military tradition was forcibly established. Soviet-type wargaming was more of an analytical game to train commanders and/or their staff in situational awareness, fast decision making and staff duty/tasks, different from the "western type" games we know and use today.[51] Since most former Warsaw Pact countries joined NATO in 1999 all the militaries started to run Course of Action (COA) analysis games, but even today not all professional wargame types are used throughout the region.

The Culture of Winning

During the second half of the last century, the CEE region was occupied by the Soviet Union for decades, and that occupation forcibly changed the whole society and still has an impact nowadays, through the older generations. During those times, there was no widespread recreational or hobby gaming like in the US or in the UK; wargaming was only for the military, and it stayed like that until recently.[52] This almost total lack of hobby gamers meant that most (professional) wargaming activities originated from the Armed Forces. While nowadays we have

[50] DANCZER Alfons (ed.): A mi hadseregünk. Az Osztrák-Magyar Monarchia népei fegyverben és a zászlók alatt. (Our Army. The Peoples of the Austrian-Hungarian Monarchy in Arms and under the Flag), *Franklin Társula*t, Budapest, 1889. 76-85.

[51] HERONYÁNYI József: A hadijátékok tervezésérõl és levezetésérõl. (About the planning and conduct of wargames). In: *Honvédelem*, 1987/11. 39–48.

[52] POGÁCSÁS Krisztián: Terepasztalos Parancsnoki Hadijáték, avagy a parancsnoki döntéshozatal és harcászati szemlélet fejlesztésének eszköze. (Sand Table Wargaming as a development tool for the commander's decision making) In: *Seregszemle,* 2013/2–3., 83–89. o.

a small and dedicated historical wargaming community, the military's traditional leading role still has an impact on today's professional wargaming culture.[53]

The Warsaw Pact era created an armed forces model, where only the winners were respected.[54] Although that might sound familiar or true for any military around the world, The Warsaw Pact culture needs some explanation since it still has an effect on how and why we play wargames. In my point of view, there exists a positive and negative winner's culture – in a positive one, the "the winner takes all" so better outcomes are rewarded, and winning is an achievement which motivates the participants to win. In a negative culture winning is the standard, and the "loser" (or simply everybody else but the best/winner) is somehow punished. Needless to say it affects morale and demotivates (ethical) competition through the fear of punishment. These two aspects (positive and negative) of military culture affected all levels of the Warsaw Pact armed forces, including the Hungarian People's Army, where young officers (now possibly senior leaders) were socialized according to these standards.

This negative kind of winning culture is still prevalent throughout the region – a prime example of this is the Polish Armed Forces' Zima (Winter)-20 exercise.[55] During the exercise held during January 2021 at the War Studies University at Warsaw, the Polish played a discovery type wargame, when they tested some of the possible Courses of Action in a large-scale conflict against a fictional Russian aggression. It was originally a classified event, but the results (serious Blue Force losses in a very short time) leaked to the public and created a heated debate, even including the competency of the country's military leaders. As professional wargamers we all know this is exactly what wargames are created for, to test under-developed or new concepts or COA's before taking any real life action. However the public reaction in Poland showed that the negative winner's culture still has an impact on the CEE region. Here we are, 30 years after the Warsaw Pact and the Soviet Union collapsed and losing (face), even during training, is still unacceptable.

Building on this I now highlight how this mentality in the region affects (military) wargaming, and how it creates an unethical atmosphere right from the beginning!

---

[53] HARANGI-TÓTH Zoltán: Nemzetközi hadijáték trendek (International Wargaming Trends). In: *Honvédségi Szemle* 2020/2. https://doi.org/10.35926%2FHSZ.2020.2.11 (Accessed: 13. February, 2021.)

[54] SZŰCS Péter: Katonai vezetői gondolkodás fejlődését befolyásoló tényezők a rendszerváltástól napjainkig (Factors Influencing the Development of Military Leadership Thinking from the Change of Regime to Date) in: *Hadtudományi Szemle*, 2020/2. https://doi.org/10.32563/hsz.2020.2.12 (Accessed: 12 February 2021.)

[55] Jedrzej GRAF: Interview with Thomasz Szatkowski (Poland's Ambassador to NATO) https://www.defence24.pl/obrona-polski-musi-sie-opierac-na-sprawdzonych-scenariuszach-a-nie-na-dogmatach?fbclid=IwAR0wiPbxmfEHte2QuqqbK-9N6rKD9RxsBhQXXuhFrLJDQE7L8RarcA9Gjsk (Accessed: 12 February 2021.)

Unethical from the Beginning

In my understanding, unethical wargaming is caused by unethical or unprofessional participant behaviour.[56] I will analyse the root causes of this in the CEE region (due to the above mentioned reasons mostly in the military domain). I will focus primarily on the connection between the set gaming environment/conditions and the participants actions/goals, and less on understanding different methods of how to do or counter the root causes. I assume that if the wargaming environment is built as an ethical one with good examples being set by leadership then that eliminates most in-game unethical acts, leaving only a few that are mostly caused by the participants' personal preferences, and each of the latter is a unique problem that requires a unique solution.

Normally, wargaming is a method that creates a creative and safe to fail environment for experimenting or learning.[57] In a military culture where only winning is respected and not-winning is punished, those two basic attributes are not present. In a military COA analysis or other type of wargame, the whole environment (multiple level leadership, superiors, etc.) is concerned about the success of the wargame, and that usually means the blue side has to win. If not (which is possible in a wargame) some loss of confidence and formal/informal punishment or other effect may take place. This creates an unbalanced, heavily biased environment that puts pressure on players on both sides where one has to win, the other has to gracefully lose. This latter creates a paradox, since if the red players are too good at it, it can "backfire" back to them which can create pressure to act unethically to secure the culturally expected outcome.

The Sum of All Fears -- The Leadership)

In the military, sponsors are often high ranking senior officers, whose authority gives them influence over the careers of more junior officers who usually make up the majority of players. Needless to say how this can have a seriously damaging effect on game play. Success becomes not just how well the COA or players performed, but also how well the game results support the personal careers of the sponsors and other senior stakeholders behind the game. It takes a professional event to a personal level, and puts a cognitive burden on all participants. The situation becomes worse if the sponsor has directly intervened in the design process (to favor some game mechanics or scenarios over others), is closely monitoring the unfolding game (to

[56] Since this research is from the point of view of professional wargaming activities I will use "unethical" and "unprofessional" interchangeably, although I recognize that there is a difference between the two.
[57] Ministry of Defence, Development, Concepts and Doctrine Center: *Wargaming Handbook.* LCSLS Headquarters and Operations Centre, Bicester, 2017. 5.

persuade players), or directly intervening in gameplay.[58] This kind of micromanagement is also a remnant of the Warsaw Pact habit of senior leaders reaching far down from their level of leadership to directly task subordinate commanders (almost like taking over command).[59]

It is worth noting that the sponsors, their representatives or other senior players of the game have the most tools and means to modify a game from the beginning to confirm their biases. This kind of involvement in the game process is unprofessional and unethical by any standard. Any bad examples of apparent acceptance of unethical behavior by senior leaders observed by junior officers, who are then under pressure to also behave unethically. This can create a reinforcing downward "unethicalness-spiral" that becomes institutionalized.

The Context of Unethical Gameplay

If the highest leadership is visibly acting unprofessionally or the whole (external) wargame environment is unethical, and the fear of the consequences for losing is stronger than the desire to win, then that puts a cognitive burden on the players and creates a moral/ethical dilemma. This can be interpreted as a cause of the internal (in-game) unethical behaviour – in my experience, in a highly professional environment none of the players tried to win by an ethically unacceptable way, and it was visible that they mostly tried to respect the rules of the game. The cause of this (the internal need to follow the good example or the fear of being caught) is irrelevant to this research's point of view.

Based on the above, if we want a more ethical wargame, first we need to create an environment that supports ethical (or denies unethical) gameplay – either by visible examples provided by senior officers, by the conditions set, or by the kind of objectives created. I usually do not support relying on "force" to ensure rules are followed. From the beginning the rules should be clear and easy to understand by everybody but unintentional mistakes always can happen.

Over-Competition, the Gateway of Unethical Gaming -- The Players

In the military, we deliberately grow people to be competitive – we measure their every step and their performance all the time. We know that in a real life fight there's rarely (if at all) a second chance to retry something but this is not the same as the above mentioned winner's

[58] Stephen DOWNES-MARTIN (2014) "Your Boss, Players, and Sponsor: The Three Witches of War Gaming," *Naval War College Review*: Vol. 67 : No. 1 , Article 5. Available at: https://digital-commons.usnwc.edu/nwc-review/vol67/iss1/5 (Accessed: 13. february 2021.)

[59] CZEGLÉDI Mihály*:* A küldetésorientált vezetési szemlélet kialakulását befolyásoló tényezők. (Factors with Influence on the Shaping of Mission-Oriented Leadership Approach) in: *Honvédségi Szemle* 2021/1. 75-88. https://doi.org/10.35926/HSZ.2021.1.6 (Accessed: 13. february 2021.)

cultures. So it is normal and expected by the organisation that people are competitive. The problem is some people are over-competitive, and their preference is to succeed in every possible situation they face,[60] sometimes not even considering the consequences of that win.[61]

In some wargames (especially in educational ones) winning or losing is simply one factor of many and usually not the most important one.[62] Over-competitive people are trying to reach their single goal, even when breaking the game, and causing it to fail achieving its primary objective. It's a personality trait, not always a deliberate behaviour,[63] but it can lead to degenerative gameplay or unethical in-game acts, due to the sheer need to overcome the opponent. In some rare cases, unprofessional and unethical can mean different things (for example causing unnecessary losses to someone's own forces to reach a certain wargame objective).

Bridge Over Troubled Water -- The Game Controllers

Game designers, adjudicators and controllers can also be unethical, but in a totally different way. The sponsor might need to confirm his/her bias, the players need to win, but the "bridge" connecting them, the people in charge of creating, running and evaluating the game (the wargame staff) must also be considered. When analysing this group it's important to define their specific personal or collective goals and the pressures on them to behave unethically. Currently the Hungarian military mostly does COA wargaming which is part of a highly structured, well written decision making process. This means the game designers and adjudicators/controllers seemingly have less motivation to "cheat" or favor one side. However, if we look into who they are outside of the wargaming environment, we likely find they are mid ranking officers, often from the same unit as the sponsors/game directors and the players. They usually know each other outside the game, which means they are possibly part of the organizational hierarchy and are not "independent" nor unbiased.

---

[60] Katie SALEN – Eric ZIMMERMANN: *Rules of Play.* MIT Press, Cambridge MA, 2003. 275-276. In their book the authors describe the five types of players. One of the aspects of their categorization is the interest in winning. From this paper's point of view, intense/extreme interest in winning is the closest equivalent to an over-competitive behaviour. A player described above is not necessarily a dedicated/intentional cheater, but (based on Salen and Zimmerman) more likely falls into the Unsportsmanlike category. The unethical environment simply makes it more plausible that the person will cheat during the gameplay.

[61] Christopher A. WEUVE et al.: Wargaming Pathologies. CNA Report, 2004. https://apps.dtic.mil/dtic/tr/fulltext/u2/a596774.pdf (Accessed: 12. February 2021.)

[62] Ibid. p. 15.

[63] CZIBOR Andrea: *Döntések társas dilemma helyzetekben: személyiségjellemzők és szituációs faktorok hatása.* (Decisions in Social Dilemmas: The Impact of Personal Traits and Situational Factors) PhD Thesis, University of Pécs, 2014. https://pea.lib.pte.hu/bitstream/handle/pea/15180/czibor-andrea-phd-2014.pdf?sequence=4&isAllowed=y (Accessed: 13. February 2021.)

Based on the above mentioned existence of a negative winning culture, it means they can (or must) support the view of their higher ranking leaders during scenario building or favour one side during adjudication to help the organization to succeed. This is not only unethical, it is also unprofessional in that it alters the wargame results and can cause real life problems during the execution of the tested COA in the real world. Moreover it can harm the trust in wargaming as an established tool for analysis in the long term. In my opinion, after the sponsor's involvement the biased gaming staff does the most harm to the acceptance of wargames (at least in Hungary).

The case is different if an outside organisation is conducting the game for the military, but this is less frequent than the previous case, since there are not many independent think-tanks or wargaming organizations in Hungary (nor in CEE). If another team runs a game, and they are financially involved, it makes them interested in not just the overall success of the game, but in the positive experience of the participants, since they are their "customers".

Conclusion

Today in Hungary we have a very small but international historical hobby gamer community,[64] an even smaller military heavy professional gaming community, and a few business wargamers. This article is based on both my current academic and my former staff position experiences, so it represents my own view on where we are now, and what the challenges we are facing. The un-acceptance of the Polish Zima-20 exercise results shows that these are not uniquely Hungarian problems, but the largest armed force of the Visegrád 4 (V4) group and the whole Central and Eastern European region is facing the same difficulties, with its roots going back to the Warsaw Pact heritage.

The question for us is how to create a more widespread, visible and influential professional community. As a first step this is inevitable to turn modern wargaming techniques and types that go beyond COA gaming and have already been proven useful elsewhere into accepted tools within the Hungarian Defence Forces and beyond. Success at this will remove the current dilemma every Hungarian military wargamer is facing: "Do I need to cheat for the Holy Cause, to deliver the expected good results in every single game, to give the participants a good experience just to get much needed future support in our quest? Or will doing that (rightfully) create a loss of trust in gaming as a whole?" What a dilemma indeed!

[64] A few large clubs are located mostly in Budapest or in the bigger cities and most hobby wargamers gather around those. For event organization and communication social media is the preferred platform. The membership of the largest Hungarian historical wargaming groups on Facebook can give a broad picture about the current state of the hobby. As of 2021 the largest historical miniature and board wargamer groups while overlap each other, all have about 2-300 members: these are the Hungarian Wargamers, 28mm Historical Wargaming Hungary and the Magyar Történelmi Hadijátékos Csoport (Group of Hungarian Historical Wargamers) groups.

Bibliography

BODA Mihály: The Hungarian Theory of Just War Based on the Idea of the Holy Crown. In: *Journal of Military Ethics,* 2020/4. http://real.mtak.hu/112918 (Accessed 13. February 2021.)

CZEGLÉDI Mihály*:* A küldetésorientált vezetési szemlélet kialakulását befolyásoló tényezők. (Factors with Influence on the Shaping of Mission-Oriented Leadership Approach) in: *Honvédségi Szemle* 2021/1. 75-88. https://doi.org/10.35926/HSZ.2021.1.6 (Accessed: 13. february 2021.)

CZIBOR Andrea: *Döntések társas dilemma helyzetekben: személyiségjellemzők és szituációs faktorok hatása.* (Decisions in Social Dilemmas: The Impact of Personal Traits and Situational Factors) PhD Thesis, University of Pécs, 2014. https://pea.lib.pte.hu/bitstream/handle/pea/15180/czibor-andrea-phd-2014.pdf?sequence=4&isAllowed=y (Accessed: 13. February 2021.)

DANCZER Alfons (ed.): *A mi hadseregünk. Az Osztrák-Magyar Monarchia népei fegyverben és a zászlók alatt.* (Our Army. The Peoples of the Austrian-Hungarian Monarchy in Arms and under the Flag), Franklin Társulat, Budapest, 1889.

DOWNES-MARTIN, Stephen (2014) "Your Boss, Players, and Sponsor: The Three Witches of War Gaming," *Naval War College Review*: Vol. 67 : No. 1 , Article 5. Available at: https://digital-commons.usnwc.edu/nwc-review/vol67/iss1/5 (Accessed: 13. february 2021.)

HERONYÁNYI József: A hadijátékok tervezésérõl és levezetésérõl. (About the planning and conduct of wargames). In: *Honvédelem*, 1987/11. 39–48.

HARANGI-TÓTH Zoltán: Nemzetközi hadijáték trendek (International Wargaming Trends). In: *Honvédségi Szemle* 2020/2. 113-124. https://doi.org/10.35926%2FHSZ.2020.2.11 (Accessed: 13. February, 2021.)

GRAF, Jedrzej: *Interview with Thomasz Szatkowski* (Polands Ambassador to NATO) https://www.defence24.pl/obrona-polski-musi-sie-opierac-na-sprawdzonych-scenariuszach-a-nie-na-dogmatach?fbclid=IwAR0wiPbxmfEHte2QuqqbK-9N6rKD9RxsBhQXXuhFrLJDQE7L8RarcA9Gjsk (Accessed: 12 February 2021.)

POGÁCSÁS Krisztián: Terepasztalos Parancsnoki Hadijáték, avagy a parancsnoki döntéshozatal és harcászati szemlélet fejlesztésének eszköze. (Sand Table Wargaming as a development tool for the commander's decision making) In: *Seregszemle*, 2013/2–3., 83–89. o.

Ministry of Defence, Development, Concepts and Doctrine Center: *Wargaming Handbook.* LCSLS Headquarters and Operations Centre, Bicester, 2017. 5.

SALEN, Katie – ZIMMERMANN, Eric: *Rules of Play.* MIT Press, Cambridge MA, 2003.

SZŰCS Péter: Katonai vezetői gondolkodás fejlődését befolyásoló tényezők a rendszerváltástól napjainkig (Factors Influencing the Development of Military Leadership Thinking from the Change of Regime to Date) in: *Hadtudományi Szemle*, 2020/2. 155-161. https://doi.org/10.32563/hsz.2020.2.12 (Accessed: 12 February 2021.)

WEUVE, Christopher A. et al.: *Wargaming Pathologies*. CNA Report, 2004. https://apps.dtic.mil/dtic/tr/fulltext/u2/a596774.pdf (Accessed: 12. February 2021.)

About the Author

Major (OF-3) Zoltán HARANGI-TÓTH is a Mechanized Infantry officer, who served in command and staff positions in infantry units in the Hungarian Defence Forces and in multiple foreign deployments. Currently he serves as a Lecturer of Military History and Leadership Studies at the Faculty of Military Sciences and Officer Training (National University of Public Service, NUPS) Budapest, Hungary. He holds a Military Leadership BSc/MSc and a History BA/MA, and he is a last year PhD candidate of the History of the Military Art at the Doctoral School of Military Sciences (NUPS, Budapest, Hungary).

Can AI Save Us from Unethical Wargames?[65]

Unethical Wargaming Working Group
Connections US 2021 Wargaming Conference

Robert Seater
MIT Lincoln Laboratory
robert.seater@ll.mit.edu

Abstract

Wargames are inherently vulnerable to unethical dynamics, both intentional and unintentional. Their nature as open ended exploration coupled with their reliance on abstraction are both strengths and weaknesses of the format. We will consider three ways in which wargames can run afoul of ethical boundaries, and how an impartial Artificial Intelligence (AI) that is myopic to the point of absurdity can help mitigate those risks. We consider game design as a form of human-machine teaming, in which the AI serves complementary roles to the human, rather than being subservient or acting as a substitute. The AI forces the human to understand the implications of their design decisions but makes no judgment on the desirability of those implications. We will describe some past successes, known limitations, and the potential future of using AI playtesters as part of a human-machine team for designing wargames.

1 Unethical Wargames

We will consider three broad categories of unethical behavior in wargames. There are other categories [1, 2, 3], but we will focus on these three major ones.

- **Unethical *Design* of Games**. Creating games to convey a message independent of truth. Presenting exposition as analysis, even accidentally, is unethical on the part of the creators.

[65] The contents of this paper represent the opinion solely of the author.

- **Unethical *Outcomes* of Games**. Creating games that produce a lesson that is difficult to properly translate into the real world. Misinterpretation and poor documentation of assumptions can provide justification for unethical behavior or policies in the real world.
- **Ethical *Dilemmas* in Games**. Creating games that force players to make moral and ethical tradeoffs and get their hands dirty. Modeling moral dilemmas can normalize unethical behavior rather than producing more informed field decisions.

We will consider each of these categories individually and then consider how an AI teammate can sometimes help resolve (or at least expose) those issues. But first, we will consider an underlying issue that is the reason why these risks are especially concerning for wargames -- the double-edged sword of abstraction.

1.1 Abstraction is a Necessary Evil

Consider game design as a modeling task. The designer takes a problem space (tactical or strategic), abstracts it (a lot or a little), and models it as a set of rules (loose or strict). In doing so, they remove detail and add higher-level rules to produce the dynamics that normally (in the full complexity of the real world) emerge from the removed detail. High-level abstract rules allow complex dynamics without cumbersome or unknowable detail -- one doesn't necessarily need to know *why* flanking, morale, or authority to use artillery improves mission outcomes to model *how* those dynamics affect higher-level decisions.

That power comes with dual risks--removing critical details or adding erroneous high level rules. Removing the details that make flanking work without adding a flanking bonus is unauthentic. So is adding an overly high reward for flanking, thus exaggerating the emergent effects. Approximation errors are acceptable when they are small or of low relevance to the primary analysis, and one job of the modeler is to judge which high level dynamics to model abstractly and how strong to make their influence. Here, we already see the three classes of unethical dynamics begin to emerge.

- **Designer**. A designer's abstraction might exaggerate a high-level rule to manipulate the game outcome consciously, unconsciously, or out of sloppiness.
- **Analyst**. An analyst might misunderstand the abstraction and take outcomes from a game designed to precisely model one dynamic but analyze it in terms of dynamics it imprecisely models.
- **Player**. An abstract representation of unethical consequences of a player's actions might be too impersonal to make a player care, thereby desensitizing them to those consequences.

These potential failings are all the result of poor abstraction choices, but *abstraction is inescapable*. A game without abstraction is quite literally the entire world, so we cannot simply add more detail whenever we have a judgment call to make. Even adding detail won't always work; very advanced combat simulators have trouble making suppressing fire emerge naturally, but nobody doubts that it is a real-world dynamic. Even when detailed models are available and practical, they are often undesirable. A detailed model is less transparent, impeding learning and analysis. A detailed model is also more brittle with respect to its assumptions and parameters, undermining its generalizability or requiring a much larger campaign of repeated runs. Detail, like abstraction, is a double-edged sword. The designer's job is to find the right level of abstraction, the right elements to abstract, and the right form for that abstraction [4, 5].

Design cannot be divorced from abstraction, and abstraction inherently involves applying one's own biases, expectations, and experience. Therefore, we should look at how to mitigate the problem, not how to avoid it. We want to allow well-intentioned designers to recognize unintended consequences of their designs, and we want external observers to have tools to catch ill-intentioned designers. As we will see in Section 2, an AI player can help provide both roles. Next, we will shift from theory to practice and consider three concrete examples of these violations, which will equip us to then return to discussing mitigations.

1.2 Unethical Design of Games

There was some news coverage years ago about a ballistic missile defense (BMD) game that was built during the Cold War to justify more spending on BMD.[66] The organizers brought in reporters to observe and participate. The game was calibrated such that most attacking missiles would be stopped ("We spend money well!"), but a few would always get through and take out recognizable cities ("We need more money!"). Players of any skill level would reliably stop 95% of the ICBMs but lose San Francisco. The reporters only played it once, so they didn't see that pattern; they went back and dutifully wrote up how great the BMD program was and how much it needed more funding. BMD likely *did* need more funding, but the game was employed in an unethical manner. It was exposition disguised as simulation.

Perhaps biased intent was involved in the BMD example, but it could also have been unconscious bias or unintentional poor design. After all, never attribute to malice that which can be explained by sheer incompetence. It is beyond our scope here to consider how to deal with malicious actors; we will focus on unintentional or unconscious skews, which can be equally damaging. A well-intentioned but sloppy engineer whose bridges keep collapsing bears

[66] I have not been able to track down a proper reference for this event. As such, it may be real, apocryphal, or part of my distorted memory. None-the-less, the situation demonstrates a concerning class of unethical behavior, regardless of its historical accuracy.

responsibility, but an engineer who follows best practice and gets unlucky is not (as) responsible. A well-intentioned designer producing a game that teaches erroneous lessons should be assessed similarly. Good intentions reduce, but never remove, the burden to mitigate risk.

A larger question for the wargaming community is *"What best practices are so foundational as to be unethical to ignore?".* This document will stick to the narrower question of *"How can we predict and mitigate unintended consequences of well-intended abstractions?*".

1.3 Unethical Interpretation of Game Outcomes

Because of the abstraction inherent in games, a strategy in the game may have lessons on the real world, but only if interpreted with care. Taking the results of a game too literally produces bad forecasts. Taking them too broadly produces vague platitudes. Somewhere in the middle lie useful lessons. Interpreting data from any experiment, exercise, or model requires care, and games are no exception [6].

In a recent course for high school students about AI and Serious Games [7], students modified a game about disease propagation, threw some reinforcement learning at it, and assessed the strategies that emerged -- good STEM education and whatnot. They built on models used in professional simulations and games, but with a zombie veneer added for entertainment value. They explored tradeoffs between resources spent on researching a cure, reducing transmission rates, evacuating healthy survivors, and threat elimination. The game included some intentionally absurd actions in order to demonstrate how quickly modern AI methods would rule them out. One was the option to firebomb a neighborhood and kill the entire population-- zombies, infected, and healthy alike. In the unmodified version of the game, the AI quickly learned to never take that action.

One student team modied disease models to describe societal differences between neighborhoods. There were wealthy gated communities, low sanitation slums, areas with religious restrictions, and so on. Those factors could affect vulnerability to the disease, mobility of the population, and which interventions would be rejected entirely. They expected to find that poor populations were more vulnerable than wealthy populations when public resources were overextended.

The AI they trained found a very efficient solution -- draw all the zombies and infected people into the gated community (using public messaging mechanics), hold them there until the end of the game (which was easy because it was already gated and reduced the chance of people wandering in or out), and then firebomb them on the last turn of the game. It calculated

that that solution saved the most lives and minimized the number of infected people in the city at the end of the game.

Although this game was played by AIs, the same outcomes could have come from human players seeking to maximize poorly designed victory conditions and poorly interpreting the in-game abstractions. In either case, how do you interpret that outcome? The answer isn't "firebomb gated communities" nor is it "completely ignore the results of this game". Rather, the lesson is something along the lines of "hold threat vectors in natural low-mobility regions to buy time to organize your response". There is a useful insight here, but it would be unethical to take the outcome literally given the abstraction. In retrospect, it might be obvious that modeling the sick as zombies will produce draconian results. However, even with that absurd stretch, there were interesting lessons uncovered, and even professional wargames are not immune to the issue of misinterpretation-- note the extensive controversy over interpretation of the Millennium Challenge wargame [8, 9].

This situation raises a deeper issue -- when a game looks realistic upon inspection, yet you get an unexpected outcome, when do you trust the result? When do you 'Fix' the game to remove the anomaly and when do you report the result as a novel finding? Is it even the game's job to make that distinction or merely to raise the issue to be evaluated by other methods? It requires judgment to recognize when an unexpected strategy in a wargame is clever and when it represents a design flaw. After all, if we threw out all unconventional solutions, we'd be losing a primary benefit of wargaming.

The broader question for the wargaming community is *"What processes and methods should designers and analysts employ to responsibly differentiate bugs from discoveries?"*. This document will focus on the narrower question of *"What tools will allow us to anticipate emergent consequences of modeling abstractions?".*

1.4 Unethical Dilemmas in Games

Another class of unethical dynamics has to do with the impact of the game on the players. Consider games that intentionally force the player to confront in-game ethical dilemmas. Perhaps resources must be allocated between search-and-rescue or evacuation, or perhaps a high value target is co-located with civilians. Such a game might be intended to understand how a command and control structure will make such decisions or simply to familiarize players with the tradeoffs involved so they will be prepared when facing similar decisions under the pressure of a real mission. Either way, the game is *intending* to produce more ethical behaviors through exposure and rehearsal.

A few years ago, we conducted a game primarily about coordinating multi-domain assets in an urban environment. Players had to combine effects from kinetic, social, cyber, and EW domains. One of the rewards for correctly correlating cross-domain information was the ability to target the enemy with indirect fire, and one of the penalties for being too hasty was the risk of civilian casualties. That sub-mechanic not only served to prevent degenerate solutions (just strike all possible targets), but it also served to establish civilian casualties as part of the decision space in an urban environment.

During the creation of that game, we saw a talk by Amnesty International at the Marine Corps University. The speaker emphasized that Amnesty's goal was not to eliminate all civilian casualties in war (a noble but unrealistic agenda), but rather to ensure that militaries do a cost-benefit analysis of military objectives and civilian losses (both lives and infrastructure). The talk was well received by our partners in the armed services, and we tried to incorporate that lesson in the game -- forcing players to do the calculus of civilian versus military priorities, even if we did not prescribe a correct or one-sided answer.

Our hope was that, through iterative play, players would build mental models of how their decisions struck different balances between mission expedience and civilian losses. Even though we, as game designers, could not prescribe the correct balance, we could help them build intuition for the tradeoffs they were making.

Hopefully it accomplished that goal, but perhaps it did the opposite. Perhaps by forcing players to repeatedly face messy ethical dilemmas, they simply got comfortable with collateral damage or numb to the civilian impacts. The abstract and schematized game mechanic of losing 'local political support points' might dehumanize the decision. By making them get their hands dirty in the game, is the game equipping them to be responsible or making them comfortable crossing moral lines?

The broader question for the wargaming community is *"How should designers assess the secondary effects of participating in a game, especially one that includes ethical dilemmas?"*. This document will focus on the narrower question of *"How can we anticipate the ethical dilemmas and compromises the players will face in a game?"*.

2 The Role of AI Teaming

AI can play a valuable role analogous to human playtesting, serving as a tool or teammate in the design process. The machine's role is to expose unintended consequences for the human team to assess and correct.

2.1 Example

Several years ago, we made an instructional game about emergency response management [10], illustrated in Figure 1. The game's goal is to teach high level emergency response decision makers how to allocate scarce personnel during a large-scale radiological disaster. Players allocate scarce resources between search and rescue, onsite medical treatment, evacuation coordination, and road clearing. Some of those tasks inefficiently serve survivors with urgent needs, whereas others efficiently serve survivors with less urgent needs. All of those tradeoffs must be adjusted over the course of the incident as needs and opportunities evolve. Players are evaluated both on lives saved and how well they maintained trust and order.

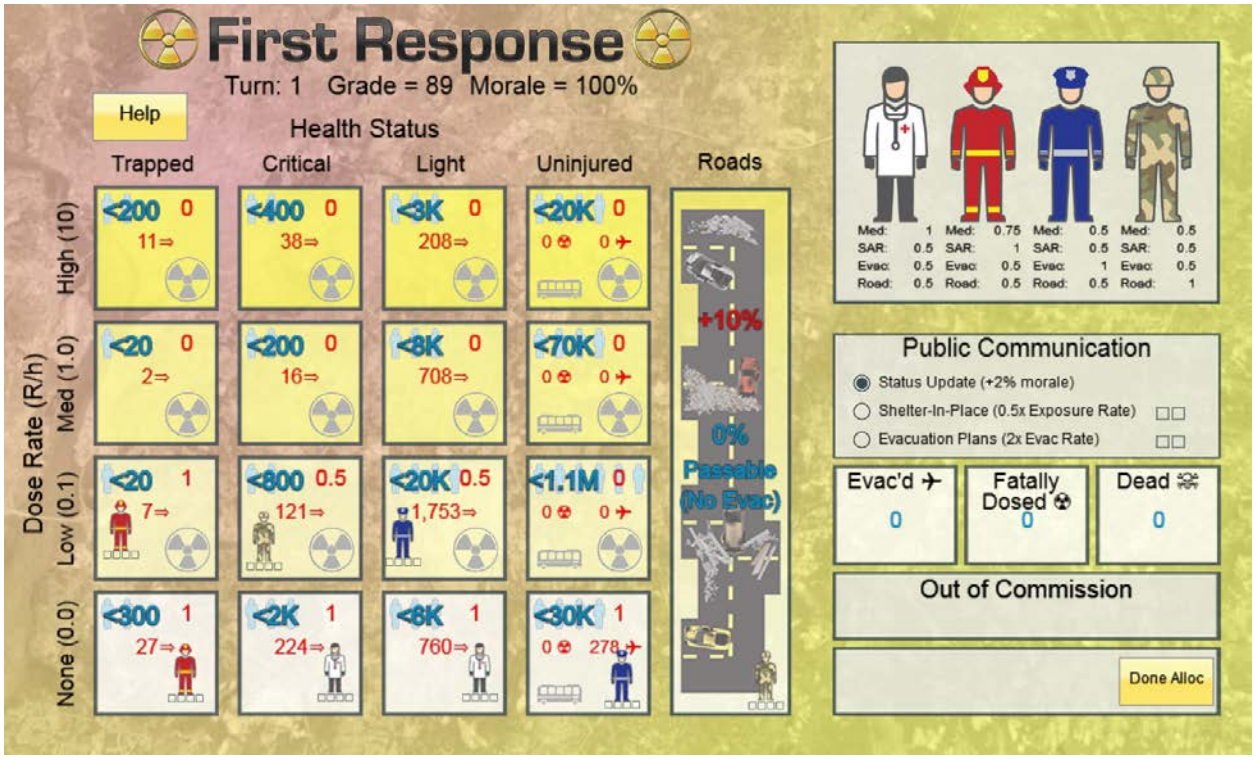


Figure 1: A game about allocating scarce resources in the wake of a terrorist nuclear attack. How do we know if the game is accidentally rewarding a degenerate strategy?

2.2 Approach

Subject matter experts in emergency management reviewed the game's assumptions, scoring, mechanics, and relevance to training needs. It passed their qualitative inspection and feedback was very positive, providing `face validity'.

However, with so many parameters and assumptions, we were unsure whether the game was teaching the right lessons. Maybe it was rewarding a degenerate solution that exploited the game abstraction, which would inadvertently teach students bad habits. The lessons it appears

to teach might not be the lessons it actually teaches, and we didn't have the resources for extensive human testing or longitudinal studies.

We trained an AI to play the game to help anticipate what lessons the game would implicitly teach to humans. We used a Monte-Carlo Tree Search to find near-optimal solutions, on the principle that the best machine strategy found was a good predictor of what greedy score-maximizing humans would eventually learn or be implicitly taught. We solved for local optima (not just global optima) to find a set of strategies a human would be likely to settle on after iterating with the game. We reviewed the discovered strategies with domain experts to determine whether they were strategies human players ought to be learning.

We solved for multiple solutions by iteratively solving the game, adding an additional term to the objective function each time to punish solutions that were similar to prior solutions. That approach found locally optimal solutions that appeared to mimic the strategies that human players converged on. Since our work, some newer AI-based methods have emerged for finding viable in-game strategies, even when no reward function is specified [11]. Regardless of the technical method used, it is important to search for sets of locally optimal solutions, not just globally optimal solutions, as that better replicates how humans explore an option space.

2.3 Findings

When we rst ran the AI, it found a completely degenerate solution that the designers, domain experts, and initial playtesters had missed. The game rewarded saving lives, evacuating people away from the dangerous region, and maintaining public trust of the response effort -- your score was an amalgam of all three of those factors. We apparently weighted the second one too high because the optimal in-game strategy was to ignore the people in need of help and to just evacuate the people who were already safe. A small overlooked error in the scoring rewards produced an unrealistic (and unethical) solution. When we re-balanced the scoring, it found solutions that appeared more credible. We reviewed that new set of solutions with the domain experts, who considered them credible and desirable to teach.

The game designers, domain experts, and AI analysis tool acted as a human-machine team [12] to create an effective game. The domain experts and designers defined the abstraction, the designers and AI predicted the consequences of that abstraction, the domain experts assessed those consequences, and the cycle iterated until the experts approved of the lessons being implicitly taught. This flow is depicted in Figure 2.

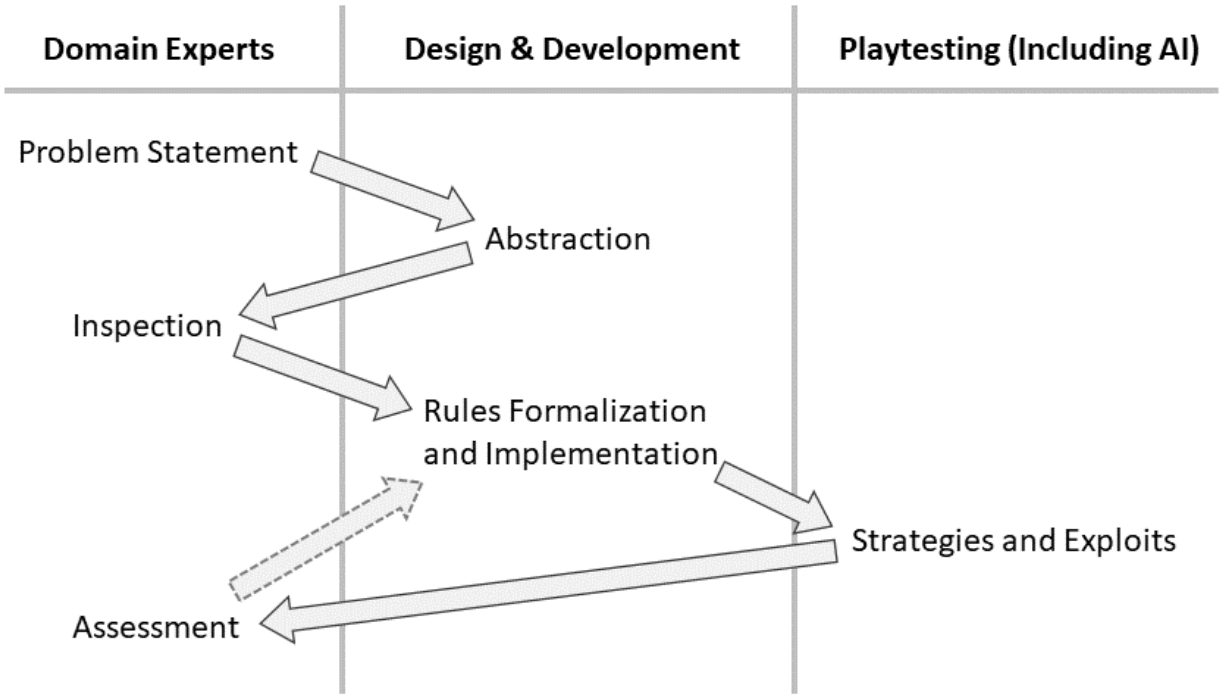


Figure 2: Workflow for integrating AI playtesting into the design process.

The design team had limited access to human playtesters. Some informal testing was done, and those results were compared to the AI. The humans found similar issues to those that the AI revealed, but they were less prolific and thorough in their analysis. The AI amplified limited playtesting resources, using human playtesting as validation of the high-volume machine testing. It allowed the designers to focus on fixing problems rather than anticipating them.

The design team had limited access to the domain experts. Discussing dominant strategies in the game was a better use of domain expert time than asking them to anticipate the lessons from a rule set. The AI analysis focused the domain experts on their specialty -- what should be taught, rather than what will be taught.

We expect that this type of human-AI teaming will become commonplace in design tasks long before such tasks are handed over for full automation, if they ever are. As Károly Zsolnai-Fehér says *"We have to make sure to formulate our problems with proper caution because the AI will try to use loopholes instead of common sense to solve them."* [13], but we see that quirk as an asset rather than a liability. Finding loopholes and exploring consequences can be exactly the skill required to complement a human's ability to build systems to satisfice vague guidance [14, 15]. An AI's myopic and literal interpretation of a given ruleset yields insights into both solutions and better problem formulation [16].

2.4 Ethics

Let's put this approach back in the context of the three categories of unethical dynamics and consider where it does and does not help.

- **Unethical *Design* of Games**. AI playtesting will show you a set of strategies that players are being implicitly rewarded for pursuing. It remains a human design task to not abuse that knowledge to rig the game and to properly assess whether those are strategies the game ought to implicitly reward. The machine anticipates, and the human interprets.
- **Unethical *Outcomes* of Games**. By knowing what in-game behaviors are rewarded by the game, the analyst has better insight into whether observed player behavior is a response to in-game pressures or revealing out-of-game bias and experience.
- **Ethical *Dilemmas* in Games**. Predicting what strategies a game rewards will not tell you how human psychology will respond to being exposed to those strategies or dilemmas. However, it does tell you whether optimal in-game strategies *require* sacrifices or crossing ethical boundaries, or whether clever play can avoid the need to compromise. The AI will help us understand the frequency and severity of the compromises the players will face, better equipping the human to assess whether that is a positive or negative of participation.

2.5 Drawbacks and Limitations

Although we have found pragmatic value in this approach at times, it isn't always practical or suitable. In reference to this document's title, keep Betteridge's Law in mind -- the answer to a question posed in a headline is never a simple *'yes'* [17]. In this case, we would answer the posed question as *'No, but it can help'*.

1. **Formalism and Structure**. AI-based analysis assumes that the game rules are formalized enough to be fed into an AI optimizer. If the action space and state space are not well defined, AI methods cannot be applied. Informal wargames can be effective [18], but this method does not apply to them.
2. **Result Interpretation**. If the AI analysis finds a weird result, that doesn't tell you if you should fix it as a problem or report it as a finding. This method helps alert designers to unexpected results, but the onus of interpretation still falls on the human.
3. **Malice and Misuse**. AI-aided design provides no protections against intentional distortion of results through game balancing. In fact, it provides a tool for a malicious designer to be even more effective at generating their desired or expected result. Once you have the AI framework set up, it is trivial to calibrate a game until the game incentivizes the behavior you want to see. Like anti-gerrymandering algorithms, formalizing a process can help or impede the elimination of bias depending on how it is employed [19].

4. **Psychological Impact**. How players respond to being forced to operate in morally gray areas is well beyond the ability of modern AI and psychological models. An AI analysis will tell you whether players are implicitly rewarded for accepting collateral damage, but not if that will produce more informed decisions or ethically numb decision makers.
5. **Repetition and Prediction**. The AI is playing and replaying the game to discover successful behaviors and strategies. If the humans will have that opportunity, we expect them to generally converge on similar strategies. However, if human players are experiencing the game just once, then the AI behavior is much less likely to predict human behavior. On a first play, players are less likely to discover exploits or internalize emergent game pressures.
6. **Time and Resources**. Training an AI requires computing resources to find good solutions -- the larger the action space and state space, the more compute cycles will be required for the AI to be competent. Many high-prole AI successes [20, 21] live on top of massive computing infrastructure and/or enormous data repositories. We have found value with smaller infrastructure on smaller games, but more complex wargames are beyond the reach of current methods and the infrastructure likely to be available.

Issues 1 through 5 apply to any method for assessing the emergent properties of written rules; they are not unique to the use of an AI playtester. Issues 4 through 6 are likely to lessen over time as of-the-shelf AI packages improve and predictive models of human cognition improve. As such, we expect AI playtesting to become increasingly powerful and viable.

3 Takeaway

AI can be a tool to assist human design, serving as design aides and teammates. By revealing unintended consequences with myopic dedication, an AI can inform human design judgments, complement expert analysis, and help predict the impact of the game on its human participants.

It is remarkably easy to inadvertently create very twisted incentive structures in a game. Manual inspection by experts is essential but inadequate, and even a simple automated analysis of the formal rules can reveal critical flaws and lead to significant improvements. In software development, including video game development, it is common for 50% of a team to be testers. Wargame design, either through necessity or culture, tends to allocate many fewer resources to testing. Wargames are most often one-off events, making high-manpower methods (like human playtesting) impractical. An AI surrogate for human playtesting may be a compromise between practical limitations of the wargaming world and best practices of the commercial world, especially as AI methods become more accessible.

We close by restating the three broader questions discussed earlier in this document, which extend beyond the scope of this document.

- What best practices are so foundational as to be unethical to ignore?
- What processes and methods should designers and analysts employ to responsibly differentiate bugs from discoveries?
- How should designers assess the secondary effects of participating in a game, especially one that includes ethical dilemmas?

Acknowledgements

We are grateful for the input from the Unethical Wargaming Working Group from Connections 2021, especially Stephen Downes-Martin for organizing the working group, Daniel Tyler Brooks for some excellent references, Rex Brynen for astute historical anecdotes, and Hiroyasu Akutsu for insightful feedback.

References

[1] Daniel Brooks. "Propaganda Games". *Extra Credit YouTube Channel*. Oct 2020 https://www.youtube.com/watch?v=UP4_bMhZ4gA

[2] Joel Kurucar; Robert Seater. "Unethical wargaming: A taxonomy for practitioners and observers", *Proceedings of Connections US 2021*, June 2021.

[3] Christopher A. Weuve; Peter P. Perla; Michael C. Markowitz; Robert Rubel; Stephen Downes-Martin; Michael Martin; Paul V. Vebber. "Wargame pathologies" https://www.cna.org/CNA les/PDF/D0010866.A1.pdf, September 2004, CNA.

[4] Paul K Davis; James H Bigelow. *"Experiments in Multiresolution Modeling (MRM)"*. RAND National Defense Research Institute, 1998.

[5] Rachel Simmons; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.), *"Zones of Control: Perspectives on Wargaming"*, chapter Goal-Driven Design and Napoleon's Triumph, pages 201-207. Game Histories. MIT Press, 2016.

[6] Robert MacDougall; Lisa Faden; Pat Harrigan (ed.); Matthew G. Kirschenbaum (ed.). *"Zones of Control: Perspectives on Wargaming"*, chapter "Simulation Literacy: The Case of Wargames in the History Classroom", pages 447-454. Game Histories. MIT Press, 2016.

[7] Amna Greaves; Joel Kurucar; Robert Seater. *"Serious Games and Artificial Intelligence"*, MIT Lincoln Laboratory and the BeaverWorks Summer Institute, 2020.

[8] Micah Zenko. "Millennium challenge: The real story of a corrupted military exercise and its legacy", *War on the Rock*s, November 2015.

[9] Kyle Mizokami. "The U.S. Lost a (Fictional) War with Iran 18 Years Ago", *Popular Mechanics,* January 2020.

[10] Robert Seater; Charles Rose. *"Skill Transfer and Virtual Training for IND Response Decision Making: Game Design for Disaster Response Training"*, MIT Lincoln Laboratory, 2016.

[11] Benjamin Eysenbach; Abhishek Gupta; Julian Ibarz; Sergey Levine. "Diversity is all you need: Learning skills without a reward function", *CoRR*, abs/1802.06070, 2018.

[12] Patricia McDermott; Cindy Dominguez; Nicholas Kasdaglis; Matthew Ryan; Isabel Trahan; Alexander Nelson. *"Human-Machine Teaming Systems Engineering Guide (MP180941)"*, MITRE Corporation, 2018.

[13] Károly Zsolnai-Feher. "4 Experiments where the AI Outsmarted Its Creators", https://www.youtube.com/watch?v=GdTBqBnqhaQ&list=RDCMUCbfYPyITQ-7l4upoX8nvctg&index=2&ab channel=TwoMinutePapers, 2018. Two Minute Papers (YouTube Channel).

[14] Nicky Case. "How to become a centaur". *Journal of Design and Science*, MIT Media Lab, MIT Press, February 2018.

[15] Wikipedians. "Herbert A. Simon", https://en.wikipedia.org/wiki/Herbert_A._Simon.
Wikipedians. "Satisficing", https://en.wikipedia.org/wiki/Satisficing.
Herbert A. Simon. "Rational Choice and the Structure of the Environment", *Psychological Review*. 63 (2): 129–138, 1956, https://uk.sagepub.com/sites/default/files/upm-binaries/25239_Chater~Vol_1~Ch_03.pdf

[16] Dario Floreano; Sara Mitri; Stephane Magnenat; Laurent Keller. "Evolutionary conditions for the emergence of communication in robots", *Current Biology* : CB, 17:514-9, 04 2007.

[17] Wikipedians. "Betteridge's Law". https://en.wikipedia.org/wiki/Betteridge%27s_law_of_headlines

[18] Jeremy Sepinsky. "Is it a wargame? it doesn't matter: Rigorous wargames versus effective wargaming", https://warontherocks.com/2021/02/is-it-a-wargame-it-doesnt-matter-rigorous-wargames-versus-effective-wargaming/, February 2021. War on the Rocks.

[19] Jordan Ellenberg. "How Computers Turned Gerrymandering into a Science", *The New York Times*, Oct 2017.

[20] David Silver; Aja Huang; Chris J. Maddison; Arthur Guez; Laurent Sifre; George van den Driessche; Julian Schrittwieser; Ioannis Antonoglou; Veda Panneershelvam; Marc Lanctot; Sander Dieleman; Dominik Grewe; John Nham; Nal Kalchbrenner; Ilya Sutskever; Timothy Lillicrap; Madeleine Leach; Koray Kavukcuoglu; Thore Graepel; Demis Hassabis. "Mastering the game of go with deep neural networks and tree search", *Nature*, 529, Jan 2016.

[21] David Silver; Julian Schrittwieser; Karen Simonyan; Ioannis Antonoglou; Aja Huang; ArthurGuez; Thomas Hubert; Lucas Baker; Matthew Lai; Adrian Bolton; Yutian Chen; Timothy Lillicrap; Fan Hui; Laurent Sifre; George van den Driessche; Thore Graepel; Demis Hassabis. "Mastering the game of go without human knowledge", *Nature* volume 550, pages354–359 (2017).

About the Author

Robert Seater is a member of the technical staff at MIT Lincoln Laboratory. He uses interactive models, serious games, and human-machine teaming as data-driven analysis tools to study expert decision making, training & education, mission-level impacts of emerging technologies, AI integration into workflows, and system-level requirements analysis. He has worked across a wide range of application domains, including emergency response management, chem-bio defense, drug smuggling interdiction, military logistics, distributed command and control, urban warfare, countering human trafficking, hypersonic missile defense, and naval missile defense. He received a BS from Haverford College in mathematics and computer science in 2002 and a PhD from MIT in computer science and requirements engineering in 2009. He has been at MIT Lincoln Laboratory since 2009.

Unethical Wargaming Discussion

The Challenge of Unethical Wargaming

Stephen Downes-Martin

The stated focus of the working group is: *"how to use unethical practices to make your wargame say what you want it to say."*

It's a thought experiment to help us identify design principles and malfeasance that we would perhaps miss if we focused on best practices of well-designed games by well-intentioned competent experts. By definition we mean competent unethical practice, else it would not fool intelligent and honest people. So we might be tempted to exclude accidents, mistakes and incompetence if it were not for the question "when do mistakes and incompetence rise to the level of unethical?" Does this occur when the results are damaging?

Note that "winning the game" is not a good synonym for "make your wargame say what you want it to say". One might deliberately lose a game using unethical means to sabotage a novel concept being gamed. I have seen that attempt made by a 2-star General (who admitted as such, not realizing I was in the room -- we were able to take that into account during adjudication and by appropriate and detailed caveats in the final report). The General believed he was "doing the right thing".

So, given a well-intentioned actor whose actions produce damaging unintended consequences, under what circumstances does one call that actor "unethical"? Can one argue that someone is unethical if they are well-intentioned and know they lack the required skill to produce good results, but nevertheless go ahead and act anyway? Do they become ethical if there is no one else to act and by not acting they risk greater damage than by acting? Then there are people subject to the Kruger-Dunning effect -- well-intentioned, ignorant, *and ignorant about their ignorance*. Are they unethical no matter how damaging or otherwise the results? When does ignorance become a lack of ethics?

It is not necessarily an unethical practice to wargame unethical behavior. By doing so one may have the opportunity to highlight the lack of ethics of the behavior being wargamed. But that raises the question "What does it mean to be unethical?"

Finally, unethical behavior has to go beyond cheating, breaking rules, doing illegal stuff, etc. It must also be *undetected* by the target. We are interested in *effective* unethical behavior. So discussion about cheating, cheating within the rules, breaking rules, doing illegal stuff, deception etc. are critical. We also need to examine how to detect when unethical "stuff" happens, so we can then examine "how to hide the fact that it has happened" (or even "how to lure the target into complicity").

Stephen Downes-Martin

I have my own definition of "unethical wargaming" that I do not impose on anyone else in this group, but we should understand each others' definitions to ensure clarity of discussion. In my view a wargame is sponsored to provide information to the sponsor or the sponsor's chain of command who has to make some decision after the game. So (again in my own view) unethical wargaming is the attempt to influence the decision making that follows the wargame to support a predetermined decision or a decision against the interests of national security (or the organization sponsoring the game for a business game). So unethical behavior is not just by players during the game, it can and does occur during pre-game discussions with the sponsor and post-game analysis and reporting. This assumes an ethical sponsor. When faced with an unethical sponsor (whose objective is to support a predetermined post-game decision) then the unethical game designer enthusiastically assists the sponsor and it is up to other stakeholders to spot what is going on.

(Note that unethical behavior does not become ethical just because the decision that results from the behavior turns out to be a good one. How a decision turns out can only be assessed after the decision, and the outcomes of any but the most trivial decisions are subject to uncertainty.)

Rex Brynen

Regarding "unethical," there's a fuzzy line between honest mistakes and carelessness.

Joel Kurucar

I think defining what is unethical is a tricky, difficult task. In my mind it boils down to intent (i.e. a mistake vs malice). I take a pretty practical position on this and just don't get too into the weeds on splitting hairs on what is unethical and what are just cases where wargames can go wrong. This is especially true since the observable behavior may be exactly the same for a mistake and an intentional malicious action; determining underlying intent is certainly not possible in a lot of cases. Therefore, I would take a broad, practical approach and err on the side of including things as unethical even if there is some gray area in the definition.

Ed McGrady

The distinction needs to be made between something that is "unethical" (a very pejorative term) and something that is due to misunderstanding, lack of knowledge, or unintentional (a mistake). Even professionals make mistakes, so even "unprofessional" carries with it a whiff of misbehavior that we may not intend.

Stephen Downes-Martin

Our legal system distinguishes between intended effects (planned murder) and accidental effects (manslaughter). The act that led to a person's unintended death may itself have been unethical (speeding, say) but one distinguishes between planned vehicular homicide and death caused by speeding.

Perhaps a way to think about this is to argue that a professional is ethically bound to be sufficiently skilled, and to know the boundaries of their skill and not cross those boundaries. Does deliberately acting outside of their skill become unethical (whether or not the result is damaging) unless sufficient justifications are available? So accidentally presenting exposition as analysis is unethical if done by a professional who should know better (but was sloppy, say), or done by someone who knows they are not qualified no matter how well-intentioned?

Anne M. Johnson

I believe unethical and immoral tend to be conflated; and they should not be. Just because an output might not be deemed worthy of consideration by one country/group/player based on moral considerations does not necessarily mean there was anything unethical done in getting to that outcome. Just because one country's morals would stop it from a course of action does not mean any other country would be so constrained.

Different categories of unethical behavior might include:

- Cheating
- Lying
- Deception
- Stealing

Unethical behaviors can fall inside or outside either the spirit or letter of the law/rule/moral code/etc. (I'll use 'the rules' to include all these). There could be a spectrum of players that need to be considered depending on the objectives of the wargame: someone who follows the letter of the rules, someone who will cheat the letter of the rules but stay within the spirit/intent of the rules (or vice versa), someone who will break/ignore the rules.

- An unethical behavior that falls outside the rules could be used to determine what additional/different perspectives are needed to better inform broader (dare I say 'more complete') understanding of a situation or topic. I'll refer to this as divergent purposes.
- An unethical behavior that falls inside the rules could be used to refine a (believed to be) fairly well understood situation or topic. I'll refer to this as convergent purposes.

Understanding how unethical behaviors could be used by players (which is why a list of unethical behaviors would be useful) could inform wargame designers how to enable or inhibit such behaviors, to support the either divergent or convergent purposes of the wargame

objectives. (Of course, there are many additional areas to consider, such as self-deception of either wargame sponsors or designers.)

I wonder if some of the cheating examples discussed elsewhere (chess, sumo wrestling) can help inform how cheating might remain at a tactical level or connect tactical to operational or even strategic levels.

Let me bring into the discussion the idea of 'amiable collaboration/competition'. An example is the National Football League (NFL). Players compete individually for spots on a team, higher salaries, awards, etc. And players come together (collaborate) as a team to compete against other teams. The teams compete against each other for playoff spots and collaborate with each other to keep the NFL going and successful. In the chess cheating example, individuals were cheating to advance their competitiveness against other individuals. Their cheating remained at the player/tactical level. In the sumo wrestling example, the cheating players collaborated with each other. I don't know much about the sumo wrestling world, but if it's anything like the world wrestling federation, the wrestlers collaborate to keep the federation going and successful (operational level).

So, I'm wondering:

(a) Does competitive cheating stay at the level of the cheater (say tactical), and does collaborative cheating cross to a level above the cheaters (say operational)?
(b) If so, would this help us better design games to understand how effects at the tactical or operational levels of war could drive the need for changes at the strategic level?

Stephen Downes-Martin

I believe it is possible for a behavior to be ethical or unethical depending on the situation and objective. Lying is normally considered unethical and indeed some philosophers claim any lie is wrong no matter the context. However, lying to a mother dying at the roadside by telling her her child survived the crash when it did not might reasonably be considered ethical. So maybe a taxonomy should include the objective of the person(s) carrying out the behavior?

Deception can be viewed as an unethical behavior, but is it also a necessary addition to an unethical behavior if the unethical behavior is to be effective. Even in situations where the perpetrator wants the target to spot the unethical behavior and thereby respond in a way desired by the perpetrator, that latter fact is hidden from the target. So the taxonomy will either have overlapping content. For example, although lying is (under certain circumstances) unethical, a lie like "I am a bright green tree frog" is not deceptive (well, I hope not!) and is thus ineffective. Unless the perpetrator was trying to deceive the target that the perpetrator was mad?

Back to your two questions. Is it useful to generalize the questions from "cheating" to competitive and collaborative deception in which successful cheating is a component of deception? It seems to me that competitive or collaborative can be tactical, operational or strategic depending on who is or are the perpetrators.

Stephen Downes-Martin

When is it in fact ethical to deceive? Some wargames are designed to explore the advantages and disadvantages of alternatives specified by the game's sponsor. Wargames are usually employed when professional judgement and subjective values feature strongly in the situation -- i.e. when for example a quantitative OR approach cannot give a definitive answer -- and there is no scientifically proven "right answer".

Critically, it is often the case that the game designer (or other stakeholders) believe they know which alternative best supports national security (which sometimes corresponds to the sponsor's objective). Is it therefore the stakeholder's ethical duty to ensure that the wargame design, execution, analysis and report pushes this alternative as high as possible up the preference ranking of alternatives in the mind of the sponsor?

Although most rational and honest people believe that it is possible to be personally wrong, most people also believe they are not wrong for the decision they are making at the time they make it -- especially if they are senior people who have worked in the relevant field. (Gladwell 2009; Chapter 10 in Wills 1994; Langer 1975). Note the research that shows "preferences are constructed in the process of making a choice or judgement" and that those preferences depend on what process is used to elicit them (Tversky & Thaler 1990).

A stakeholder's ethical duty certainly is to explain this to their boss and colleagues. However, if the latter deliberately chooses to ignore the research and its implications then under what circumstances (if any) is it the stakeholder's ethical duty to game the outer and inner games to get the best possible result for national security? After all, it is no one's ethical duty to promote an alternative they believe is not best for national security.

Daniel Tyler Brooks

Unethical Wargaming is hacking the epistemology of the game for the sake of advocating an a priori belief rather than discovering insight into the phenomenon being modeled. It's a form of intellectual dishonesty that game designers are all probably guilty of to some degree because all models are wrong, but some are useful. The degree to which they are wrong shouldn't be a secret, but probably is if we don't want to undermine the faith in the methodology. And yet how many times have we witnessed a pathological use of a wargame's results as evidence for a claim outside of its original research objectives?

Joel Kurucar

We should spend some time looking into how other fields combat (un)ethical issues. P-value hacking in academic publishing, game balance in major titles, tournament rules for things like Magic the Gathering, etc. To me it comes down to humans being humans and being susceptible to their biases and goals. Wargames are no different. People want to be right. People seek information that agrees with what they already think. I think in some ways, a lot of ethical wargaming may come down to how players are recruited, how the game is framed to them, what controls are put in place so that analysts or those running the game can't cherry pick observations to back up their a priori beliefs, etc. I don't think the games themselves, no matter how well designed, can solve the problem..

Good research requires good methods and integrity. There are controls put in place to help make sure that is happening. I think we need to lean into the topic of ethics for the field as if wargaming is a research methodology. That can open doors to a lot of good thoughts on ethics in research to help tackle this issue.

Paul Vebber

I would add to the list of challenges "ways to prevent exploiting interactions the game designer did not account for" which is technically cheating in general, but can sometimes be rationalized by "all's fair in love and war". But is it unethical? Does doing so generate useful insights or subvert the objectives of the game with no added value produced?

Phillip Reiman

In my opinion 'Ethics' in the context of this Working Group is more about rules and rule-following than classical living right. Salen and Zimmerman divide players into 5 categories:

A. Standard - right attitude and interested in winning
B. Dedicated - Zealot and intense interest in winning
C. Unsportsmanlike - Not quite a cheater, but intense interest in winning may cause the player to violate the spirit of the rules, if not the rules outright
D. Cheats - pretends to have the right attitude and intense interest in winning.
E. Spoil Sport - does not have the appropriate attitude and no interest in winning

As to lawyers, we tend to fall into B and C because we are advocates and take winning seriously - perhaps too seriously. I have seen it a lot - I have seen it do horrible things. This is why we have a Bar, a Code of Ethics and a Disciplinary Board. The tendency to reward 'wins' over justice encourages type C players which is precisely what is wrong with my profession.

Stephen Downes-Martin

Phil's comments about lawyers being B and C type players raised a question in my mind. How can players be B and C type without letting the other players realize that is what they are? Answering that question is probably part of designing an unethical wargame (when the designer is the type B or C player in the outer game) and part of playing (when the stakeholder playing the inner game is type B or C).

Phillip Reiman

I saw this on RCD, "Honest words mistaken as truth are the shackles on a gullible mind. There is an important distinction between truth and honesty often ignored by conventional wisdom, which traditionally presumes one leads to the other." Nice review of a book about USG intervention in commercial media - but the review dives into the author's discussion of 'personal truth'. One can deceive with the truth.
https://www.realcleardefense.com/articles/2020/11/11/docu-fictions_of_war_583691.html

Mark Flanagan

I am struck by the thought that the "anti-system" approach (imagining unethical behaviour being the norm) describes most if not all large computer systems I have worked with. There seems to be more (professional, aka often paid for by the public purse) energy put into saying something is finished - when it clearly isn't. Sadly the vested interests seem to be far better at proving their mal-point than the guardians [sadly disorientated management] countering them.

Frames and Ethics in Wargames

Ed McGrady

While there are as many ways to think about professional games as there are stars in the sky, one way to look at them is according to whether the discussion is about stuff that is inside the game, or outside the game. I, because I have a strong academic bent, refer to this as the internal and external frames of the game. The internal frame looks at things like game components, player behavior, sponsor behavior, and how the game is interpreted or presents itself to the external world. The external frame is how the world looks at the game. I think the distinction is important not only for this discussion but also because we tend to dwell a lot in the internal frame and don't think a lot about the external frames given to our games. To put it another way internal frames occur within the magic circle, external ones occur outside the magic circle.

So, how does this deal with ethics? As I see it there are several ways in which the question of "ethics" – roughly defined as "right action and behavior" – intersects with games and my "frames":

Internal - mechanical: Does the game somehow do something unethical through the way in which it manipulates players actions or information? In other words, has the game been designed in some way that is not ethical? This is the canonical "let's make our weapon unstoppable, and then say the game found the weapon to be unstoppable." I believe this is actually the most common way in which games are "thrown" or manipulated for a cause. Simply because it is the easiest, particularly with system parameters. And, no matter how much computer iron and 3d simulgraphics you throw at the game, the weapon will still be unstoppable.

Internal - players: Here the players cheat within the construct of the game to do something unethical. This is classical cheating, rules lawyering, as well as manipulating their own or other's actions to throw the game.

Internal - analysis: I place the analytic portions of the game, all of them, both before and after the game, as part of the internal frame of the game. Here someone takes the results of the game and manipulates the results to accomplish some evil or disreputable goal. In truth I think this happens a lot, but mostly in an innocent way where people simply mis-characterize the game due to a lack of understanding or context. But it's super common.

From my look at the discussion here I think most of the discussion has centered around one or more of these internal frames of the game. Which makes sense, given that the charter for the group is to figure out how to manipulate games in a way that does something "manipulative."

The external frame discussion is less common probably because DoD is not inclined, nor chartered, to have that discussion most of the time. The question externally is typically whether something should be done, not how. Now you could say "well questions of whether come up all the time in strategy games" and you would help me a lot by doing that to demonstrate what the external frame is. Questions within the game of whether to take one action or another are still tied to the internal frame of the game – because they are being asked within the game. To get to the external frame we have to ask about the game, not about the internals of the game. As I see it there are only a few places where the external frame of the game intersects with ethical issues:

External - contextual: Should the game have occurred? Is it ethical to game whatever it is that is being gamed? An example might be gaming the use of violence as a means of suppressing a democratically elected government.

External - behavioral: Did the game have a real-world effect that resulted in something unethical happening? The clearest example of that would be some form of bias or harassment of an individual involved in the game.

External - standards and norms: Did the game somehow violate some standard or norm that it should not have. Presenting, for example, behavior widely considered reprehensible as the preferred course of action.

There are probably others that I'm missing but these are pretty broad in scope and representative. Now you may say "ah-ha" a lot of what you discuss happens within the game. Of course it does, it would be difficult to talk about the game without actually talking about the game. Internally or externally. The external frame for the game can still deal with stuff that happens within the game (someone punched someone) but it's now being evaluated by external ethical criteria that have nothing to do with the game (that is an actual crime). That is why I refer to these as "frames" which is information/media speak for how you spin something.

I think I could also make an argument that we often miss external frames in DoD because:

- It's not DoD's job, most of the time, to worry about such things.
- Most professional games are on topics where these issues don't get raised.
- We spend a lot of time worrying about "accuracy" or how the game should be done, and don't have a lot of bandwidth left over to ask whether we should be doing the game in the first place.

Stephen Downes-Martin

As far as I can make out Ed's definitions of internal and external frames are close to the definitions of inner and outer game that first appeared in the literature at the Naval War College in the early 2000's, and which have frequently appeared since then in discussions on wargaming pathologies (see the paper "Exploit Group Dynamics to Corrupt a Wargame" in this document). We have also discussed the inner and outer game in this forum. However Ed's discussion goes much deeper and picks apart the components in a really useful way.

I think a key difference between "Inner and outer game" and "internal and external framework", is that the outer game is part of "the game". It is the game played by the sponsor, the sponsor's chain of command, the game designer and the game designer's chain of command. The external framework is a much larger domain within which both the outer and inner game components of the wargame lie. It may be that the outer game insulates the inner game from the external framework, but I am not sure of that. In any case we are interested in unethical behavior in both the inner and the outer game.

Ed McGrady

"Making wargames say what you want it to say" implies that something is happening within the world of the wargame. Even pointing to the wargame and saying "the wargame says this" implicitly accepts the legitimacy of the wargame in the first place. What I am talking about is questioning the ethics of the overall problem, not the particulars of the game itself.

Calling a wargame racist, for example, is not referring to the internals of the game, but it is accusing it of being unethical. Or a wargame for Wikileaks designed to understand how to exfil secrets. Pick the thing you despise most, criticizing the wargames ethics requires you to stand outside of the wargame to do it.

You can, of course, throw a wargame so it says something externally objectionable, or do something within the game that generates an external criticism. I just don't think that understanding the mechanics of how you throw the game is equal to understanding why the game or action is morally wrong. Those are two different things.

The external frame of a game is not a pathology, it is understanding the game as a semiotic symbol that has multiple resonances in the political, social, and human spaces as well as whatever the stated purpose of the game is. For example, you can discuss Joseph Conrad's "Heart of Darkness" as a piece of (late) 19th century literature, or you can discuss it as part of a system that upheld and promoted colonialism. I don't think we are engaged in anything like the latter project here.

Zoltán Harangi-Tóth

What Ed mentions (external framework eg. the broader environment of the wargame) is more of a fast changing environment (see the colonial background example and it's acceptance in 1980 or in 2020), a dynamic thing, but the internal framework is something static/still. Cheating in-the-game is unethical all the time whatever the exact method used, but the external framework (like the mentioned social factors) are rapidly changing. This kind of external framework could rapidly become obsolete, so we would have to be very specific about the details.

Ed McGrady

This is exactly what I'm saying (leaving all the time dependencies out of it). You can structure it as "ethical" vs. "moral" if you assume ethics is somehow fixed by rules independent of culture (it is not) and that moral floats according to culture and beliefs. But it is that distinction between the "ethics/internal" of the game and the "morality/external" of the game that I'm getting at.

Breaking the Rules versus Cheating

Stephen Downes-Martin

Chapter 21 (Breaking the Rules) in "Rules of Play: Game Design Fundamentals" by Katie Salen & Eric Zimmerman (pp 267--285) is worth a read. It describes five player types and compares their relation to the formal game system, and explores a range of behaviors.

Paul Vebber

The "meta" around the execution of a game is fraught with the specter of ethical dilemmas caused by conflicting motivations of:

- the players (the white cell being a player, as Stephen reminds us),
- the sponsor paying for the game (who has an agenda for doing so)
- the customer of the game - often not the sponsor who similarly has a agenda for the game event and its analytic artifacts,
- the game designer, who has an agenda for designing the game a certain way (which may simply be retaining their sanity),
- the game designer's boss (Stephen's 3 witches are all in the soup now)
- and those who have capabilities or concepts included in the game and part of the "decision space" being explored.

Then to make things really interesting you have the "outside influencer" who is out to make the game fail and will play with the intent of blocking easy avenues to victory (generally not an adversary country, but - the real adversary - another Service). There are also the 2nd and 3rd order influencers, but I will hold it at those 7 for my contribution to the efforts here - a game about analytical gaming - highlighting where the competition between those involved in the "meta-game" surrounding the game produce a soup of ethical dilemmas - totally separate from those involved in the actual playing of the game and the associated rules, lawyering and shenanigans.

Stephen Downes-Martin

The players' objectives should not be the same as the sponsor objectives. The players' objectives are to win the game. The sponsors' objectives are to achieve some insight or result that satisfies the purpose of the game. I.e. gain an analytic insight, achieve an educational or training objective, deliver a useful experience, or convince some senior leaders that a concept is useful and worth further investigation or not. Malfeasance occurs when the players play to satisfy or derail the sponsor's objectives rather than playing to "win the game". Good game design makes this difficult but not impossible. I have observed this several times, the most egregious was a game with the analytic purpose of comparing two C2 structures. I overheard

the 2-star General ordering his service subordinates who would be playing the game to make no game decision that could support the concept his service opposed and to only make decisions that supported the concept his service wanted, even if it meant "losing the game" by making dumb moves. What I did about this is another story, suffice it to say I successfully countered the General's move in "the outer game". Which brings us to "the outer game".

The inner game is what we think of as "the game". The outer game is the one played by the sponsor and other stakeholders with the recipients of the game results and insights. Ideally the game designer (along with their bosses) and the adjudication cell play in both games simultaneously. The sponsor should only play in the outer game. However sponsors and senior outer game players will try to influence the inner game by various means including placing "their people" into the inner game as stooges.

Kayla Capps

Coming at this from an intelligence perspective, I would say the number one thing to be on the watch for during the design phase are the assumptions that get built into the game. Faulty assumptions, often unconscious and unarticulated, are a component of essentially every major intelligence failure (and probably most operational failures). This is not to argue that we should not incorporate assumptions into games. The fact is they are almost always needed for a variety of reasons such as missing data, a need to approximate factors, or a need to control the scope of the problem under investigation in terms of space or time. However, assumptions need to be clearly stated and sensitivity analysis should be done on them to increase understanding of relations between key variables and outputs and the overall robustness of the game system. Similarly, scoping assumptions regarding the applicability of a game need to be articulated so that it is clear to what and how long the output is applicable.

Chad Briggs

I agree with Kayla about built-in assumptions, and from my experience both in defense wargames and hobby games, hard-wiring rules to fit certain assumptions can easily lead to blind spots, and at times motivate people to CWR out of frustration and their own sense of achieving realism and plausibility. Some games can add so many rules, charts, tables, etc. that actually playing through becomes unworkable. Others (like matrix games or one-page rules like Cthulhu Dark) rely on cooperation and agreement among the players to work.

This is a huge challenge when wargaming irregular warfare and disinformation, which by their nature encourages players to cheat. My American students, especially military and intel officers, keep asking for hard rules to follow even when they're assigned to the red cell as pirates, they keep wanting to play within the rules even when told not to. By contrast, players in the Balkans and Eastern Europe (mostly Kosovo and Ukraine) have to be given wide latitude for

actions, because for them the rules don't seem to exist very well in real life. The benefit is when players create pathways I never would have thought of: inserting spies into other cells, a embezzlement scheme that siphons off humanitarian aid and launders it to fund mercenaries, creating disinfo videos targeting specific players, etc. - and I say benefit in terms of red team ethics. All those actions would have broken rules from my American perspective, but the point of hybrid warfare is that we have to be prepared for actions we don't expect. I can't come up with a random action chart of those things in advance. In the end, participants are better able to spot when rules are broken in real life because they've done it or experienced it already.

In other words, CWR can be useful, but it depends greatly on expectations of the sponsor, referee, and especially participants. Also, fairness, plausibility and realism are different things, but all have to be considered. For example, it was entirely realistic for me to include misleading data in a cell's briefing (intel reports do get things wrong), but the American players complained that it was unfair of me as referee to do so. In Ukraine or the Balkans, the question was much more "could this happen?" and fairness was secondary. To that end, in Europe I typically construct the green cell with internal divisions, corruption and competing loyalties, a more realistic portrayal of many governments (this is easiest in fictional settings) but again one that most American players are uncomfortable with.

Stephen Downes-Martin

The takeaway I get from Kayla's and Chad's posts is that manipulating assumptions can be a method of unethical wargaming in several ways:

1. Invent assumptions that encourage the sponsor and/or the players to think in certain ways. I say "invent" because they are not necessary for the game or necessarily true. Use plenty of jargon and include accepted facts. Use similar words and phrases in the accepted facts that you use in the invented assumption so the fast reader does not notice the lack of logic connecting accepted fact to your invented assumption. And, similar to propaganda, make sure there's a kernel of truth in your invented assumption.
2. Overlook real and necessary assumptions.
3. Don't mention faulty assumptions that you notice are in play.
4. During play, assert assumptions starting with the words "we all know that ..." and say it with a casual but firm and authoritative tone of voice.

Hiroyasu Akutsu

How common are "shenanigans or rule-breaking" during wargaming in your community? Is it widely-observed in the western wargaming community? Is it cultural? I am asking these questions because I have not seen such behaviors in our professional policy simulations/gaming environment. It may be that I have not recognized "shenanigans" and that they do occur and

are often overlooked. Personally, I pay more attention to distinguishing between intervention from either outside authority or the controller (including ourselves) designed to influence the players' decision-making and the necessary injections by the controller as part of its efforts to make adjustments to the way a game proceeds.

Cheating within the Rules

Stephen Downes-Martin

What are the differences between cheating and unethical behavior? Is cheating "breaking the written rules"? People have pointed out the existence of unwritten rules aka societal norms, expected behavior, obvious behavior etc. One cannot as a practical matter write down everything one is not supposed to do. We have a values problem when discussing unwritten rules or societal norms. What is a norm to you might not be one to me, or we might disagree on the relative importance of two norms and thus disagree on whether breaking one norm versus another is unethical or ethical. So, is "being unethical while working within the rules" simply "cheating within the rules" (CWR)? And, Is it always unethical to break the rules?

Although players using CWR might prefer their cheating to be covert in order to protect the appearance of the integrity of the game's results, they are protected from personal punishment if their behavior becomes overt since they are not breaking the game's rules. It seems to me there are several areas of investigation, including:

1. How can players, including the adjudicators, cheat within the rules to achieve a game outcome desired by the player (not necessarily the game's official victory condition)?
2. How can game designers deliberately encourage CWR by carefully crafting the game rules so that the game outcome better supports a desired agenda?
3. What does it mean to cheat within the rules in the outer game with the sponsor and other stakeholders?
4. When does CWR provide ethical benefit to the sponsor?

Phillip Reiman

Does anyone have thoughts on creating games where Cheating within the Rules (CWR) is encouraged by design? I can see results of area 4 (above in Stephen's post) benefiting taxation, policing, political institutions, policy creation and any other area where participants 'struggle against the rules' directly as opposed to between individuals. A vs. Society contrasted to A vs. B contrasted to A vs. Nature.

For example, see the old game "Illuminati".[67] According to Salen and Zimmerman, this is a 'cheat to win' game, and provides a great foundation for discussing Unethical Wargaming Design and CWR.

[67] http://www.sjgames.com/illuminati/img/illuminati_rules.pdf

Hiroyasu Akutsu

To me, it sounds like the Blue team and its allied teams trying to list up the Red team's "gray zone tactics" and to get it to agree to a code of conduct. The Red team may say it would accept such a code but actually ignores and continues to cross the line while the Blue team and its allies cannot successfully stop the Red's "gray zone" actions. After all, no one on the Blue side can check the Red because it is not easy for the Blue side to draw the line between CWR and rule-breaking and also because the Red may be talking to each Blue ally about having a separate deal behind the scenes.

Robert Seater

That's an interesting observation. Enforcing rules is a game tactic, especially if they are informally enforced rules along the lines of 'play in the spirit of the game' or 'play realistically'. If you are studying grayzone warfare, red's big advantage (and realistic options) all involve pushing the limits of what is considered fair or realistic. So, enforcing rules strictly or discouraging players from pushing the limits of the game could be an unethical bias to pretend that blue will win when really red will 'cheat' in the real world.

Hiroyasu Akutsu

My focus is not really on bias, but I would rather find it interesting to see how the Blue's counter-gray zone measures fail by emphasizing the importance of complying with the existing rules and paradoxically creating even more space that the Red can exploit to its advantage. Some scholars of international relations might call both the Blue's and the Red's behavior as "organized hypocrisy."

Phillip Reiman

The current emphasis on operations below the threshold of armed conflict and multi-domain operations (MDO) has brought up discussion of lawfare and gamesmanship. Frankly, speaking for myself and probably most wargamers - gamers, wargamers and social scientists – it's a form of cheating. Or maybe better described as cheating within the rules (CWR). Instances are everywhere: from hockey's 'chirping' or basketball's 'trash-talking' to hyper-defining concepts in business contracts or international treaties. One favorite example comes from golf, where if you feel your opponent favors a faster game, you ensure the game is slow by taking a long time to hunt for lost balls. Of course, only for your opponent's ball, as that's only sporting. It's not cheating – it is CWR.

Take a look at the Snopes article "How To 'Weaponize' Misleading Narratives About Voting: Lessons from Trump-Linked PILF Playbook".[68] The interesting idea here for us is (skip the politics

[68] https://www.snopes.com/news/2020/10/15/the-fog-of-war/

if you can) - Creating false information - exploring the idea that in 'deception' there are still 'unwritten rules'. Cheater's Rules: not rules as written and not easily described. Example: how much cheating is too much against a given opponent? How much cheating undermines the legitimacy of a win? Can you cheat on behalf of an unknowing participant in the more legitimate game (aka 5th column)? It comes up in multi-domain ops in that the Russian idea of MDO is to 'cheat' [in a Western view] just up to the point below where things get kinetic. How far can you push the rules?

I use 'cheating' as a term meaning 'act dishonestly' or violate rules in order to gain an advantage. I am not talking about degenerate strategies or simply misdirection. Now I am thinking about it - fake tanks presented to enemy = decoys - totally ok, but fake information presented to non belligerents = er, informational decoys?

It's also worth taking a look at SOCOM's "Resistance Operating Concept".[69] It provides a narrative, a method of recruitment and suggests tactics while setting out a larger strategy. Impressive. Politics aside, this is how it's done. This is not really 'unethical' wargaming, but it is MDO, resistance actions which are often seen as 'unfair' in a hypothetical force-on-force world.

So, we're examining players' relationships with rules. In the unwritten, unspoken social contract that we engage in when we explore decision-making through games there is the implicit agreement that we will follow the rules. It applies in all cultural settings. Cheat and you are out. Opprobrium or Ostracism. Or someone declares shenanigans. My point is that CWR is real. It's a big deal and we need to explore it.

Daniel Tyler Brooks

Cheating within the Rules can actually be the primary insight of a wargame, if the goal is finding players with Strategic brains, and nurturing them like say, at the Basic Strategic Arts Program (BSAP). A good example of this from a few years back was when CYBERCOM and West Point (of all places) collaborated on an exercise inspired by Star Trek's Kobayashi Maru and Naruto's Chunin Exam (https://youtu.be/7902m6MTu0E).

A game design that encourages this is Magic the Gathering, where in tournament play, the Meta-game and card economy is built around exploiting assumptions about who is bringing what to a tournament, and the Paper-Rock-Scissors of discovering the "Best Deck" which then causes everyone to play the "Best Deck" which causes others to build decks that are terrible at everything except beating the "Best Deck." As a case study, there's a lot of mileage out of Magic the Gathering as a game about rule making and rule breaking.

[69] https://www.hsdl.org/?abstract&did=838149

Consider the idea of the "Spirit of the Commander Format" in the Commander Casual format of Magic the Gathering. It's a multiplayer, singleton, eternal format where there are cards that can be used for things that are considered "Degenerate," and there are playground expectations about what the group is ok with as a casual format. Some argue "If the official committee for the format hasn't banned the card, then I should be allowed to run it in my deck..." and then she sits down to play a turn 2 win condition, and then the rest of the table says "good game" scoops their cards, and then refuses to play with that person until they pick a different deck. If the offending player refuses, the group shrugs and starts a new game without her. The social contract is enforced by sequential games... if you nuke all of the lands on the board, or play cards like "Winter Orb", everyone gets annoyed at you and doesn't invite you for future games. Political norms are like that. Why do Presidents not pack the supreme court? Because their opponents will do it when its their turn. Why are filibusters allowed? Because some day you'll be the minority and want it as a tool in your belt. Reciprocation in sportsmanship over infinitely repeatable games is how you get invited to keep playing. Because CWR might allow you to win a game, it also might cause you to never get to play again... which doesn't maximize your win percentages over time if you don't get to play. What does this teach us about strategy in wargames? Perhaps international norms in a system of international anarchy are just about reciprocation in infinitely repeated games in a multi-polar world... but then, what does that mean when an autocratic ruler is dying? This might be her last game, and she doesn't care if the whole thing burns down because she has nothing to lose. Same thing for a lame duck. I suspect childless world leaders are even more prone to negative-sum thinking, which strikes everyone else playing sequential games thinking her to be irrational, when in fact she's just got a different preference order and incentive structure.

Stephen Downes-Martin

Tyler's comments reminded me of the game theory research behind the 1984 book "the evolution of cooperation" by Robert Axelrod (Axelrod 1984). He looked into a simple prisoners' dilemma game and examined what strategies evolve and are optimal when you have to play the game many times (finite and infinite) with the same opponent and accumulate your scores. Start honest and use tit-for-tat seems the optimal strategy (caveat - for this game).

Paul Vebber

As Stephen offers – "if you are certain that it is in the best interests of national security for a certain outcome to occur, is it not then your responsibility to ensure that outcome is chosen?" Is it more ethical to "play by the game rules" and let a bad decision be made, or "CWR" (or "cheat when you hope no one is looking") to ensure the correct decision is made. What happens when multiple people involved are morally certain they are right and others are wrong?

Daniel Tyler Brooks

When you already know what is supposed to happen in the execution of a game, and as a designer you start making the game to ensure that outcome, at a certain point it becomes pointless to actually play the game ... and, I'd argue, you've put the cart before the horse and have created a "propaganda game."[70]

Stephen Downes-Martin

I believe egregious cheating within the rules and degenerate strategies (taking advantage of an oversight in the game design or rules) are very useful to the game designer and analyst, so long as it is spotted! It provides the designer with lessons learned about game design for next time and the analyst with additional insights (maybe) from the game about how people might behave in the real world in addition to guidance on caveating the analytic results.

And then there are games perhaps designed to encourage out of the box thinking. In which case cheating within the rules is the desirable behavior?

Montana Hunter

I think you're identifying three forms of cheating here:

1. Cheating that takes advantage of flaws in the game design or system (i.e. the example of the naval wargaming group that noticed text on their screen whenever a missile was firing). I argue that this is not cheating "cheating", it's more taking advantage of a flaw in the system/game design.
2. Cheating that is actively encouraged by the game design. The card game "Cheat" is perhaps the single simplest version of that (though even there the "cheating" is limited by rules. You're allowed to lie about the number of cards you have, but you can't hide cards underneath your seat or palm them to your opponent).
3. Cheating that is actively breaking the established rules of a wargame. Falsifying dice, making an illegal move etc

Of these three, I believe that only the third one should really be frowned upon in wargames/games.

[70] https://www.youtube.com/watch?v=UP4_bMhZ4gA

Exclude Real World Data

William "Wilf" Owen

If wargaming is primarily about understanding the totality of decision making and consequences, then how come the input to those decisions are game models and not real-world information?

Stephen Downes-Martin

In most of the professional wargaming I have been involved with over the last 30 years the inputs were real world data. One can and should argue the validity of that data. All too often it is optimistic range, exercise or manufacturers' data for equipment performance. Often however it is data from contemporary and historical conflict, and often the "range data" for Blue has been deliberately degraded by the game designer to be more realistic. Degrading optimistic data is done by old cynics using a mix of professional judgment and modeling and simulation.

William "Wilf" Owen

The real-world data I am more concerned about are those that drive decisions, such as cartography, communications, medical and logistics. Additionally just sources of hard information such as meters of usable dock frontage etc. If real-world knowledge and ability is not creating an advantage in the "game," then why would I trust the game?

Stephen Downes-Martin

I have seen each of the areas of data you mention not included in games for which the sponsor's objective did not need that data. Logistics is a good example. If the game is about comms over a short intense battle in which in the real world you would not be resupplied, then playing logistics is pointless. The danger you mention is when logistics is left out despite it being needed. The issue I would like to see addressed is how would an unethical game designer, player or sponsor exclude, wrongly edit, or manipulate the need for real world data in such a way that it supported a poor post-game decision and in such a way that the unethical behavior was not noticed. The more innovative and cunning the unethical behavior the better because then we have examples we might not have previously thought about and can start devising inoculation methods.

William "Wilf" Owen

I would express some caution when assuming there are professional wargames that do not need to model logistics, casualty evacuation or similar. They may not be a part of the game, but their consideration does drive real-world decision making. "Ethically" professional wargames should replicate or actually use real-world command drivers. I get very energised when I see Divisional command decisions being made by someone not trained or highly familiar with

Divisional Command. Likewise, I've seen a Brigade Staff perform a whole planning cycle that was completely irrelevant to the mechanics of the game they were part of. Thus when people gain "insights" from games where the commanders did not or could not execute COAs built from real-world procedures, then how safe are those insights?

Stephen Downes-Martin

Your points are well made and well taken. A problem faced by any wargame designer is the trade-off between playability, realism, and usefulness. The game designer and sponsor "have to pick two", meaning they cannot have the best of all three. Occasionally a breakthrough occurs that allows an improvement of all three, but the trade-off after the breakthrough must still be made. The game with ultimate realism, every can of beans, every bullet and it's trajectory, every decision by every private firing that bullet, is realistic but unplayable. Every game, every simulation makes simplifications in its representation of reality. You have raised an interesting question: how do you use the playability, realism, and usefulness trade-off unethically to make your wargame say what you want it to say?

(An interesting program for implementing and playing logistics in analog wargames is provided by Michael Hugos at SCM Globe.[71])

Anne M. Johnson

I think there are multiple factors that contribute to excluding logistics considerations in wargames - even as a component in the narrative leading up to the scenario of the game. [Note: I'm including lifecycle maintenance in the logistics category, not sure if it usually is.] Most players want to fight the enemy during the game, not the 'boring little details' that get forces into the initial game position at full mission readiness.

Perhaps a subset of games could (a) game the transition of force location and readiness from an 'everyday in the navy/army/air force/space force' to initial location and assumed mission readiness at other games initial starting points, and/or (b) include a game participant survey to collect player assumptions of force readiness and how close in their actual work the force meets those assumptions. I am aware of at least one TTX that tried the former, but (in my opinion) the ingoing assumptions contained a lot of self-deceit and hubris.

If either of these options were done, they could be used to determine a more realistic initial force set up. It seems quite obvious to do this, so maybe it is already being done. But, I haven't seen it in the games I've been exposed to.

[71] https://www.scmglobe.com/battle-smolensk-1941-invasion-russia/

Methinks the Referee Doth Protest Too Much

Mark Flanagan

Here are three wargame examples lying on what I thought as a continuum of ethics: From [Ethical] to [Partly Ethical] to [Unethical]

1. An Ethical Example: The Gallant Knight Exercises
    - A real problem that the NATO forces were unable to stop (1980s)
    - A Soviet push from Afghanistan to the oil rich Middle East severing a logistic artery
    - "Blue" did not possess the "lift" capacity to stop "Red" in time
    - Over a series of wargames over the best part of a decade "Blue" practical changes came into effect to remedy this
    - [Source: Matt Caffrey .. On Wargaming and Connections UK and Connections conferences]
    - Positive Takeaway: Evaluation and Feedback loop to a perceived threat produced operational capacity that was used in Desert Storm to good effect. It was a "warts and all" appreciation
2. A Partially Ethical (but very political) example: Millennium Challenge 2002
    - A huge series of wargames and practical operational exercises (of huge expense) that were grouped together
    - The "Red" team possessed an out of the box thinker (Van Riper) who was given free reign
    - Rewards for "Red" success were 'reset' partly due to operational exercise constraints and "the show must go on" to a timetable
    - Red subsequently felt constrained and had to conform to conventional expected behavior
    - [Source: Blink - Malcolm Gladwell]
    - Positive: Practical steps seem to be taken from "lessons learned" as the subsequent picture of a LAV-25 chained down on the deck of an amphibious assault ship indicates there was "some" value to Van Riper's concerns https://www.marinecorpstimes.com/news/your-marine-corps/2018/12/03/lavs-on-the-flight-deck-the-corps-is-testing-ways-to-boost-ship-security/
    - Negative: Huge scale of exercise meant that the results would themselves become a BOGSAT and the "results" were 'marketed' with inter-service bias.
3. An Unethical example: The Midway Wargames by Imperial Japanese Navy
    - As part of the IJN best practice all major operations were wargamed, Midway preparations followed best practice in a procedural way
    - Most famously the umpire overturned chance results that went unfavourably to the IJN and refloated two CVs (Kaga and Akagi).. but that was actually "ethical"

- The "sinkings" were the result of high level bombing from land-based B-17's from Midway - which from a technical perspective this was not deemed a very likely outcome
- The wargame was "unethical" because the assumption of an unopposed IJN first-strike on Midway was never challenged, that as Paul Webber pointed out, the IJN submarine screen was never thought to be compromised and "would" tell of USN CV movement from Pearl Harbour
- In fact something that the IJN were not aware of (an intelligence coup of the Purple Cipler and the infamous "distillation tower" ruse to determine where "AF" was) gave the USN "an edge" and the CVs had sailed before the IJN sub-screen was in place (which echoes the WWI 'Room 40' advantage the RN had over the HSF in the North Sea)
- An opportunity was missed to "scale up" the search capability to at least be on a level footing which also lead to the tragic circumstances of catching the First and Second Air Fleets tooled for the wrong job, at the wrong moment
- [Source: Peter Perla (Art of Wargame), Robert Berks (The Craft of Wargaming)]
- Negative: Wargaming became rote, tick-box and part of the "Victory Disease" (as Tameichi Hara - Japanese Destroyer Captain reported) overconfidence and respect for the enemy's capability waned
- See pages 23 -- 43 of https://apps.dtic.mil/dtic/tr/fulltext/u2/1021990.pdf for more detailed information on Midway. The author also nicely contrasts the influence of the Pearl Harbour wargames (akin to "Gallant Knight", in the way they positively changed operational planning and outcome by having a challenging Red player) to the Midway planning and games (where Yamamoto ended up spreading resources to include the Aleutians)

Michael Hugos

Concerning the three gaming exercises raised by Mark, in each of these games it seems one of the strongest signals of unethical intent was the way unexpected or undesired outcomes were denied or dismissed by the game referees for reasons lacking in intellectual integrity. Reasons given by the referees in these three games were that the actions of the opposing force were simply not allowed, not done, or unrealistic just because the referees said so.

It seems that after the shock and potential embarrassment caused by a game opponent who really is thinking out of the box, all game participants need to think about whether such out of the box actions could realistically happen. The actions that allowed the Iran team to win the Millennium Challenge in the opening round would have been easy enough to carry out in the real world.

When such moves are dismissed by referees simply as "unlikely", "improbable", or "people wouldn't do that..." etc. then that is evidence in its own right that the game is unethical and

predetermined for a desired outcome. It's just "refloating the fleet" and carrying on with a predetermined charade.

A counterpoint to the results of those wargame exercises is the results that were achieved by the German Army as they were conducting a tabletop wargame exercise on the defense of the area around the Hurtgen Forest. As they were running the wargame American forces actually did invade the Hurtgen Forest and the Germans continued their exercise incorporating what the Americans were doing in the real world. That wargame resulted in a tactical victory for the Germans. Did this game enable the Germans to devise effective tactics to counter the Americans? Did they come up with any out of the box moves? Perhaps there was not much out of the box thinking on the part of the American Army. Perhaps that was an example of a game working so well because the enemy did what you expected they would do?

Rex Brynen

Millennium Challenge is an interesting case. The Van Riper account of the game, whereby White disallows some unorthodox Red moves because they defeat Blue, is the one that has gone down in wargaming legend.

There's another version, however—the one where a combination of (1) real life civilian shipping in the exercise region, (2) narrow navigational safety constraints on the deployment of the actual ships being used in the exercise component, and (3) the failure of the JSAF model to be able to handle all this, meant that the ships were not meant to be where the wargame depicted erroneously them. On top of which, (4) much of the defensive fire adjudication wasn't even executed, since that part of the model had been turned off. Finally, (5) the game did not limit what weapons RED could mount on what small boats, so unlike the actual small cannon, ATGMs and RPGs that the IRGC mounts on most of its small boat fleet, these were packing destroyer-size naval SSMs.

The USN version therefore is that Red sunk a ghost fleet that wasn't actually in the location Red thought it was, using weapons Red couldn't have mounted on small boats, and even then only succeeded because Blue's defensive weapons across the entire fleet were turned off.

I wasn't there to form my own judgment, of course, but I'm struck that the wargaming community often seems to treat the complicated story of MC '02 much like we accuse sponsors of misusing our games -- we pick out the version of the story that best fits our purposes.

Stephen Downes-Martin

One of the problems with an event like MC02 is the presence of large numbers of people and very expensive platforms (think entire Carrier Battle Groups) scheduled for several weeks. One cannot simply end the event on day one if Red or Blue has a cunning plan that ends the

war. Too expensive, and the information sought about the objectives of the event is lost. So, the Carrier is refloated, the nuked city is resurrected, whatever, and the event proceeds to explore what would happen if that cunning plan failed or was not thought of, or to explore the original objectives of the event if the cunning plan simply circumvented those objectives. (The cunning planners should get a BZ however). Then the cunning plan should be explored in detail later.

During the WWII German Army rehearsal wargame on the defense of the area around the Hurtgen Forest, the Americans really did attack. Officers at the game who were in command of the 116th Panzer Division were sent back to their units, and the Game seamlessly transitioned to commanding that division via the 5th Panzer Army chain of command using officers who were held back at the game. Red (i.e. US) moves from field reports were fed to the game players, and Blue (i.e. German) responses were rapidly gamed and the selected responses issued as orders to the officers who shortly before had been at the game. The game became the live planning cell. The German Army took wargaming very seriously starting after WWI when they were restricted in actual forces, and used the technique for officer education, training, character development, input to promotion, and then effectively in combat. Some references: Hoffman, R. M. (General der Infanterie). 1952 "War Gaming", US Army Foreign Military Studies Branch Document P-094; Vego, M. 2012. "German War Gaming", Naval War College Review Vol 65, No 4 (see pages 34-35); Downes-Martin, S. 2019. "How an opponent wargames is an intelligence collection requirement", Intelligence Community Forum, June (See slide 7)

William "Wilf" Owen

I strongly agree with Rex, that MC02 has turned into something of myth and legend. I had lunch with Van Riper just over a year later and he was clear on what vulnerability he had exploited and why he felt it important. At the time I didn't buy it and I still don't and that speaks to a raft of major problems which I am not sure there are "ethics" per se.

If the purpose of a game is to provide insights that can be used in the real world, then what people do in the real world is really important. Assuming referees have some real-world experience or reasonable path to the knowledge, then anyone saying X or Y is "unlikely", "improbable", or "people wouldn't do that..." is critical.

The fact that a wargame model fails, or appears to fail when faced with real-world evidence is often a problem many folks do not wish to address, often in defence of wargaming as a method. That lack of moral courage is an ethics issue. Without getting into embarrassing detail, I've seen a game designed to test formation force structures be conducted with no logistical, maintenance or medical model. - and the game was not designed to include one. The ethical thing to do is to stand up and say that the game is flawed to the point of being useless … or do you hang in there for the insights that may exist?

The predetermined charade can actually be the game itself and the only elements of sanity are those people saying "unlikely", "improbable", or "people wouldn't do that…"

Hiroyasu Akutsu

I agree with Rex as well when he says "we pick out the version of the story that best fits our purposes." This also applies to the Midway Wargame. One of the "unethical" aspects of the game is that in the initial phase the umpire (Vice Admiral Ugaki) minimized the damage on the Blue (Japanese) ships to save the face of the sponsor (Admiral Yamamoto) or to mislead the operational planning. The "reasonable" part of it was that the actual damage on the Blue ships taught a lesson that Blue should avoid such a situation by doing or not doing this, that, and etc.

Antoine Bourguilleau

On the Midway exercice, there is also another problem : Namely that its findings were published by Mitsuo Fuchida, who insisted on the blindness of the Japanese sponsors, but, in the last paragraphs states something that usually goes under the radar, because everybody likes a good story: The referee decided to refloat the two lost CV's because they had been sent to the bottom by B-17 bombers flying from Midway, which the referee felt was very unlikely. But he asked Nagumo's staff whether they thought about the eventuality of a US CV strike from the flanks. The answers showed no real provisions were made and he asked them to think about it - to no avail, obviously.

Decision-Based Evidence Making

Ben Taylor

Sometimes games are set up to prove a point, rather than to explore a problem. We all know that the scientific method is about developing a hypothesis about the real world and then conducting experiments to generate evidence that either proves or disproves the hypothesis. However sometimes the researcher deliberately skews the experiment to generate the evidence that will give them the proof that they seek. This is bad science ethics, but possibly good business since that is what marketing people do all the time anyway - and the ones who are good at it probably make more money than ethical scientists do!

Skewing supposedly objective research to come up with the answer that you first thought of is not a problem unique to wargames. As a 30-year practitioner of operations research in the defence environment the vast majority of my work has been with tools other than games (yes, there really are some!). But if a study including a game is being skewed that potentially results in an unethical game and so I believe is fairly within our scope.

There are (at least) two ways that this can come about.

1. The game designer is complicit in the process and designs a game knowing that there is some inherent bias built in
2. The game designer is manipulated by the study's stakeholders or high-level leader to introduce a bias

The game could end up being the same whether or not the game designer is in on the fix. Here are some examples of things that could happen:

- The scenario chosen for the game is biased in favour of (or against) certain options that are being investigated:
    - If your Army wants to get a long-range precision weapon bought then make sure that the enemy deploys in a flat open terrain where weapons can engage at maximum range. Also make sure that there are no weather rules and that the battle is fought with perfect visibility and ideal conditions for your preferred weapon's seekers.
    - Conversely if you prefer the close-in stuff, make sure that the key battle is set in a rugged jungle during the rainy season thereby hampering every fancy technical widget imaginable (and grounding the Air Force).
    - If the system that you want to look good relies upon networked sensors, make sure that it assumed in the game rules that the network is invulnerable to enemy interference
    - If you want to favour ship-based systems over land based systems make sure that the scenario is set on an enemy-held island archipelago beyond the range of your Air Force's tactical land-based aircraft

    - Conversely if you want to favour land-based systems make sure that the key targets are in a land—locked country out of the Navy's reach (or better yet just define it as an Army and Air Force only scenario and just assume the whole Navy away!)
    - Make the "robust assumption" that the enemy will be organized and equipped with present day equipment, while our people will have the latest and best of everything, if you want our forces to look good. Of course if you want to increase budgets and make the case for a big new project then do it the other way around. The robust assumption is now that the enemy will field all of their futuristics wonder weapons in vast numbers and we'll fight them with our current stuff "to explore potential capability gaps".
    - Allies can be assumed to make up for any capability shortfalls that we may have in our own forces and don't want to have exposed, or they can be assumed to not show up at all if that suits our aims better.

- Make sure that the operating concept favours your preferred solution:
    - Have it written into the scenario that the host nation can provide infinite amounts of fuel if you don't want awkward questions about where your fuel supplies are coming from. Oh, and the host nation's airbases have infinite capacity, as do their seaports, unless of course it suits your objective for them to have none of either - this is called "testing our capabilities in an austere environment".
    - If you can't avoid considering logistics in the land domain and want to pursue a fuel and/or ammunition hungry solution then just assume that there is an infinite stockpile just off the map and that the Navy or Air Force can just deliver them by air to wherever you want them. Far simpler than being constrained by all of those pesky supply vehicles and vulnerable lines of communication - and nobody likes to play logistics in wargames anyway.
    - Future systems can be operated in new ways quite unlike current systems. Aircraft could be operated so as to have pit stops like racing cars - as they pull off the runway on landing the crew jump out one side while a new crew climb in the other side. By the time they're buckled in the plane is refuelled and rearmed and can fly straight into battle again. Time on the ground is just a few minutes. Your new aircraft will be almost permanently over the battlefield just like real ones … never are

- Quantitative performance data is always a good area to influence
    - It is a widely held belief in defence procurement that new technology is always more reliable and simpler to maintain than existing ones. Therefore, new equipment that has yet to be used in real life can always be assumed to be 100% reliable with no down time for maintenance or requirement for spares.
    - It is said that weapons fired on a firing range are an order of magnitude more effective than when used in live exercises, and are in turn an order of magnitude more effective in exercises than when used in actual combat. Naturally our favoured system will be assumed to function like its test range data. Other contenders,

especially those already in service, will use "realistic" data drawn from operational experience

Clearly this is just an illustration of what someone could try on. The list is not supposed to be exhaustive. If your sponsor or helpful stakeholder on a mission smuggles a few of these design assumptions into your game then the players can be as honest and ethical in their play as they like, the fix is already in on the data and assumptions. The players and controllers could be completely unaware of what they are contributing to.

What kind of defences can we provide our ethical game designer with?

Nina Kollars

Ben's point here can't be said too strongly. Decision-based evidence making is alive and well in the Pentagon and is an entire industry in the beltway. My gripe with the return of wargaming as an analytical tool, is that it isn't'. It is fashion dressed like a tool. Finding the most easily misinterpreted, easily mal-designed, and least transparent method to prove to SecDef that their future warfighting concept is valid is most of what I am observing in the shops these days. This is in spite of ethical designers who insist on being transparent about the limitations of wargaming.

When the discourse of client, sponsor, and audience are so deeply rooted in the language of wargaming there exists a tension for the researcher. That's where ethics is important. I'm not saying it is any different from any of the think tank reports out there, everyone has their service-godfather for which they kiss the ring, but the manipulability of the tool and the fact that the Pentagon is a house of politics more often than a house of analysis means that I'm deeply skeptical of any wargame and its findings.

Stephen Downes-Martin

Nina introduces the idea of using language to (a) deceive the target that the wargame is ethical when it is not, or (b) lure the wishful thinking stakeholder into accepting a wargame as credible, or (c) providing an unethical stakeholder with a smokescreen. Examples of language use that alter decision making are contained in the "preference reversal" literature. In general any behavioral anomaly is a useful place to start crafting effective unethical behaviors. For example see Chapter 7 ("Preference Reversals") in "The Winner's Curse: Paradoxes and Anomalies of Economic Life" by Richard H. Thaler (free PDF download of the published article on which chapter 7 is based is at https://pubs.aeaweb.org/doi/pdf/10.1257/jep.4.2.201).

Is Wargaming a Real Profession?

Rex Brynen

Wilf's paper "Unethical Wargaming: Let us be Incompetent!" raises some important issues for professional (war)gaming.

I strongly agree there is a danger in promoting wargaming as a self-contained, self-referential "profession." It's a tool. It's good for some things. It's not so good for others. It is sometimes done well. It is sometimes done poorly. It almost never works well when done in isolation (whether that's the "cycle of research" or effective integration into a broader curriculum). I think Wilf underestimates the degree of self-criticism, however—Yuna Wong is only one of several leading figures warning that we need to be much more rigorous in measuring the utility of our games.

Wilf has more confidence than I do in the relevant experience of military officers. The US last fought a maneuver war twenty years ago, against a second-rate opponent. Most NATO allies haven't fought one since Korea or WWII. No one in NATO, US included, has ever faced an opponent with strong EW/ECM/cyber; responsive, massed indirect fires (with UAVs and other sensors plus SFMs for effect); modern AD2/AD; widespread PGMs; modern fighter aircraft; modern surface and subsurface combatants; etc. This isn't to say that staff and command and planning expertise isn't also highly valuable, because it is (and absolutely essential to the actual conduct of operations). However there's a reason why geeky OR analysts playing Red often hand officers playing Blue their a** on a plate: it's not because they're exploiting game mechanics but because they have far more insight into how these systems might affect future battlefields.

(I'll also add a plug here for diversity and thinking outside the box because you haven't been socialized into box-like thinking. The reason why WATU worked was not only the few experienced RN officers it had, but also because of the young WRENS who had never been to sea on a submarine but could apply rigorous logic and original thinking to U-Boat tactics and ASW.)

Finally, don't get me started on the degree of insight that senior officers have into strategy, or more precisely POL-MIL strategy. While some are indeed very insightful, I'm frequently appalled at how little understanding others have of how and why governments pursue goals, how this is embedded in local politics, the decision-processes involved (especially outside their own countries), or the potential second and third order consequences of military actions. Let us all remember that Wesley ("let's use military force to stop the Russians landing in Pristina") Clark was SACEUR.

Unethical wargames also exist, however, where the game places excessive confidence in "military experience" and fails to adequately question whether that experience is sound or relevant.

Ed McGrady

I disagree with Rex. Most professions are "tools" in one form or another. Medical, engineering, law, they are topic independent tools that are used (usually under a set of standards and wickets, but not always) to accomplish something specific. Medical professionals do a lot of different things. So do lawyers (not to really compare us to either...).

In that sense I would claim that Game Design is a profession, although it is not well regulated, and it does not have standards. In the same way that Graphic Design is a profession. Or screenwriting. In fact I'd say that analogy is pretty close. You can take the narrow view that professions require a license, but then that would leave out Professors and I certainly wouldn't want to do that!

To hit the nail one more time, here is the random Internet definition of "profession": "a paid occupation, especially one that involves prolonged training and a formal qualification." While we don't have degrees in this (yet, though some places do but mostly video focused) but we do have Phil's rather useful set of definitions for experience as a game designer. Certainly the training is prolonged, and we occasionally get paid (though not much...).

Patrick Garrett

I want to footstomp one of Rex's points of disagreement with Wilf regarding the experience of military officers or of "expert" players. In addition to his observations, I'd really like to focus on "Red" or however we want to characterize the foreign adversary depicted in the game. Whether it is the Soviet Union or more modern adversaries, the Red Team will often play Red imperfectly. Even with the most experienced and expert area specialists or military analysts covering that adversary, they almost certainly have profound gaps in understanding the doctrine, military capabilities, organizational culture, and command personalities (to name a few) that a real "Red" knows intuitively … or has discarded.

It's also worth remembering that not every Red expert will play the same as another...even if they are from the same organization but especially if they are from different ones. Analytic disagreements are natural...and while those differences could be very small, it can affect gameplay and maybe the insights drawn from the experience. I am really taken with Rex's observation about WATU and think it worth reflecting on how things might have gone if WRNS members of the team...hadn't been part of the effort. I for one think that their willingness to consider new ideas contributed to WATU's success.

But bringing it back to the above on Red...my thought is that the more we can conduct several iterations of a game, using different players...people from different backgrounds, different experiences, different degrees of expertise...the more likely we are to identify valuable insights.

Zoltán Harangi-Tóth

I think wargaming is a profession since people get paid for designing and conducting wargames. It is a tool in the hand of a unit staff to measure COA's, but it is a profession for the designer. It has some degree of standards, but still lacks some academic underpinning.

Also, I don't agree with Rex when he writes: *"Unethical wargames exist where wargamers unethically prioritize their own hobby experiences or pet design approaches at the cost of the utility of the game."* I think it is unethical if it is intentional, and simply unprofessional if it's unintentional (if we accept that game design is a profession).

Rex writes *"Unethical wargames also exist, however, where the game places excessive confidence in "military experience" and fails to adequately question whether that experience is sound or relevant."* Well it goes back directly to the core of the military: command is an art, since (among others) it is strongly built on the experience of the commander. Experience is not directly the opposite of creative thinking! The level of "mental rigidity" depends on the person(ality) and how he/she was trained and socialized. Whatever the result of the decision, the commander should know how to make confident decisions fast - but the execution still can fail for numerous reasons.

Thinking outside of the box directly goes against the normally pattern heavy military thinking - decision making in a time constrained environment is often characterized by patterns and standard answers to certain problems - at least at the beginning of the process (at the tactical level, these are called drills). At least at the operational level, military leaders should be trained to achieve a certain level of creativity (mental agility?), not just critical thinking. Wargaming is a good educational tool for that.

Ed McGrady

The military always thinks inside the box when it is in garrison. When stuff starts happening is when the officers who are unconventional, and the unconventional ideas, come into their own. Gaming, by somewhat duplicating that operational environment, gives officers the free pass to be that unconventional officer that they can't be in garrison.

Rex Brynen

I accept the point made by Ed and Zoltán that there are wargaming professionals or professional wargamers -- I wrote imprecisely. I was emphasizing the "self-contained, self-referential" part. My concern is that there are risks in becoming an insular guild. Regarding "unethical," there's a fuzzy line between honest mistakes and carelessness.

Stephen Downes-Martin

On the topic of what does "professional" mean. A profession is **not** a tool or selection of tools. Professions use tools, but a profession has three characteristics:

1. It has an established and accepted official body of documented knowledge.
2. It has a written code of ethics.
3. has a governing body that licenses and admits members to the profession, and expels those that behave incompetently or break the code of ethics.

Examples are the medical profession, the profession of arms, financial advisors. However "Wargaming" is *not* a profession. We satisfy 1, we have the beginnings of 2, but thankfully we don't have 3. Nevertheless we choose to "behave professionally", we know who does not and we avoid those. So we refer to "professional wargamers". Much of the complaints we have internally about "bad wargames" (and we are always complaining about them) occurs precisely because we are technically *not* a profession and so cannot label and evict incompetent designers or organizations from the DOD marketplace. We know how to behave professionally because we belong to one or more *other* professions, or because we have established professional level credentials through performance.

So there are lots of gray areas and overlap between incompetent, unprofessional and unethical. We are focused on being deliberately unethical in order to gain an advantage. This implies a high level of competency at being unethical (in order to effectively gain the advantage) and enough competency at wargaming (in order to win contracts), and it removes accidents, mistakes, and incompetence from the discussion.

Ed McGrady

There are always multiple ways to use a word. "Professional" in common use refers to everyone from engineers to lawyers to graphic designers. And Professors. No one gets kicked out of the graphic design or professor "profession" because they screw something up. There are standards, some of which are obvious, some of which are not. I agree that we don't have standards for what a good game is, but I would also contend other professions face this challenge and don't just give up and declare themselves tools.

In specific use "professional" would refer to someone who is licensed to practice and enjoys certain legal protections regarding liability. (e.g. a Professional Engineer). We are not that. But neither are professors. Or graphic artists.

I continually fail to understand why we beat our profession up when we should be busy identifying what bugs us about "bad" games and calling them out. Instead we get all wrapped up with repeatability, social science, and real science and ignore the aesthetic values that we seem to be actually applying.

Hiroyasu Akutsu

On the point of Stephen's saying that "Wargaming" is *not* a profession, I have a simple question. MORS provides wargaming courses and provides a wargaming certificate. I came across a graduate student at Connections UK in 2018 who was about to complete his PhD in wargaming and I played his cyber board wargame which he said was for a playtest for his dissertation. I enjoyed listening to a presentation by one of Rex's students at Connections North in 2019 who told me that he wishes to establish his career as a wargamer (or wargaming "professional") at DND. Are these efforts not for making wargaming a profession? If not, what are they for? This may prove Stephen's saying that wargaming is not a profession at least for now. But I came across several wargame "professionals" who work as a wargaming consultant at consultant companies in the US.

Exploit Hobby Game Mechanics

Stephen Downes-Martin

Wilf's paper "Unethical Wargaming: Let us be Incompetent!" makes two controversial claims that must be addressed if we believe hobby games can provide insight to military professionals. First he asks "does any professional make decisions based on terrain as represented by a Hex-map?" and answers "never, and it can only mislead, yet hex maps are not unheard of in professional wargames, and yet people have claimed professional insights from hex-map games." He goes on to ask "Does any commander base their decisions on their experience of a dice-role modifier (DRM)?" and answers "No, they do not, nor have they based their understanding of manoeuvre on what the "rules" tell them." He concludes "If real-life military experience and knowledge is not the primary enabler for someone to compete in a professional wargame, then there is a problem."

If he is correct, then I believe he is arguing that basing professional military wargames on hobby techniques is fundamentally unethical in that it is promoting a personal hobby for pay while placing lives at risk. He argues that one cause of this is that many hobby gamers do not have the required military experience and skill to know when a hobby technique is applicable and when it is not -- I believe this overlaps arguments by Ed McGrady and Jon Compton that too many wargames do not properly address military sponsors' real-world military problems.

So the question for us is under what circumstances is Wilf correct, or incorrect, and what to do about it?

Zoltán Harangi-Tóth

I don't think that argument is new.

Military professionals don't count DRM's, but weight decisions based on risk assessment, assumption, analysis, etc. The DRM is an abstract way to model risk mitigation, and historical planning ratios, etc. It means they are to represent a lot of factors (results of extensive staff work) in a SINGLE number, to get a brief look at what could happen in an engagement. Counting DRM's is the equivalent of asking staff members about their assessment, opinion, running COA wargames, briefs, etc. If you want to test that in real life run a staff ex, and you will get all the details.

The importance of this is that DRM's present a fast and easy first picture of a future situation, that participant decision makers can understand in a second - if the wargame design is right, the DRM's, and the result of the engagement will be close to real life. The method is irrelevant, the decision-result pair is what is important at some levels. Again, if you want a proper method too, you need staffs to do their job instead of the abstraction.

The problem is that COTS wargames are balanced in a way that all players have the same chance to win a game (symmetric or asymmetric), instead of real life situations. This can be handled by modifying them to the needs of PME, staffs, etc. I think this is the missing link here. For military planning, you run COA wargames -- never hex+DRM styled wargames. So real world (operational) decisions are based on a specific kind of wargaming (and its results) not in COTS style abstract wargames.

About hexagonal maps - it is the same. It is just a platform, it is irrelevant how you count distance, range, etc. Measuring on the map, in a computer simulation or counting grids. The main problem with hexes is that they give a "gamey" feel to wargames, and sometimes make professional wargames hard to sell - to players and to sponsors. It is the decision of the designer what to use - based on his/her experience, the participants, the results, etc.

Also worthy of note, the facilitators role is very important if playing a COTS wargame with inexperienced participants - he/she has to translate every aspect to their level of understanding, to get them "into" the scenario, and that might seem odd at first. The AAR is also more important to better understand the result. Still, I feel that with young people (especially in former education) (M)COTS wargames are the best way to introduce them to wargaming.

So I think it is important to separate the aim of the game - is it to simulate something? Or is it to get some insight? Is it tactical or strategic? Or educational for BSc or MSc students? The list goes on. The platform (map, hex, counter, mini, computer,) all serves to achieve the aim - and it is not the reason/driver of why we play, it is just to physically represent what we are doing (what Peter Perla calls "the instrumentality" of the game). Discovery type, strategic wargames, or classroom wargames are certainly based on the instrumentality. You learn from the result, you don't teach/learn game mechanics (well except at game design courses)! However a proper AAR could fix it if something goes wrong.

My problem with civilian wargames is more about casualties - they are totally insensitive to casualties, so they play wargames to win (and there's an ENDEX after each game), and never deal with force protection, collateral damage, etc.

John Curry

During the Cold War, the BAOR played many wargames using the 1 km grid squares on standard maps to define movement, engagement ranges, retreat routes etc. It was sufficient to note the approximate position of the battlegroup for the purposes of these games. Rolling forward to today, many UK wargames still use the grids on maps as a shorthand summary of units location, for movement, engagement ranges etc. However, hexes are in many ways superior. I have seen them used in recent times for professional wargames in the UK; twice for airborne games, six times for land warfare games and once for naval by the maritime warfare

centre in the UK. So hex based maps are used in the UK, despite the abstraction and generalisation they introduce, and are satisfactory for many types of professional wargames.

CRT modifiers (DRM). When playing a game without an umpire with current combat experience, they are a shorthand generalisation of key factors. If you look in military aide memoires, they are full of lists of factors for different types of military operations. e.g. advance, attack, bridge demolition, etc. CRT modifiers are a way of generalising these factors sufficient to convey their average importance in standard military situations for the purposes of a wargame. e.g. if defending a hill + 1, if dug in for 24 hours +1, IDF fire support on call +2. Of course, in reality the relative weight of these factors will vary, e.g. if a unit has no defense stores (wire, mines, etc) then they will not gain as much advantage as a unit that does have these.

In the UK hex maps are used professionally from my experience and so are CRT modifiers. They are part of the necessary abstraction and generalisations essential to make wargames work.

William "Wilf" Owen

John, I am well aware the UK used HEX Wargames. I've been on those games (see my other post on Rex's thread). You wrote *"In the UK hex maps are used from my experience and so are CRT modifiers. They are part of the necessary abstraction and generalisations essential to make wargames work."* I disagree. I think there are better ways. I think HEX is just cheap, quick and not very accurate. I also think there is little objective evidence to suggest that it produces a safe quality of insights.

The Divisional Wargame did indeed use 1km boxes but was overlaid with a ton of information relevant to the actual ground which was well known to the user who was usually a real Divisional Commander who exercised his Division on that ground 1-2 times a year and did at least one other major CPX. Plus the "game" was a minor part of a far wider and well-resourced process including primitive computers and it was VERY slow. We wouldn't use that system today or rather we should not and how safe and accurate was the work it did had that been taken in isolation, versus the ton of other information that fed it?

Zoltán Harangi-Tóth

We once used Next War Poland to get insight - how? It was to test the applicability of manual wargames and possible COTS games at the strategic level - it was inspired by the USMC War College events + Rand's 2014 wargames about the Baltics. It was quite fun, and was part of an educational experiment, so no taxpayer money was spent, although there were some serious lessons learned! We gave large maps to theater level commanders to plan operations, and DRAW them on acetate, just like in the old days (EW/Cyber environment - really, it was a cheap

and unclassified event), they felt quite nostalgic and they surprisingly enjoyed the session. They gave their orders to liaison officers, and we played the turns and phases, and calculated losses, movement by using the NWP game (89$ at the time). It was surprisingly accurate (based on the AAR comments by the training audience - almost all of them had high profile NATO exercise experiences), the small details were inaccurate, but the whole outcome was more or less realistic. Yeah, it surprised me too.

So how did we use a hex & counter wargame for professional wargaming? Like a pre-made black box, in a “kriegspielish” way. The participants were aware of what we did, since at day 1 we introduced them to NWP (so it was not unethical at all), but they did not play the game, the wargame staff translated their map based decisions and concepts to the game, and rolled the dice. And that was all - we just shared some pictures at the AAR about what they saw and what the staff had seen during the game. At the end, some of them, previously not familiar with that battlefield, got some insight about the context of the terrain, capability, infrastructure, the possible level of attrition, importance of specific LOC's, etc, so it was to get a broad picture, and not "real analysis". After the event all of them filled a questionnaire, and it was clear that they were aware it was just a game, and not more - and agreed that it still was useful to create some insight and familiarization with the topic.

I agree with Stephen that it is highly unethical to sell a hobby wargame as the Holy Grail, but I deny that they don't have any use. It depends on the aim of the wargame, and this is where the ones in charge of the game have their responsibility. Also, the AAR is to understand what is important from the results, and to fix the possible misinterpretation/misunderstandings. I think it is sometimes unintentional to not correct those mistakes - especially if someone uses his/her own favourite hobby game, and it looks well designed from the look of the components, map, etc. Honestly, I really can't imagine a situation when someone intentionally uses a hobby game to deceive professional decision makers. I heard about people playing A Distant Plain in Afghanistan, but never heard that they tried to replicate the events in real life.

So I think it is totally viable to use COTS games in a professional environment, but not in the way the commercial designers created them. Do the participants have to deal with hexes? Why not - a hex map, although pretty general, is literally a (Combined) Obstacle Overlay, since it represents the most important terrain features, evaluated by the designer. Also, it depends on the level of the game - yeah, squads need a VR environment to train, but for division or above the map is fine and cheap, does not matter if it is hex or not.

It is also worth noting that Warsaw Pact planning was very math heavy: every piece of equipment (as far as I remember, for us the T-55 tank was the equivalent of 1, while Leopard 2 versions were 2+) every unit (morale, training, C2, etc simply existed as a multipliers) had a

number to represent its strength -- a single number just like on a wargame counter! The planner used those numbers to count the ammo, strikes, etc. to cause a certain level of loss to that unit - and if approved, they simply stockpiled the forces/material to conduct the mission. Sounds like a DRM/CRT, isn't it? Well, at the end they did not roll the dice, but they had another table to assess the losses.

We planned to run a multiple level (Army-Division and Division-Brigade - The higher level staff will do the planning on real maps, the lower level staff will play the wargame at the game map) experimental event with MSc student classes to test TRL Game's "Less Than 60 Miles C3" games applicability to PME. Sadly COVID postponed it for next year.

William "Wilf" Owen

OK, so let's look at Next War Poland. It uses 12km Hex, and 3-day game turns yet models down to Battalion/Unit. Sorry, but there is so much being abstracted there I would really worry as to how safe insights are being produced. If the outcomes seemed realistic, how was that judged? Were they realistic? If you got insights, how safe were those insights? Could those same insights have come from simply looking at maps and extant OA and Staff Data, thus be safer and more supportable? I merely ask the rhetorical question. From my perspective this is not about how well Hex works or does not. The fact that Hex is as popular and widely used is to my mind a symptom of a far wider problem which is really about cheap, simple and "fun" versus provably fit for purpose. We may be talking about Hex, but this is not really about Hex.

Ed McGrady

I totally agree that hobby games are not applicable to modern joint and integrated operations. They cannot possibly simulate the complex C3, decision-making, and ISR stacks that you really need to simulate to get anywhere close to an actual modern battlespace. Whether that's peer warfare or you are delivering meds to tsunami survivors. The modern, US/NATO, battlespace is dominated by a complex set of organizational, procedural, and C3 machinery that most hobby games, based more or less on a WWII model, totally ignore. And why wouldn't they? They are built for entertainment, while running an ISR stack against enemy signatures is neither for the faint of heart nor terribly compelling as a game. (Trust me...it isn't).

In fact I'm about to launch (somewhere, sometime) my own call out that we totally ignore the most important factors (say, Congress, say ROE, say information) in a rush toward the kinetic. That is a huge mistake in most games since what really dominates the decision-making, Pk or whether you are going to get into trouble? Yes, including it is hard for multiform reasons, but we are professionals so we get it done.

Now, when you are talking about specific *mechanics* I have to utterly disagree. Hexes simply provide an easy way to regulate relationships on a map. You can count them and see what the (approximate) distances are. That is a mechanical play aid that does little to shape the overall run of the game. Likewise DRMs, but if you are giving the players specific DRMs you are already so lost in your representation of the modern, joint, battlespace that you have other problems. Like thinking such a complex set of machinery can be represented by relatively simple rules. These are tools that designers deploy, for good or ill, and are not inherently somehow "evil" or "misleading". Just tools.

I believe the underlying argument is that we don't represent the decision-making at the echelon level we are placing the players at. Of course we don't. Most fleet echelon commands will receive *thousands* of inputs from intel, operational, logistics, C3, headquarters, and other sources each DAY during a conflict (if not each hour). Sorting through and managing that information, and making the large number of necessary decisions, is the role of the commander at that echelon. We can't possibly simulate that in a *game*, that is what a CPX is for. In a game we decide where the focus will be, and what decisions we want to place before the commander. Most of these, logically, involve the "fun ones" of resource allocation, commanders focus and intent, and maneuver/combat. I don't want to make the JFACC player review mission plans for B-52 strikes launching over Germany because, while they would do so in great detail in the real world, I don't expect my players to be playing 24 hours a day for weeks at a time. Instead they just go: launch this here at that target there. But *even at that level* there is a *lot* of stuff to take into account, from the dynamics of time (interestingly we pick on hexes but not what a turn represents) to the entire SPACE/CYBER/C3/EMMW environment that the launch will occur in. Hobby games have no clue how such things actually work, which makes hobby designers totally unfit to design such games. Which is why I have raised, on multiple occasions, that perhaps Connections is too close to the hobby community (there are other reasons too). But, this has nothing to do with hexes or DRMs. Or dice. Or other mechanical aspects to our games.

So, I agree that hobby gamers probably don't understand operational level command requirements in a joint/coalition operating environment, much less those requirements pushed into 2035/2050. But why would they. They are not experienced military operations analysts who have been to war, sailed on ships, and dealt with real world operations. That's why you need the unicorn I always talk about: someone with hobby game experience, who is a good military operations analyst (at CNA we'd distinguish between "analysis of operations" and "operations analysis" one is something you do, the other is a profession), and who can actually behave themselves in public. A unicorn. Lacking the unicorn you are going to get some unfortunate games.

I don't think this is an ethical problem. Lack of understanding is not an ethical problem, it's a problem that needs to be fixed through education, consideration, and better understanding. At least in my definition of "ethical". It's like saying "it's unethical to use a hatchet because of Lizzy Borden". Sure I can give you 20 whacks, but I can also use it to make firewood. It's a tool. Use the mechanics correctly for the problem that is in front of you. The issue we seem to be discussing is a lack of understanding of the problem - the joint operational environment - which is understandable given the difficulty most people who have not actually worked inside it have understanding it (and many who do work inside it have the same challenges...).

Rex Brynen

I also agree that hobby game mechanisms are often inappropriately applied to serious wargames, leading to a situation where knowing how to leverage zones of control (ZOCs) or the quirks of the combat results table (CRT) is a primary determinant of victory. Leaving my former career as a MacDonald's employee aside, Wilf is right that the ability of civilian hobbyists to beat military professionals in a wargame is irrelevant if the game is nothing like the real conflict and the skill-set needed to win is different too.

That being said, there is nothing inherently problematic about hexes or DRMs. A well-designed hex map is simply a summary of geographic and route information, and is not inherently more misleading than a contour map or overhead imagery. At the tactical level, it makes sense to game with terrain displays that are familiar and transfer well to real-world planning and maneuver. At higher levels, less so. DRMs are simply an indicator of the magnitude of causal effect, and I suspect any officer who has trouble processing how adjustments to probability affect the balance of risks has bigger problems with Bayesian updating and information processing. There is a reason why Clausewitz thought the best analogy to warfare was a game of cards.

(Side note here: there is robust evidence that understanding probability and updating is strongly associated with the ability to forecast outcomes. There is also some evidence that military officers do not come equipped with these skills -- indeed, that there is a recruitment bias against them -- and they have to be trained.)

The really key question is whether the game design is meeting the analytical or training objective, and how we know that. Unethical wargames exist where wargamers unethically prioritize their own hobby experiences or pet design approaches at the cost of the utility of the game.

William "Wilf" Owen

I am going to limit my response here to the HEX issue, otherwise, the replies will get too long and unmanageable. I'll deal with the military experience bit in another post.

My work is about developing, equipping, training and thus understanding land warfare formations in the 21st Century. I write from that perspective. I am not a "Wargamer." I use "Wargaming" as a tool. This is what I actually wrote *" does any professional make decisions based on terrain as represented by a Hex-map? The answer is never, and it can only mislead, yet hex maps are not unheard of in professional wargames, and yet people have claimed professional insights from hex-map games."* So given that was the extent of what I wrote, how "wrong" was I? I merely asked the question. The issue for me is not fast, cheap or accurate, but good insights supportable/verifiable by evidence or bad insight that is not. If you opt for bad, then that has ethical implications. If HEX games are supposed to be fast and cheap, how can they be accurate? Not accurate means not useful. (accurate/insightful?). Yes, all games' models are "approximations of reality" At what point does the approximation become misleading? Yet there is stuff out there on the net where men with Government paychecks claim real-world insights from two-player "Next War Poland" type HEX games.

It is my experience that HEX games produce unsafe, unverifiable insights. For example, one UK force development game involving some 12-15 people in the room concluded in the report that X-vehicle was a poor choice for reconnaissance. That is not something a HEX game can do with any safety. It is literally impossible to make that judgement accurately even if the vehicle is the major equipment in a formation. What the report meant to say was that the vehicle failed in the game given how the rules were used. That has almost no implications for the real world. I think HEX is an unsafe tool compared to other methods.

I'll ask four simple questions as concerns HEX and also allied DRM.

1. How do you obtain safe and usable combat power numbers for a counter representing a Russian BMP-3 Battalion or even sub-unit?
    a. How do you reduce combat power to a single number?
    b. How is that number different for say BTR-90A?
2. If I am investigating Brigade structures what size of HEX is optimal?
    a. What produces the best insights?
    b. How many terrain HEX types do I need to get those insights?
3. How do you calculate the DRM used for a sub-unit being dug in woodland versus being in a prepared defence in "urban"?

… and those are all things I have seen crop up in government-funded Force Development work that used HEX games.

Rex Brynen

It's ironic I'm answering this since I almost never use hexes in a game design for many of the reasons Wilf suggests, but here goes!

1) If you're modelling a BMP- or BTR-equipped motor rifle battalion at one counter/one combat factor you are (or should be) conducting games at the divisional or corps level. I would suggest that there is not a lot of evidence from high intensity campaigns that divisional or corps commanders fighting actual wars spend a lot of time considering the APC/IFV configuration of Red infantry battalions. Instead, they think more holistically about the relative combat power of the formations involved (including not only equipment but training, terrain, losses and fatigue, supply, etc.) In other words, cognitively, they do much the same thing as summarizing relative combat power with a number.

Obviously if you are running a game in which you are trying to determine the relative merits of BMP- and BTR- MR battalions, you're not using a 1 counter/1 battalion/1 combat factor scale. Even when I do this casually (as I did in a series of games last summer) a MR battalion is disaggregated down to the vehicle/fire-team scale, with multiple variables assigned for aspects of mobility, armour, primary and secondary weapons, sensors, fire control, etc. If you're wargaming it digitally you can get even more complex.

On the other hand, if you're running a game for divisional/corps commanders, you shouldn't be bogging them down with technical details of the capabilities of a BPK-3-42 vs BPK-M.

2) This is exactly the same question as "what map/scale should I give them" if you're not using hexes. I agree that if you're running an evaluation of recce platforms, capabilities, and organization you shouldn't be using hexes. You also shouldn't just be using maps, either—they're often fundamentally misleading about visibility and line of sight because of vegetation and ground clutter. (In more casual games, we've been known to use Google Earth and Street View to adjudicate actual line of sight for ISR purposes, and it's highlighted how misleading maps and even overhead imagery can be.)

The Divisional Wargame did indeed use 1km boxes but was overlaid with a ton of information relevant to the actual ground which was well known to the user who was usually a real Divisional Commander who exercised his Division on that ground 1-2 times a year and did at least one other major CPX. Oh, that could be a problem. If we ever fight the Russians in the Baltics it certainly won't be over terrain that the divisional commander knows well and on which he or she has exercised several times….

3) A DRM is simply a measure of "how much difference does X make" and parallels the way people cognitively make decisions. Certainly, most DRMs are rough estimates, usually on the basis of some expertise. Sometimes they're rooted in serious OR work. Sometimes they're invented out of thin air and should be challenged.

The use of DRMs and for what depends on the game you're running. If you're running a corps/division game with units depicted at the battalion level, all you need is a rough, informed approximation of the impact of hasty defence in woods or prepared defence in an urban area. No divisional commander cares exactly what trees ("The Larch") are where (https://www.youtube.com/watch?v=H0zVsxUbbjM)

Overall, I agree that folks can be too quick to apply familiar/hobby techniques to serious wargaming challenges without asking if they are fit for purpose. I can certainly think of MOD wargames that fall into that trap. Equally, however, neither hexes nor DRMs nor dice are necessarily the wrong answer either. It depends on what you are trying to do.

Jon Compton

The conversation around hobby games is curious. I'm not really aware of many folks who use them in the way described in some of the threads, but as a practitioner (setting aside the discussion over whether it is a "practice") who has often used hobby wargames in the professional setting to look at serious issues that face the US, I think I can add something to the conversation about how they can be used unethically.

First off, when I say I've used hobby games, the first impression that people seem to get is that I pulled a copy of Next War off the shelf and had at it. Frankly, that idea is silly. I'm not aware of, nor have I ever even heard of someone doing something like that to generate real operational or strategic insight. What is useful, and in my view, essential about having a working knowledge of hobby games is not the application of the specific games, but rather as a library of representative systems for otherwise complex processes. If I need to represent force on force interactions at a certain level of analysis with a certain plausibility of adjudication, why should a practitioner reinvent that wheel, when there is a very good chance it's already been represented in a hobby game?

So using a hobby game in practice, at least as it was done in my experience, used no content inside the box other than the rule book, and even then, only a certain portion of it. The utility of a hobby game designer on staff was that such a person could rapidly adapt, strip-down, and reassemble such systems into a new workable game design that we could then apply to whatever our specific problem set was in a relatively short amount of time. That is a very useful application of hobby games in my view and has nothing to do with trying to convince anyone

that playing Squad Leader would teach anyone about tactical combat. Would it be unethical to do something like that? Sure. But I've never seen it.

But what I have seen is something like the following. I once used a double blind system to explore some ISR issues that we lifted right out of hobby wargame designs. The double blind system allowed players to think about how they would use and integrate ISR systems for both defense and for offense. However, there was a problem in the design itself that happens to be one of my personal pet peeves, but in this case did not get addressed. In two instances a very important asset faced a 90% probability of a mission kill. However, the design we used applied a uniform probability distribution (i.e. is rolled a die). Now the statistician in me hates that because in wargames there is not enough die rolling in most instances for the law of large numbers to come into effect. And, sure enough, in both cases a 10 was rolled, and the asset survived both times unscathed.

Now we could have intervened and simply overruled that outcome, but elected not to in order to explore the low probability outcome. But in our report we were very clear that those assets should not have survived that encounter but for very lucky die rolls. However, that was not what observers took away. It came out in the AARs of other observers that those assets were survivable in those circumstances and that the game provided evidence to that effect (neglecting to mention entirely that in both instances the chance of survival were 1 in 10).

I believe that this is a very clear example of unethical behavior with respect to a wargame that, frankly, could have been accounted for in the design simply by using better probability curves than simple uniform distributions. This is an example of how a simple design expedient can lead to unethical behavior.

Stephen Downes-Martin

One approach to using hobby game instrumentalities is to answer the question *"How can one unethically use the gap between operational reality and wargame instrumentalities (DRM, hexes, dice, victory points, whatever) to pervert the decision making of the game's sponsor (whether it is an operational Commander gaming a COA, or a department researching a problem, or .... whatever)"*.

Chess, Poker, Wrestling and Politics

Stephen Downes-Martin

Since chess is syntactically simple (but semantically complicated) it is a great place to start obtaining insight into cheating and algorithms to spot it. I suggest two next steps;

1. identify specific insights from cheating at chess that can (with suitable adaptations) be applied to wargaming, and
2. examine Poker for the same.

Poker is a closed information game that allows deception in the form of bluffing similar to warfare (and thus wargaming). Differences abound of course, table Poker and Chess are (obviously) face to face allowing immediate psychological interaction with a few specified personalities, while in warfare (and wargaming) the decision makers are usually not sitting across a table directly observing each others body language, facial expressions and tone of voice and sometimes have only intel about their opponent decision makers' personalities. There appears to be a large literature about cheating at poker both in print and online.

Online and table competitions seem to have different modes of cheating, and that raises the question whether distributed wargaming introduces different modes of cheating. At some level any seminar-design wargame is distributed since Red and Blue are usually in different rooms while they concoct their cunning plans.

Peter Perla

Physical cheating at "home" games is probably not uncommon. This can range from collusion (why I never, ever play for serious money with a group of strangers who may know each other) to physical manipulation. A colleague of mine refused\s to play poker because his background in close up magic is simply too tempting.

There are so many things going on with technology these days to help people cheat that it is questionable to me whether any casino is free of cheaters, many of them in the pay of the casinos themselves. I know guys in my old poker group who have won a few hundreds of dollars at casinos like the MGM near here but I have never gone despite the temptations. I'm just too paranoid.

I think for your Unethical Wargaming working group purposes the key techniques involve collusion and inside information. Collusion is usually defined as cooperation among two or more players to ensure that they can collectively take loads of money from the outsiders in the game. Inside information is what I term all forms of skulduggery involving one or more players and some other agents not playing at the table. There are many specific forms of each of these but all of them share those core characteristics.

Some quick thoughts off the top of my head as to how they might apply to wargaming.

- An example of Collusion: Two players from different agencies share a common interest in having the wargame support their different—but not opposed—individual positions. During the play of the game they each espouse the other's position, apparently disinterestedly. This is probably most insidious in matrix games, where their actions automatically have double the support enjoyed a priori by other players. This can easily reinforce the tendencies toward self-fulfilling prophecy.
- An example of Inside Information: The wargame designer and a key member of the sponsor's staff are secretly—or not so secretly—working together to push the sponsor to a particular decision. (Of course this works even better when it's the sponsor who is pushing the gamers to produce the desired outcome.) The staff member feeds the gamers the kinds of things the sponsor seems to think will support the desired position. The gamers then build those into the game. Most blatant, of course, are the estimates of systems performance.

I'm sure you can come up with lots of other examples of how these sorts of "cheats" can drive and magnify the usual dangers of advocacy, self-fulfilling prophecy, and pseudo realism.

Anne M. Johnson

Coordination between competition and cooperation is needed across scales in complex systems. For example in professional football - players compete for positions (and pay) but cooperate as a team on game day. Teams compete against each other, but cooperate to maintain the football league. Where does this type of cooperation come into play within the field of wargaming?

Stephen Downes-Martin

Our interest would be when does cooperation turn into collusion. In the football league example, when does cooperation between teams designed to maintain interest in the league become cheating -- such as throwing a match, say, to make the overall season more exciting to spectators?) The discussion on "Kayfabe" and "Nash in Najaf" provide some insights into this kind of cooperation..

Ian Robinson

Based on a history of gameplay algorithms can spot if you are deviating from your past behaviour (cheating). Luke McShane writes in the Spectator magazine "In recent years, forensic statistical analysis of chess games has become a niche area of research. With the right software, one can dissect an entire collection of games, measuring each move against the suggestions of a powerful chess computer. This offers an interesting historical perspective, and a semblance of

objectivity, in comparing players from different eras, like Capablanca and Fischer. It also offers the prospect of a tool to detect cheating. Unmasking a primitive cheat, who copies the computer's favourite move at every turn, is easy. But we can also quantify, for a given strength of player, how far short of that ideal we would expect them to fall over the course of a game." (https://www.spectator.co.uk/article/a-trout-in-the-milk-)

For additional news on cheating in chess see:

- "Chess's cheating crisis: 'paranoia has become the culture'", https://www.theguardian.com/sport/2020/oct/16/chesss-cheating-crisis-paranoia-has-become-the-culture
- "FIDE President address on anti-cheating policies", https://fide.com/news/703

The FIDE President asks:

"Below are the main questions we would like to have your opinion on:

1. Our methods of detection, although very advanced and ever-improving, can't provide a 100% confirmation. In many cases, the probability estimated is higher than the one for DNA tests. Do you believe a statistical algorithm (or a combination of those) giving close to 100% probability of cheating could stand as sufficient grounds for banning a player? If yes - what odds would you find sufficient?
2. Shall FIDE apply sanctions for alleged online violations to over-the-board-play (and vice versa)?
3. Shall we apply sanctions for alleged violations at platforms' own events, and other unofficial online events, to official FIDE online events (and vice versa)?
4. Shall we publish the names of alleged violators after the very first conviction?
5. Shall the violators be punished retroactively, with their prize money, rating and titles been revoked for some period preceding the verdict? And, if yes, how far back should these actions go?
6. What would you consider a reasonable banning period for first-time violators, and for repeat offenders? How strict should the measures be in youth competitions? "

It's a very serious issue for FIDE. For us, you would have to have a track record of games played to see if the player was acting out of character. Perhaps a game script for each player (play like you wish within these boundaries, if you don't know that we know you are out of character) might help here?

See also "Anti-Cheating in Online Chess" by Alex Holowczak, August 2020, *Chess Magazine*, (https://www.newinchess.com/media/wysiwyg/product_pdf/8312.pdf) for an explanation of using player's Z scores.

Stephen Downes-Martin

Perhaps this kind of pattern matching algorithm to detect cheating could be adapted to spotting early when the commander of a military unit in combat has been replaced by another, or when the enemy has changed its concept of operations? The number of variables is hugely larger than in chess, and the amount of data hugely less, but the idea is fascinating.

The whole issue of using algorithms to look at behavior and scoring to detect cheating reminded me about cheating in Sumo wrestling, see the Wikipedia article at: https://en.wikipedia.org/wiki/Match-fixing_in_professional_sumo in which econometrics was used to prove the existence of cheating. Interestingly, the statistics returned to what they should be in the absence of cheating after the articles and book chapters about it were published!

Kristan Wheaton

The McShane article reminded me of a portion from the 1954 novel "Sweet Thursday" by John Steinbeck:

> "Why can't you rig a chess game?" the Patrón asked.
>
> "Oh, you can, you can! Or at least you can rearrange your opponent. Comes to the same thing."

Stephen Downes-Martin

"Rearrange your opponent" is a brilliant description of deception (which is not the same as lying) in which one "rearranges" the beliefs of the opponent in order to elicit a desired behavior. In chess of course rearranging the pieces forces a change in behavior while rearranging beliefs might not result in new behavior.

Paul Vebber

An excellent example of "rearrange your opponent" as a Chess cheat is demonstrated in the Intercontinental Ballistic Missile Variation of the Tennison Gambit: https://youtu.be/E2xNlzsnPCQ

Joel Kurucar

There is no rule in the chess rule book saying you cannot bomb the chess board. Rob Seater and I were talking about a similar case last week. Chess boxing is actually a constrained rule set for Chess. In normal Chess, nothing prohibits me from punching you midway through a move except societal expectations and norms. In Chess boxing, the rules on that are actually very clear -- you can (https://www.rulesofsport.com/sports/chess-boxing.html)

I actually think from Chess and Poker we can start to characterize the fundamental different ways players can cheat. For example, here's a very hot first take:

1. Gain additional information (Poker = witness signaling opponents cards)
2. Gain material advantage (Poker = counterfeit chips or fake cards up sleeves)
3. Alter/Bend rules (Chess = lie about how castling works to your opponent who hasn't heard about it before)
4. Violate unspoken rules/norms (Chess = punch your opponent until they resign)
5. Reward/Incentive Theft (Poker = players/workers stealing from the prize pot)

I would say this is starting to get at some fundamentally different ways to cheat. The idea being, if we can reason about cheating abstractly, maybe we can apply them systematically to wargaming and identify vulnerabilities.

Montana Hunter

I think that Joel's observation that there is no rule in the chess rule book saying you cannot bomb the chess board creates an interesting question for those developing rules for wargames. How specific are you with the rules, and how much freedom do you give your opponent to experiment. More freeform games encourage creativity, but that creativity could result in actions that are highly unlikely/impossible for an actual actor to carry out. On the other hand, too much specificity in developing rules has the potential to remove freedom and creativity from players - potentially hindering the development of unconventional tactics and strategies. This brings us back to more historical examples. A (theoretical) French wargame pre-WWII could have identified the Ardennes as impassable terrain for armour, thus making the German attack through that area technically "cheating". Embracing flexibility (even in rigid rule sets) could be an interesting way to create creativity in wargames that are run multiple times. Perhaps players could be allowed to "break" a single rule in subsequent playthroughs, to test whether the assumptions created by a wargame still hold true in unusual circumstances.

Hiroyasu Akutsu

Reference Robert Seater's paper "Can AI Save Us from Unethical Wargames?", Artificial Intelligence (AI) has been used by professional Chess, Changi (Korean chess), and Shogi (Japanese chess) players to cheat. These cases have raised the issue of unethical practice in professional games. The latter two cases became controversial and were big news in Korea and Japan. On the other hand, AI can be employed to detect and deter unethical practices in wargaming. So there should be both negative and positive aspects of "human-machine teaming" in wargaming. In a most extreme imaginary case, it could be AI vs. AI proxy wars or rat races that are fought between those who try to prevent unethical practice and those who try to cheat. In the meantime, both Korean and Japanese professional chess communities have

created new rules to prohibit the possession of smartphones and the use of other "suspicious" AI-based gadgets during play. If ill-motivated players continue to successfully outplay their opponents and match organizers and sponsors with emerging new AI-based technologies, new rules would have to be added or new technologies be employed to counter the ill-motivated players. Among the younger generation of professional Shogi players in Japan, it is now common to play against an AI and learn how that AI "thinks" to defeat them on their PC. One of the top Shogi professionals in Japan, namely Sota Fujii, is known to play like an AI. How does he think, and how would we know?

Additionally, in the Japanese Professional Shogi "scandal" mentioned above, the suspected player was not punished because the Japanese Shogi Association was unable to determine if he had relied on AI through his smartphone. It was his actions that were problematic, and it could happen that even well-intended/intentioned players with good "playership" or "Shogi spirits" might be tempted to use the power of AI or any new technologies to survive and thrive in such a competitive world. Or ill-intended/intentioned players can pretend to have used such technologies only by accident during play. In this case, it would not be easy to determine their intention unless the suspected players honestly confess.

Some comparative analysis between Chinese, Korean, and Japanese chess (Xiangqi, Chanki, and Shogi) may be useful in terms of their structure, rules, how cheating occurs or does not occur, and how it has (or not) been dealt with.

Similarities between Western and Chinese chess are widely-known, and may also be interesting to compare. This reminds me of an episode of Bob Hardy, a former Marine during WWII and OR analyst at the LTV Corporation, in which a Chinese was surprised that Hardy was able to play Chinese chess despite his inability to speak Chinese while he was in Tianjin. (Thomas B. Allen, War Games (Reading, Berks: Cox & Wyman Ltd, 1987), p. 73)

Robert Seater

I saw a talk by Jennifer Shahade, a top-ranked chess player (and a top-ranked poker player). She said she stopped playing chess online because it was too easy for players to cheat by using an engine. She didn't mind at all that computers were better at chess than her -- after all, cars are faster than humans but that doesn't remove the competitive relevance of racing. However, using a computer was cheating. I share her intuition, despite it feeling quite odd to make it cheating to use a tool. After all, tool use is one of the defining properties of humanity and civilization.

Stephen Downes-Martin

I think it is cheating when a banned tool is used. Using a racing car during a foot race would be fun to watch (!) but probably count as cheating. One way this kind of cheating is done in wargames is to give Blue what the Naval War College calls a "technology victory machine". Blue has GPS enabled cruise missile and modern comms PLUS the concept being explored, while Red has Scud missiles and rotary phones. These are all tools. Hey presto, Blue wins and the super-whamodyne concept clearly is critical and must be funded. I've made this example obvious, but have seen some quite subtle victory machines in use.

Daniel Tyler Brooks

Is metagaming a road to insight or fraud? See the description of "Kayfabe" in professional wrestling -- "in continuous development for more than a century and supports an intricate multi-billion dollar business empire of pure hokum" -- in which competitors "are actually close collaborators who must form a closed system sealed against outsiders" and where (unwelcome) sporting behavior is called "Breaking Kayfabe" (https://www.edge.org/response-detail/11783). This also relates to the "Frames and Ethics in Wargaming" discussion.

Stephen Downes-Martin

What I got out of the Kayfabe paper by Eric Weinstein is that behavior that looks corrupt and breaks explicit laws can have benefits to all the stakeholders -- even the people the laws being broken are meant to protect (for example the consumers in the Kayfabe case who knowingly replace fair gambling opportunities with pure spectacular entertainment). Kayfabe and particularly its variants open up a world of interesting unethical techniques.

Moving from entertainment to warfare ...

This reminds me of the "corruption between players" described by Hank Brightman in "Nash in Najaf: Game Theory and Its Applicability to the Iraqi Conflict", Air & Space Power Journal, Fall 2007, Vol. XX1, No. 3, Pp 35-41 (https://www.files.ethz.ch/isn/120990/fal07.pdf). See also https://www.analytic-education.com/armed-insurgency-threat/.

Note that "corruption" in the Brightman paper does NOT necessarily imply moral corruption, it refers to mathematical corruption of Nash equilibrium. In the Nash in Najaf case the players do the opposite to professional wrestling along two dimensions to end up in a similar place. Unlike Kayfabe (1) the players do not communicate their cooperation, and (2) they reduce their official expected payoffs (mission success) in return for their personal survival.

The Nash in Najaf case has similar characteristics to the tactic used by the British during the Malaya anti-terrorism campaign of "losing" ammunition in the jungle that had been booby-trapped (grenades that would explode immediately when the pin was pulled, rifle

ammunition designed to explode the rifle breech). But not 100% of the ammo. Very rapidly the insurgents knew the ammo was booby trapped. The deception was in how much was booby-trapped. The payoffs explained why this worked. The British believed they would kill more terrorists than the terrorists would kill British troops using the booby trapped ammo, and the terrorists believed it was worth it as they ran out of ammo to use what they knew was booby-trapped ammo. There was no explicit communication, deception was limited to the percentage of ammo that was booby-trapped, and that percentage was altered according to intelligence about terrorist logistics.

Weinstein argues that Kayfabe "gives us potentially the most complete example of the general process" and so how does Nash in Najaf (and other combat examples) fit? I suggest Nash in Najaf (and thus certain behaviors in warfare) does so by first allowing in outsiders, and then reducing the open and close cooperation to implicit signaling along with the killing of people who didn't get the message (resulting in an evolutionary Nash equilibrium, and pretty quickly too).

Hiroyasu Akutsu

Take a look at "Cheating Cheaters: Malware Delivered as Call of Duty Cheats" (https://www.activision.com/cdn/research/cheating_cheaters_final.pdf). From the report's Summary:

> " The video gaming industry is a popular target for various threat actors. Players as well as studios and publishers themselves are at risk for both opportunistic and targeted cyber-attacks – tactics range from leveraging fake APKs of popular mobile games, to compromising accounts for resale. Even APT (Advanced Persistent Threat) actors have been known to target the video gaming industry.
>
> This report will examine a hacking tool being promoted for use against gamers by masquerading as a cheat for Call of Duty: Warzone. This particular tool is considered a dropper, a piece of malware that is used to install or deliver an additional payload, such as credential stealing malware, on a target system or device. A dropper is a means to an end, rather than the end itself – but still is a critical link in the chain. The dropper examined in this report, "Cod Dropper v0.1", can be customized to install other, more destructive, malware onto the targets' machines."

Stephen Downes-Martin

Does this form of cheating belong in the same class as the Nigerian banking scam? I.e. appealing to someone's dishonest inclinations in order to cheat them? That raises the question "is it unethical to cheat a cheater?" Which raises the question of motive. Cheating the cheater for personal gain versus to punish the cheater. If we define cheating as breaking mutually agreed on rules (for example game design rules for wargames, or international norms for real war) then deception and trickery are probably not cheating. But what we learn about human behavior from dishonest behavior can guide us in crafting deceptive practices in games and war that rely on human moral weaknesses.

Hiroyasu Akutsu

Yes, you can say that the "cheating cheaters" approach has similarities with the Nigerian Banking scam. In this sense, this is only about cheating or deception that is totally acceptable in a wargame as a tool or choice of action for the players or controller to employ in the game.

If this kind of cyber-attack (or perhaps any other type intended to steal information) is used against the servers or computers that contain the wargame rules or materials by the potential players to steal information about them before the wargames are conducted, then this is an unethical practice by the potential players. This is like a Kobayashi Maru situation (https://en.wikipedia.org/wiki/Kobayashi_Maru) in a geographically-distant distributed wargame in which a player steals some essential information about the wargame he or she is not supposed to know beforehand.

Paul Vebber

Here's a fascinating take on Q-Anon as a "mirror of a game" - an activity where "players" are lead deliberately down a rabbit hole so that rather than finding themselves in the "magic circle" of game play and experience immersion in imaginative ways, it leads them into believing that it is *reality* that is fake and that the truth about what is real is being revealed as epiphany by a puppet master in the truest sense (Berkowitz, R. 2020. "A Gamer Designer's Analysis of QAnon: Playing with reality", September 30)
https://medium.com/curiouserinstitute/a-game-designers-analysis-of-qanon-580972548be5

Daniel Tyler Brooks

My social science brain has been watching the 2020 U.S. Election as an example of how the concept of an Edge, Corner, or Boundary case that "breaks" the rules of the game is what strategic thinking is. "The Electoral College result should be ignored, the members should faithlessly vote towards the will of the popular vote." vs "Hey, I based my strategy on the existence of the Electoral College, and you didn't have a problem playing by that rule until you

lost" from 2016. In 2020, we are not just in a corner case of all four corners, but all eight (including the ceiling of a cubed room). Just take the mathematical possibility that the winner of the Electoral College could not only lose the popular vote, but also die of COID-19 after November the 3rd but before December 7th (the day the EC votes). I doubt even https://fivethirtyeight.com/ has calculated the probability of that in a decision tree for a voter. But the possibility is statistically significant.

Phillip Reiman

The big idea in dirty politics - as opposed to idealized political systems - is legitimacy. How much can a candidate or party 'get away with' and retain legitimacy in the public eye? Too much foul play and you're unable to govern or worse - OTOH, if you take no risks you may be unable to influence anyone and unable to access power but . . . there's your problem. Negative ads are a good example. Notoriously reflect badly on the proponent *as well as* the target. Voter disenfranchisement is another example - do you suppress the vote lawfully in order to get or retain power and if so, at what point does the vote become a sham in the eyes of the public? It's the essence of influence ops.

The history of the "20% off Bed, Bath and Beyond Coupon" makes for fascinating reading, and makes me think that there are no rules other than a pragmatic 'influence somebody' ethos in marketing. Likewise, the movie "Queen's Gambit" is all *about* rules - the types and the transgressions - inside and outside of games.

This video (https://youtu.be/dDYFiq1l5Dg) about "Supreme Court Shenanigans" is a nice example of rule-bending, breaking and stretching. See also https://youtu.be/fLq86hnwqF0 "Fixing Broken Politics" which uses a lot of rules derived from games to counter degenerate strategies as defined in Salen and Zimmerman's "Rules of Play".[72]

Andrew McGee

I come to the topic as a lawyer - both academic and practical - of over 40 years' experience. I began to get a bit twitchy quite early on in this thread when there was a video referring to 'shenanigans' in the process for appointing Supreme Court Justices. That is of course very topical at present. My concern was the apparent suggestion that some people might be doing something wrong by advancing or delaying the confirmation process. Yet there seems no doubt that no one has yet broken any of the rules. In a strict legal context I do not see that anyone can be criticised for operating the rules to their own advantage. To put it bluntly, that is what lawyers do all the time - we study the rules carefully, we look to see how they can work to our

[72] Ed McGrady raised the topic of 'shenanigans' in wargaming in the "In-Stride Adjudication" working group report 2018 generating significant discussion. The Working Group report is on PAXsims at: https://paxsims.files.wordpress.com/2018/09/in-stride-adjudication-working-group-report-20180908.pdf.

client's advantage. Where necessary we search for -and then argue for - the interpretation which will suit our client. And we do that even where we do not like or do not agree with the interpretation for which we have to argue. I could point you to a few reported cases in English Law where courts have, at my urging, adopted views of the law which I think are plain wrong. But all I did was my professional duty to fight for my client.

Now, this is all good clean fun, but what does it have to do with wargaming and in particular unethical wargaming? Having shown myself to be in a professional context, quite happy to resort to 'shenanigans', I will go on to say that I do not at all approach wargaming in this way. I am purely a hobby wargamer, and there is no prospect at all that anything I do in a wargame is going to have any impact on any real-world decision by any state power. But I do regard wargaming as part of a serious study of military history in a broad sense. Although I will read the rules with a lawyer's eye for ambiguities and loopholes (can't help it at this time of my life) it does not follow that I will take advantage of all such problems. Fundamentally, I do not play to win, but to learn. I know that I am not alone in this, though equally I know some players who are entirely driven by the need to win.

A particularly instructive wargaming example is MMP's Operational Combat Series That has been around for nearly 30 years. The Rules have gone through many iterations, and at any given time there is someone, referred to as the 'Series Honcho', who is empowered to rule on any Rules disputes which arise. He is, in effect, a one-man Supreme Court, and players of the system seek and sometimes debate his rulings as assiduously as lawyers consider the work of the Washington Nine. But in addition to rulings on specific points, successive Honchos have maintained a list of actions which are within the letter of the rules, but which have proved controversial as possible abuses. For each such action the list categorises them as 'gamey' or 'not gamey', the clear implication being that players really should not resort to 'gamey' actions, even though they do not contravene the letter of the rules. ,Such actions are, if you like, 'unethical'.

I have come to all this fairly late, having started to play the systems only after it was well-established. But it strikes me as a fascinating example of a system which covers a large range of games and whose dedicated players (who are numerous) take very serious technical legal questions of rule interpretation, but then superimpose on that a notion of ethics. I very much like the way this works, and I find it an interesting example of players looking to develop ethical standards in areas where the rules (may) fall short of complete clarity and perfection in their aims.

Ian Robinson

Robert Conquest's Three Laws of Politics:[73]

1. Everyone is conservative about what he knows best.
2. Any organization not explicitly right-wing sooner or later becomes left-wing.
3. The simplest way to explain the behavior of any bureaucratic organization is to assume that it is controlled by a cabal of its enemies.

John Derbyshire adds this: "Of the Second Law, Conquest gave the Church of England and Amnesty International as examples. Of the Third, he noted that a bureaucracy sometimes actually is controlled by a secret cabal of its enemies — e.g. the postwar British secret service."

John Moore thinks the third law is almost right; it should read "assume that it is controlled by a cabal of the enemies of the stated purpose of that bureaucracy."[74]

Miltion Friedman noted that bureaucratic resource allocation involves spending other people's money on other people, so there are no compelling reasons to control either cost or quality — but a bureaucrat will learn, given time, how to "spend on others" in such a fashion that the primary benefit flows to himself.

[73] https://www.isegoria.net/2008/07/robert-conquests-three-laws-of-politics/
[74] http://www.tinyvital.com/blog/2003/10/26/conquests-second-law-of-politics

Civil Emergency Planning

John Curry

From my experience unethical practice happens more in wargames run for civil emergency planning than for the military.

Stephen Downes-Martin

What have you seen that makes you believe it does?

What kinds of unethical behaviors?

Were they effective (i.e. not detected and rejected)?

Why does it happen more often?

Ed McGrady

One of the most important considerations in doing anything "across the river" (I.e. for non DoD Departments/Agencies) is the difference in liability between the DoD and other parts of government. INAL (I'm not a lawyer) but there are special prohibitions about organizational and individual liability when DoD does something that causes harm (for pretty obvious reasons) while other agencies can get sued. As part of the discovery process a game (or analysis or model or simulation or aggregation of data or lessons learned or log book etc.) can be used as evidence that you should have known not to do whatever it is you just did. That would be bad. Therefore there is a strong desire not to (a) do such things in the first place, (b) not document them when they have to be done, (c) not keep the documents, and (d) have such documentation, if it must be done and kept to be of such anodyne nature that no one can make any case with it. This is not necessarily any more unethical than any other legal gambit, but it can greatly affect the nature of "lessons learned" in agencies and departments outside of DoD (of which emergency response is a subset).

Stephen Downes-Martin

(I'm also not a lawyer!) With regard to the claims about gaming for civil emergency planning that there is a strong desire with regards to legally dangerous issues:

(a) "not to do such things in the first place"

Does this imply active connivance between the game designer and the civil emergency planning sponsor. What are the mechanisms by which this occurs when the sponsor and design team are not part of the same organization, or is it just that the sponsor pays and gets what they want? The latter is a common pressure in DoD games. To be unethical, the civil emergency

planning sponsors and/or the game designers must be intentionally creating deceptive or uninformative games in order to avoid legal problems.

Question: What effective mechanisms can we design to avoid gaming something that needs to be gamed while appearing to game the stuff that needs to be gamed, what mechanisms can we design to avoid detection, and finally what are the warnings and indicators that this is happening?

(b) "not document them when they have to be done", and
(c) "not keep the documents"

To decide when it is advisable not to document something from a game requires legal advice, or a routine protocol of never documenting games. The former lays one open to a charge of conspiracy, so will throwing away documents one should have kept. Someone *will* talk.

Question: What effective mechanisms can we design that will reduce the legal risk to an organization when it decides not to document something or to destroy what it knows are legally controversial documents.

d) "have such documentation, if it must be done and kept to be of such anodyne nature that no one can make any case with it"

If one only sometimes produces anodyne reports then when one does that will immediately raise suspicions. If one always produces anodyne reports why bother funding the games to start with?

Question: What effective mechanisms can we design that maintain a civil emergency planners reputation while producing anodyne reports?

Final Question: What are the visible warnings and indicators that such mechanisms are in place, and how can one deal with them?

Phillip Reiman

(I <u>am</u> a lawyer!) The question as I understand it, is: Do civil government wargames or staff exercises lead to more unethical conduct among the participants than those held by the military? If so, why and how can we mitigate this issue?

Answer: Yes – if this is the question – IMO it is probable that an observer, controller, trainer or facilitator (OC) running a civilian government exercise would encounter more game-rule breaking than in a military setting. The cause would be organizational culture. Because:

1. Civilian Exercises are rare and therefore:
    a. Like kids who haven't played games before, they aren't comfortable with losing in any form – no magic circle for them
    b. They are very concerned about success. It's bad press to say, "Local Government can't stop crisis in large scale test". If the 3 BN of the 4 BDE can't clear an obstacle, the failure is real, but the repercussions are not a matter for public debate.
    c. Their staff is very concerned about success. No one wants to be seen as the weak link that caused 1b. Think "Veep" – they are worried about exterior appearance
    d. Likewise, the rarity of exercises means that there is more at stake – hence the anodyne nature of the reports. "Everyone was above average"
2. Add in that civ exercises are 'catastrophic' events – not 'what we normally do'. Compare this to the Brigade staff that trains to seize an enemy-held objective and then wargames this task. They should be good at it. Sure, exercise might be more complex or involve an opponent, but civilian officials outside of the emergency response crowd, don't really practice for dirty bombs or pandemics. We/they talk about them, but we don't have time to regularly work on those sorts of problems. As a result:
    a. Civ Officials have little or no experience handling complex crises, and so they rely on the closest experts they have – the first responders.
    b. Authority is essentially ceded entirely to first responders and the chain of command is then hyperlocal – this police chief or that fire chief.
    c. Yet, higher level coordination is the key to these exercises. That requires a common operating picture which, in turn, requires a plan to create such and that plan needs to be tailored to the crisis involved. Your pandemic plan probably won't help too much in a dirty bomb situation.
    d. This leads to a *lot* of improv at the intermediate levels – the OCs record this and the reports are generally not flattering to the higher-level institution that organized the exercise. No one likes to read, "You are the problem".
3. Result? They try to cheat.
4. What effective mechanisms can we design that maintain a civil emergency planners' reputation while producing anodyne reports?
    a. Tell them 2a-d before the exercise. Manage expectations.
    b. Exercise design for a very narrow group – just the Command Post or even a portion of the command post allows more controllable interpretations of the result and does not tarnish the reputation of the entire local/state/federal entity. "We have some problems we need to address on our team in short passing, but offense and defense are improving".

5. What effective mechanisms can we design to avoid gaming something that needs to be gamed while appearing to game the stuff that needs to be gamed, what mechanisms can we design to avoid detection, and finally what are the warnings and indicators that this is happening?
    a. IMO – people come in with expectations over what the topic or focus of the game will be and how they'll participate.
    b. In classic deception operations, the OC would reinforce these assumptions – encourage cheat sheets or hint on what sort of crisis plan should review. EX: You're going to play the role of the local police chief in a pandemic in order to evaluate community response times. Better get familiar with XYZ! However --The exercise is designed 'secretly' to test . . . some meta element – such as questionable access to 'answers' or a chance to 'alter the result' of the exercise. In other words, a chance to cheat.
6. What effective mechanisms can we design that will reduce the legal risk to an organization when it decides not to document something or to destroy what it knows are legally controversial documents.
    a. Ah, if an organization is creating documents – even in game – and they are controversial
        i. They need to be dealt with in game or in real life as either
            1. Justified
            2. Unjustified and in error
        ii. Destroying documents to cover up an unjust or 'controversial' decision is simply not good government and quite possibly illegal.
    b. So, let's get hypothetical: a government entity receives a report or generates a report that says a future pandemic causing grave illness XZX attacks people with a BBB genetic code. However, it is readily transmissible by people with AAA genetic codes who are generally asymptomatic. Is this report controversial? Yes, because it could trigger anti-AAA hysteria. Can the Gov hide it? Basically, no – but it CAN address it. What if the Gov decides to segregate AAA from BBB – a risky decision socially – but one that the Gov would eventually have to own. There's no getting around it.
7. What are the visible warnings and indicators that such mechanisms are in place, and how can one deal with them?
    a. You are the magician – as OC, you've got to be aware when the participants are working outside the space you've set up for them. So, these indicators would vary based on your exercise.
    b. The best corrective measure is to ask and then immediately redirect. Are you watching my hands? Cuz that's where the magic is.

Understand Losing Behaviors

Ian Robinson

In the Western Desert, the armour of the British Army were keen at the start of the campaign to 'Balaklaver' - heads down charge at enemy positions, even when defended by anti-tank gun positions. This worked reasonably well against the original enemy, the Italians, but when they tried it on Rommel's DAK, armed with 88mm guns, the results were disastrous.

('Balaklavering' is a euphemism used by the British when attacking a strong defensive enemy position, particularly held by artillery. It's derived from the Battle of Balaclava in 1854, when the Light Brigade misunderstood its orders and charged into a valley flanked and headed by Russian artillery.)

When wargaming this, it takes courage to do exactly as the RTR did, rather than as most wargamers do, and advance but stop just outside 'efficiency range' boundaries, then continue the advance until the next one. This is gamesmanship - unethical behaviour compared to reality based on a too precise knowledge of the rules and the enemies dispositions.

How many wargames are calibrated by historical actions in such a way in the expectation that the same results happen? And who would willingly adopt a strategy such as this at the outset of a game, despite being ordered to do so? One might blanch at this order, consider it 'unethical' and yet orders are orders. Would it be unethical to subvert this strategy in a wargame to preserve yourself?

In the context here, we are considering whether it is unethical to subvert a wargame order that you know would devastate your own command in the expectation that there is only a slight probability of success, which is believed to lead on to greater success for the whole army. It's the wargamer's test of self sacrifice. Is it unethical to disobey these orders, follow self preservation and significantly amend the plan? (It's a measure of whether in a wargame 'Auftragstaktik' (i.e. Mission Tactics[75]) supersedes 'Befehlstaktik' (i.e. order driven tactics). How much initiative can you exercise before you are subverting the game: 'orders are orders', or orders are to be interpreted to achieve an end, even if this means subverting current doctrine?

If you seek to use a wargame to understand likely outcomes in combat, one must test and understand losing behaviours before adopting newer, better ones, based on the imperfection of reality through a set of rules. So, how to overcome the innate will to win and survive and do as they did, and risk the same fate in a wargame?

[75] https://en.wikipedia.org/wiki/Mission-type_tactics

Robert Seater

So, the interesting tidbit I got out of all of that was that it is desirable to force players into losing positions, to build their intuition and mental models of why such positions are bad. If you always win by doing the right thing, you don't actually learn why it is the right thing. So it is responsible (if not actually ethical) to rig games to be unwinnable.

Stephen Downes-Martin

I think it depends on the objective of the game. If the sponsor's objective is "how could one side win (if at all) despite using 'Balaklavering' (for either educational or analytic purposes) then that side would be constrained by the game design to use 'Balaklavering' . On the other hand if the sponsor's objective was "how could the side that used 'Balaklavering' do better by not using 'Balaklavering' in that historic situation" then the side that originally used 'Balaklavering' would be in free play. It seems to me there are three ways of obtaining something like 'Balaklavering' -- how many others are there?

1. Start the game just after 'Balaklavering' has occurred by putting it into the road to war.
2. A variant of #1, inform one side that in move one they WILL use it. I's seen this done in nuclear de-escalation games -- the lead of the cell playing the "first nuclear use" side is recruited to use a nuclear strike in move 1, but has freedom of how, what target, where, timing, yield, etc., to fit within that player's Pol/Mil game objectives.
3. Use some form of "hidden scenario game" approach that encourages the use of it without the players realizing it. Hard to do but if successful the best method as the players remain in the magic circle.

Ian's point of understanding losing behaviors is very interesting and important given the difficulties of recovering from failure. It also fits with the idea that a wargame is most informative about a new concept or behavior when that concept of behavior fails in the game. When this happens we have at least one way the new concept or behavior might fail and we can explore that. If a new concept or behavior does not fail all we know is that under the exact game circumstances and given the exact decisions made by the opposing team and adjudication it participated in success along with everything else (including the opposing side's decisions). Since any wargame is a single trajectory through all possible decisions by all sides we have no idea about the likelihood of success nor how that success might not occur.

Ian's other point of having "a too precise knowledge of the rules and the enemies dispositions" is again a matter of sponsor objectives (eg. "What might happen if Blue knows 'this' about Red?" or vice versa), poor or unethical game design (letting one side know more

than is acceptable for the sponsor objectives or to support the sponsor agenda), and unethical player behavior (conversations between opposing side members in the restrooms).

For both the Inner and Outer Games I see three questions:

1. What other methods of achieving 'Balaklavering' are there?
2. When is 'Balaklavering' unethical?
3. How would one detect and deal with unethical 'Balaklavering' in a game, or design them out of a game?

Ian asks "Would it be unethical to subvert this strategy in a wargame to preserve yourself?" If the sponsor's objective is to explore different ways the 'Balaklavering' might have turned out or to educate officers about the historical results of such 'Balaklavering', then yes it would be unethical for the Blue player to subvert the strategy. On the other hand if the sponsor's objective is to explore how Blue might have fared by *not* using 'Balaklavering' then it would be unethical for the Blue player to *not* subvert the historical strategy.

In a game I observed some time ago the three-star blue player (who was also the sponsor -- but that is a discussion for a different time!) kept on using standard doctrine, which resulted in catastrophic losses. It took us a while to realize he was being neither unethical nor unintelligent. He was very intelligently driving home to his staff the insufficiency of standard doctrine and their experience, and driving them to think hard during the months of post-game activity on how they should fight that particular war. If he had instead used novel tactics or doctrine and won, that would have been both ineffective and hence unethical. He would not have demonstrated the failure of the standard doctrine and thus it would be impossible to drive the change from an untested standard doctrine to a new doctrine cooked up in a single wargame. It all depends on the *real* objective of the sponsor.

Ian Robinson

Reference your comment "In a game I observed some time ago the three-star kept on using standard doctrine, which resulted in catastrophic losses. It took us a while to realize he was being neither unethical nor unintelligent." That's fascinating. Failing in such circumstances teaches you what you don't know and then helps you learn. Far better to do that, rather than play unethically as you suggested above. Outside of professional wargaming, it's rare to replay the same battle and test different scenarios to see what would happen.

Antoine Bourguilleau

Ten years ago at my club I once wanted to replay the battle of the Java Sea, when the Allies faced the so-called "long lance" Japanese torpedoes for the first time. Of course, knowing that, the allies will try to close quickly on the Japanese fleet to avoid being destroyed at long range.

So I proposed to the participants to have a game in the Mediterranean in 1915, with a Turkish fleet meeting a French fleet just outside the Straits. The Turks had double range torpedoes (and the French were not informed of that). When the first torpedoes hit French cruisers, the French were totally disoriented. Thinking they probably ran in a minefield they reduced speed to check. Each time torpedoes were fired by the Turks, I threw one die per torpedo, considering that on a 5+ the French were able to spot the torpedoes trails.

The result was close enough to what I expected: The French players were totally surprised, it took them some time to get their nerves back and try to find a solution in face of the impending disaster. The Turks much enjoyed the game but I must say I chose them carefully from the pool of players that I knew would throw the table away if they were in the French team!

I'm currently working with some friends on a reversed Austerlitz because no Austro-Russian player wants to leave the Pratzen Heights. But if the battlefield is turned upside down and left to right, the French replaced by Austrians and the Austro-Russians by Prussians during the Seven years War and the Austrian right flank exposed …

Ian Robinson

Quite so - it takes a degree of courage to walk willingly into a trap to see what will happen, and for that reason historical battles are always difficult to refight, especially at 'club' level where the sense of pride often prevents that. Personally I think there is plenty of mileage in repeating historical battles and tactics at the start, and allowing the 'decision tree' in each combat exchange to work its way through. This can allow for changes to the battle flow emerging (imagine if the French lost the cavalry battle at Austerlitz for example). The real benefit lies in understanding how quickly you can turn a situation around if it starts to go wrong, or conversely how quickly you can exploit a situation if things go better than expected.

Stephen Downes-Martin

The unethical approach suggested by this thread is to deliberately design a game in which Blue wins (or Blue's concept is successful) in a way that looks credible and hides the weaknesses in Blue. Since people like winning this might not be hard to do. The motives for doing this are *very* high if the game is to compare two concepts, one sponsored by one military community and the other concept sponsored by another military community (for example, hypothetically a Navy C2 concept versus a Marine Corps C2 concept to be used during a Maritime Operation). The shenanigans during the outer game and the inner game, by senior people (who, hypothetically of course, have several stars on their shoulders) who believe themselves honest, would (hypothetically of course) be astonishing.

Ian Robinson

In a nutshell. And, of course, the stakes are high in all senses. The other aspect is the extent to which when you are playing, are you playing their plan or the man? Someone who deliberately subverts or ignores the plan in order to win is (a) showing their initiative by (b) subverting a failing plan. It might be better for all to allow the plan to fail to show how bad it is, rather than paper over the cracks. But that means losing in public, which we are not wired up to do.

Exploit Logical Fallacies

Anne M. Johnson

I came across a YouTube video (https://www.youtube.com/watch?v=Qf03U04rqGQ) that speaks about logical fallacies. Understanding these could contribute to insight of unethical wargaming.

Robert Seater

Here is a proposal for a Meta-Ethical Wargame. Run it like a Matrix Style Game, in which in-game bonuses to die rolls are assigned based on the strength of arguments made by both teams. But add in a rule that you also get an in-game bonus for each logical fallacy they identify in the opposing team's arguments.

You said you have a better position because you were using a tried-and-true C2 structure to produce faster decisions? Well, that's +1 to my die roll for Argumentum ad antiquitatem (appeal to convention) and another +1 for naturalistic fallacy (concluding a value from factual statement, by assuming that faster decisions mean a better position).

A real trouble maker could gain additional in-game advantage by finding flaws in the adjudicators' stated reasoning. You just said that you'll give the enemy a bonus because 2 of the 3 judges agreed? Ha! Argumentum ad numerum (appeal to numeric popularity) gives me +1. If you brought in guest adjudicators, I claim another +1 for Argumentum ad populum (appeal to popular opinion). Even when I lose, I win.

I suspect such a game would teach only a little about military success, but it would get the players to be very attuned to weak logical arguments. I wonder if it would, empirically, produce better decisions or just more sluggishness and hesitation on the part of the players?

Stephen Downes-Martin

Reference the Meta-Ethical Matrix Wargame, there's a game called "The Mind's Lie" by Kris Wheaton (a member of our group) that explicitly deals with biases via analysis of scenarios. I believe it will be published soon. The point I want to make however is that I ran about half a dozen Colonels and Navy Captains through the game a few times. Initially skeptical, they became enthusiastic players and the arguments about what was believable or not about the scenarios based on the information being provided became fascinating. They then started talking about applying the logical arguments, but most importantly using the game mechanic, to analyze documents they had to deal with in their careers. Board Game Geek has a brief description (https://boardgamegeek.com/boardgame/139711/minds-lie).

Daniel Tyler Brooks

There's a difference between a formal and informal fallacy. The only difference between a bias and a heuristic is whether or not it bites you. As it turns out, the invocation of one informal fallacy can often be countered with the claim of the inverse informal fallacy. (My favorite contemporary example being "false equivalence" vs "special pleading") For example, via induction, I've determined that a fast partial answer is often better than a more perfect answer later. Of course, sometimes the fast answer is wrong and the later answer would have been right. I get away with my intuition most of the time because we satisfice solutions rather than optimize. Especially in combat, where being wrong fast can be better than being right slow. It can also be disastrous. Is it better to have lots of cheap tanks, or a few top of the line ones? I dare say, most officers know just enough about logical fallacies to hurt themselves.

Stephen Downes-Martin

We are interested in how to use logical arguments, and the claim that the other person is committing a fallacy, for unethical purposes. Careful working of the fallacy might perhaps be used to fool someone into believing a fallacy had or had not been committed, countering with another informal fallacy even when knowing one is wrong, and so on. The fast partial answer versus a more perfect answer later is a matter of deadlines and depth. Do you mean the unethical behavior is the perpetrator arguing for one of them to lure the target decision maker into a decision that favors the perpetrator?

Daniel Tyler Brooks

I think the logical fallacy discussion is a bit of a red herring, because what you are really interested in are people arguing in "bad faith." In other words, someone making a disingenuous argument for the sake of winning the game. Perhaps the argument is disingenuous because the arguer is aware of a fallacy he neglects to make known, one that in fact causes him to be unpersuaded by his own argument, but is sufficient to persuade his competitors, but the fallacy is incidental. The key point is that a disingenuous argument is an act of intellectual dishonesty. This comes up quite a bit in atheist vs theist debates. When you become aware of an opponent's bias, one way you can win is to nurture it in order to exploit it later. I don't think that's cheating in warfare, but the very hallmark of strategy, and that Sun Tzu would nod in approval. However in a Matrix game, a hyper-competitive player making disingenuous arguments because they want to win can torpedo the research objectives.

Stephen Downes-Martin

If it is possible to exploit any stakeholders' (player, sponsor, sponsor's chain of command, adjudicator) inability to spot a logical fallacy then logical fallacies (as well as psychological and cognitive biases) become a weapon in the hands of the unethical wargamer.

Change the Rules

John Curry

I am interested in cheating where the 'rules' of the 'game' have changed leading to one side being badly defeated. For example:

- Custer at Little Bighorn (1876) - the standard rules of Indian warfare had changed, the cavalry were facing better armed and determined defenders.
- The Battle of Grozny (1994/5) - the Russians faced a mobile urban defence.

Recognising when the other side has changed the military paradigm (i.e. the rules) seems a most interesting challenge to avoid being surprised in the first battle of the next war.

Stephen Downes-Martin

A corollary to that might be deceiving the opponent that the rules have changed when they have not. Is there a historical example of that?

Mark Flanagan

Consider the following historical example:

> You are in the position of a commander of a small coastal port in WW2. One fine morning, as dawn breaks you find yourself under bombardment from an enemy destroyer lying off-shore. The destroyer is at extreme range but confidently the shore defense batteries return fire. The punishing barrage lasts for half an hour and many valuable areas of the port and small ships are damaged. There will be hell to pay. To add insult to injury the destroyer turns its attention to the towns AA batteries and a flight of enemy fighter bombers commences an attack as the destroyer departs unscathed.

What has happened? The town's defenses were not insignificant and the coastal batteries were sighted and ranged to defend against such an attack (aka a destroyer's bombardment). Have the rules of engagement changed? Or do you merely think they have as per your corollary above?

Stephen Downes-Martin

I think there are two types of rule being conflated; physics (how far can a gun fire) and tactics (best allocation of assets). Neither of which changed in the example. What went wrong was a lack of thinking by the town defense planners and whoever was in charge.

The town defenses would naturally be overwhelmed by an attacking destroyer with an accurate range greater than that of the shore defense batteries, and the attack you describe is predictable assuming a smart opponent and assuming intelligence about enemy destroyer

weapons is accurate. The defender failed to spot the approach of the destroyer and failed to launch an aircraft attack on the destroyer before it got within range. (If I was the destroyer commander I would have engaged the ground based AA batteries first before engaging the port and its contents.)

I don't think any rule was changed nor did anyone think one had. I think the attackers exploited a fact of reality -- their guns outranged the town guns -- and the town did not plan for that or did not have the assets or information to effectively defend themselves. Or the town was not important enough to reallocate more assets to its defense and so had to rely on luck -- hoping the town would be ignored , and if not hoping a lucky shot from the shore batteries might save the day.

Mark Flanagan

This was the outline from Operation "Decomposed Two" (pp 170-172 from "Motor Gunboat 658: The Small Boat War in the Mediterranean"'). It was a bombardment of the now Croatian ports of Orebic and Korcula on the Peljesac Peninsula , in 1944 (30th/31st July). The "destroyer" was in fact a decoy, being a mocked up LCI (254) that "had a magnificent canvas at a rakish angle and two large canvas gun turrets". Landing craft had dropped a battery of eight 25 pounders in a remote beach from where they would bombard the towns with a RA FOO in place to correct and direct. At which point the "destroyer" would "let off fireworks which simulated gun flashes, from behind each turret". The operation worked to a charm. The Hurricanes' strafing was additional noise to distract and confuse. The landing craft safely embarked the 25 pounders and no "Red" was ever aware of their existence as they returned undetected.

Yet a rule had seemed to change, even though "the physics (how far can a gun fire)" for the coastal defenses were perfectly adequate to see off a destroyer. You may call it a "shenanigan" (McGrady) or a "wizard-wheeze" (Mouat), and I may have been unethical to lead you down the Barnum path of "what you see in not what you get" with the description. However "Red" is now reconsidering his coastal defenses based on the counterfactual superior gun ranges of RN destroyers in theatre.

Stephen Downes-Martin

To make sure I understand, this is an example of tricking the enemy using an M-type deception op into believing one has more powerful and different location weapons in such a way that the enemy response hides the deception (by itself generating confusion) and is led into defeat.

(Dis)Respect the Player

Robert Seater

Here is a completely different take on the topic. Consider how a designer should think of their players. It is tempting to think of the design role as in service to the player, especially if the game's goal is to provide education or training to the players. But that can lead to poor designs. I like to maintain an opposing slogan – "Maintain a healthy disrespect for the player". I have it written on a sticky note on my wall so I won't accidentally try to be the player's friend.

The idea is that you are not there to make the player's life easier. A player wants to win, for their decisions to all be obvious, and for the theory they had before playing the game to be validated by the game. That is rarely if ever the goal of a professional or serious game. In fact, you want players to sometimes lose, to agonize over decisions, and to learn new perspectives.

Really, this is just the obvious principle of staying focused on your programmatic objectives (or 'learning objectives' or 'experimental goals' or 'requirements'), as the program objectives might not be the same as the end-user objectives. But, at first blush, it sounds unethical to encourage designers to disrespect the wishes of the audience they seek to serve!

Stephen Downes-Martin

I would add the other two witches of wargaming to your dictum about maintaining a healthy disrespect for them as well -- the sponsor and the game designer's boss (Downes-Martin 2014). Being disrespectful for the reasons you state is not unethical. It guards against the unethical influencing of the game design by the three witches (who do so for well-intentioned reasons).

Our interest is in deliberately exploring how we can use unethical practices to get the game sponsor to make a decision we prefer -- in order to identify such behavior, and thus defenses against them, that we might miss by just exploring what is ethical. Making the player's life easier might in fact be part of a useful unethical practice for making the game say something the designer (or the designer's boss) wants it to say. For example if the game was to compare two competing concepts, and the game design made it hard for the player to generate the advantages of one and easier to generate its disadvantages (and vice versa for the other concept), one could, by time-stressing the players, influence the conclusions of the game without the players or sponsor noticing. (I have not fully thought this through).

An actual example I saw was not unethical but provided exactly the method your note suggests. There were two Blue teams and one Red (played by the adjudicator cell, but that is not the issue I am discussing). It was an Operational level of war game. When the Blue cells sent requests for information (RFI) to the Control cell, we answered one cell (B1) with detailed tactical information (that Blue might reasonably have). When B2 asked for information, we

responded with vague promises that Intelligence would look into it, that it was unavailable, or we simply provided operational level information with little to no tactical detail. Initially we received the same volume of RFIs from each cell, but over the week the number of RFIs from B1 rose and the number from B2 dropped. We did not tell the players about this aspect of the game design. The game was an experiment about when accurate information requested by military commanders damages their decision making. B2 was significantly more successful at achieving their operational objectives in the game because they focused on the operational level. B1 became ever more mired in tactical detail, ran out of time each move and made hurried and uninformed decisions.

This result is backed up by experiments run by the CIA into the mismatch between when decision makers feel comfortable they have enough information to make a good decision and when they actually have enough information to make a good decision. Turns out that in general the latter occurs before the former (Heuer 1999). So, an unethical technique could be to "respect the players" by deliberately providing detailed information in response to a players' requests in order to influence the trajectory of the game by distracting the player.

Hiroyasu Akutsu

Has anyone discussed or worked on facilitators' unethical behavior? I am interested in "over-facilitation" or the kind of facilitation with some unethical intentions that goes beyond what the facilitators are expected to do in a team they are supposed to help. Facilitators can also help implement and support game designer's unethical plans for the game and player's decisions.

Stephen Downes-Martin

Facilitators are supposed to facilitate the players through the game's process as designed. It seems to me that the main mechanism for them to be unethical is to have their own agenda and to subvert the game process (whether or not that process was itself unethical) in opposition to the game designer. An opportunity for subversion by the facilitator occurs in the necessary ambiguity in the Facilitator's chapter of the gamebook that lays out the facilitation process. Like the law, not every circumstance can be reasonably covered in the gamebook and good facilitators must be allowed initiative, hence the "ambiguity". For example consider a seminar game design with group discussion inside each player cell. An expert unethical facilitator might suggest questions for the group (or collection of subgroups) to consider, worded in terms of gains or losses of alternatives -- but *not* both (Downes-Martin 2020). The human tendency to respond differently to gains than one does to losses will affect the group decision. Other pathologies generated by group discussion exist, each of which can be exploited by a smooth

facilitator. I have seen facilitators engage in some of these behaviors, whether from ineptitude or deep cunning (!) I don't know.

Stephen Downes-Martin

The Hidden Scenario approach is a useful wargaming method which may be open to interesting unethical manipulation of the player. The ethical intent of the approach is to benignly deceive the players as to the nature of the scenario in order to "hide the benefits of hindsight or the pitfalls of prejudice" gained from previous experience with real world or wargamed events (Curry 2012). Two examples of hidden scenario games along with a description of how players react to the scenario are "Home Front 86" (Price 1985) and the "Carter Racing Study" (Brittain & Sitkin 1989).

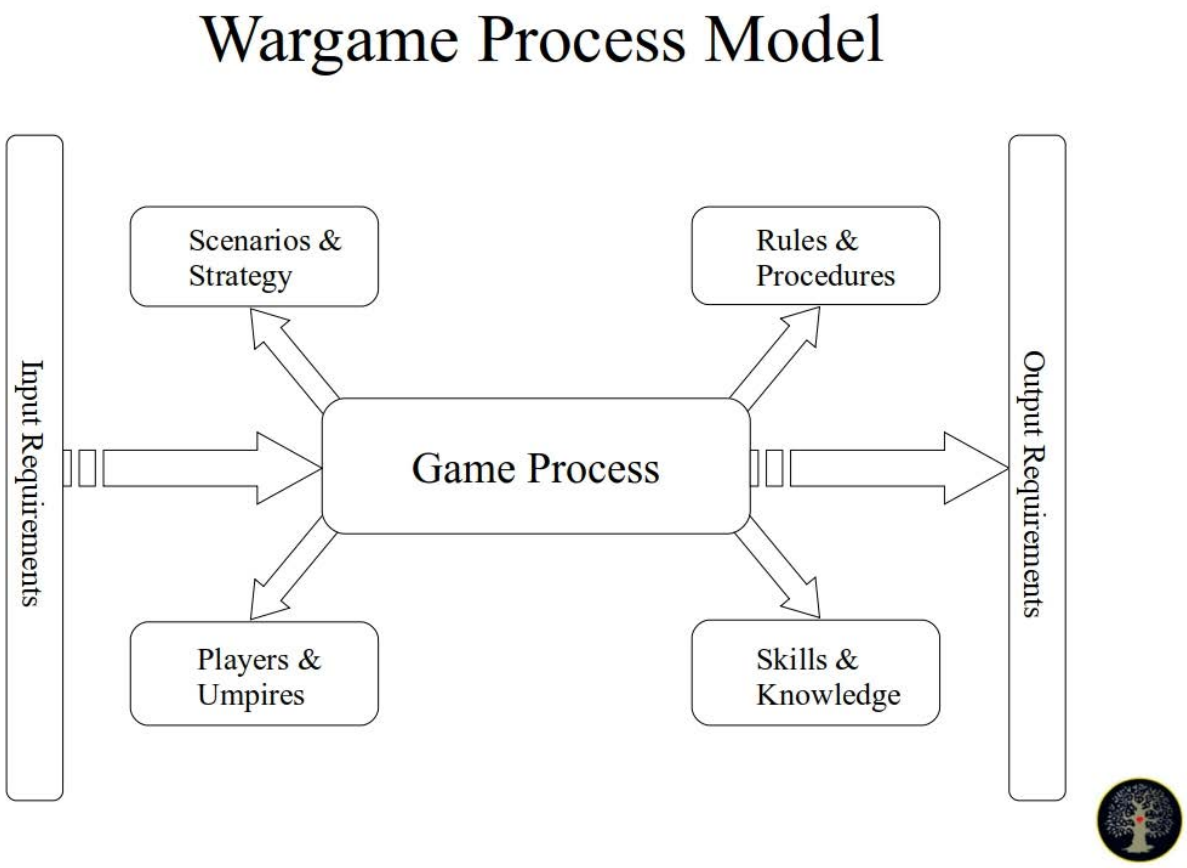
In a hidden scenario game it is easy for players to deviate from ethical norms or to put others' lives at risk without realizing it. In "Homefront 86" players end up running a death camp without realizing what they have done. In the "Carter Racing" exercise the players launch the Challenger Shuttle (disguised as a high performance racing car) despite having all the information necessary to work out the risk. In the latter it often happens that someone notices and states that "this is identical to the Challenger Shuttle" -- but is usually ignored.

The unethical game designer may be able to amplify the effects of the Risky and Dishonesty shifts using s hidden scenario approach, to lure the players into making large shifts before they realize it. A player may be able to imply to other players that a hidden scenario is in play when it is not in order to encourage decisions that are inappropriately risk averse.

A Wargames Process Model

Ian Robinson

I come from working 30+ years in the UK Chemicals Industry (ICI) and we are keen on using 'Quality Process Models' to describe workflow. Here's what a Wargame Process might look like that our business (Hearts of Oak) uses.



So we can begin to hang our hat on the process ...

Inputs ('interfacing' stakeholders. There will be more further back in a chain of command). Do they define what they want (tell the game exactly what they want) or describe broadly what they'd like (sell to the game their expectations, but don't care how it's delivered). Also is bias/unethical starting, and merely propagating through the process?

Game process

Scenarios and Strategy. Imposed or defined by the team? Are they realistic (Red vs Blue will be different), played to a script? An interaction here with Skills and Knowledge of players.

Players and Umpires. Obvious scope for bias/unethical influence here.

Rules and Procedures. Obvious scope for bias/unethical influence here. Has the game been systematically tested before this project started? Mark Flangan and I have an academic paper about to be published exploring a method for doing this in an unbiased way, and we found surprising nuggets.

Skills and Knowledge. Military people are trained in their own traditions - and these are the basis for Red/Blue differences, and 'shades of blue' opinions. Non-military people come with a different background. Again scope for bias/unethical influence here. Also there is a direct link to scenarios and strategy based on your experience.

Outputs (team to initial stakeholders). Are all game outcomes communicated in an unbiased way, including unusual or 'bad outcomes' vs expectations? These include low probability results. failure to do so would count as bias/unethical behaviour to me. Is the communication easy to understand, to both initial stakeholders and their chain of command? A failure to get this right might mean initial stakeholders garbling or changing the outputs when communicating onwards. Again scope for bias/unethical influence here.

On a continuous thought, the process model outputs become the inputs for the next step of the entire process and so on. Juran (quality guru) spoke of the virus of variability infecting industrial processes, from an earlier stage to the next one and so on. Bias/unethical behaviour is just the same here, infecting all subsequent stages.

Anne M. Johnson

I thought the Wall Street Journal article "What the Pandemic Has Taught Us About Science" might be relevant to this discussion ("The scientific method remains the best way to solve many problems, but bias, overconfidence and politics can sometimes lead scientists astray").[76] While not specifically about un/ethics in wargaming; believability and un/intended deception in scientific research and publishing has some similarities.

Ian Robinson

I agree with the article. As a professional scientist, I think we offer expert opinion in the legal sense, based on data and 'accepted' understanding. In truth no scientific theory can ever be proved true, but they can be proved false (Popper's falsifiability). The more difficult it is to prove a theory wrong, the more reliable the explanation. So do scientists act in unethical/biased ways? You bet, but they tend to get exposed quickly. Far more unlikely to be uncovered is a poorly designed or executed wargame, simply because it's unlikely to be repeated. Scientists do

[76] https://www.wsj.com/articles/what-the-pandemic-has-taught-us-about-science-11602255638

often repeat work just to check - they calibrate. Do we calibrate a wargame against our previous experience before we believe the results?

Stephen Downes-Martin

Ian's note about repeatability reminded me of David Goodstein's work dealing with scientific fraud. When David Goodstein was Provost at Caltech he was responsible for investigating all allegations of scientific misconduct. He identified three risk factors as present in nearly all cases of scientific fraud (Goodstein 2010):

1. The perpetrators "knew, or thought they knew, what the answer to the problem they were considering would turn out to be if they went to all the trouble of doing the work properly;
2. were under career pressure; and
3. were working in a field where individual experiments are not expected to be precisely reproducible."

The presence of these risk factors does *not* mean fraud is present. However when fraud has occurred these factors are nearly always present. They are therefore worth monitoring. I argue in "Your Boss, Players and Sponsor: the Three Witches of Wargaming" (Downes-Martin 2014) that these three factors are nearly always present in DoD wargaming:

1. Many senior officers are confident in their intuition or gut feel about matters in general and about topics being gamed.
2. For many important DoD wargames the sponsors are under career pressure to make the programs being gamed look good and be successful.
3. Repeatability in wargames is hard at best, and even if a wargame design is repeatable the DoD has little appetite for repeated gaming about a topic and all concerned know the likelihood of game repetition to be extremely low.

None of this means the stakeholders are automatically engaging in intellectual fraud, but it does mean the wargaming enterprise is vulnerable to intentional, albeit non-cynical, warping of intellectual honesty at best and at worst is vulnerable to deliberate intellectual fraud driven by careerism -- and in either case as Ian points out it is hard to spot.

Therefore we should be examining the three risk factors individually and in combination, ask ourselves where and how these might trigger intellectual fraud, and who would be vulnerable to temptation. Then do the thought experiment "how would I deliberately exploit the vulnerabilities in new and exciting ways?" Finally put in place active mitigation measures.

Ian Robinson

An unethical/biased game is a low quality outcome, as it doesn't meet specification (the LHS of the process model). It's a waste of time except for the biased person/group inside, and could lead to profound consequences if used for planning etc. When planes crash we have flight recorders to try to figure out why. A wargame with no equivalent cannot be investigated thereafter for signs of weakness. Game recording is certainly an active mitigation measure.

There is a culture clash between 'gut feelers' and 'measure and learn'. The first is quicker for sure but certainly leads to hubris and mistakes. The second (lessons learned) is slower but will lead to best use of resources available idc. It took the British Army 1916 and 1917 to figure out how to use the men and weapons to defeat the Germans in 1918.

Stephen Downes-Martin

I must challenge the idea that an unethical/biased game is a low quality outcome or that it is a waste of time. The idea that the ends justifies the means is both an ancient one and is alive today. Good people who believe they have the answer sometimes engage in unethical practice for the greater good and sometimes they turn out to be correct. Furthermore, good and ethical decisions sometimes turn out to be wrong. The quality of the outcome depends entirely on the outcome, not on whether the mechanism that produced that outcome satisfies some specification or the outcome was achieved by unethical means.

Unethical practices can be hard to detect, the outcomes from a process often cannot be checked for quality of "rightness" before implementing them, and the final outcome (this C2 structure, or that C2 structure, for example) often cannot be assessed for years -- until we go to war using the concept that a wargame influenced.

Ian Robinson

It's a low quality outcome inasmuch as it would fail to meet input specifications, one of which should be ethical play from the players? It is certainly a 'lost opportunity' if we think we won't repeat a game. Perhaps lie detector tests at the end of a game (did you play unethically) might be a simpler way to detect such behaviour.

Stephen Downes-Martin

To make sure I understand you, are you defining "low quality" to mean "does not meet input specifications, one of which should be ethical play"? If so, what happens if the unethical game designer designs a game that does permit unethical play? Is the outcome of such a game high quality?

Amna Greaves

Regarding lie detectors, there are two concerns:

1. The need for baseline data
2. The fact that many engineers/scientists suffer from "flatline" condition where we tend to be less emotional than more center-of-the-curve humans (meaning I don't know if you could accurately detect someone lying without taking a lot of data)

Design For Evil

Mark Simpkins

I am a visiting research fellow with Design Against Crime (DAC), a research centre based at Central Saint Martins, University of the Arts London. The main output of DAC has been in the area of product design and criminal misuse or abuse, to the end of studying and even talking to criminals about how they operate. This has also included, for example setting up observation points and watching how people use ATM points. Whilst this kind of ethnographic fieldwork is the norm the possible 'unethical' element is that it is being performed through the lens of possible criminal intent. In fact the one of the key parts of DACs work is about being as a designer more aware of these other 'uses', which recently has come to the fore in lots of discussions on the ethics of design and implementation of digital systems, AI/ML etc. (https://doteveryone.org.uk/project/consequence-scanning/). I was looking back at some earlier writing on design and sustainability and was reminded of Bruce Sterling's Viridian Design Project, where he outlined a set of design principles, Futurist, Politics, Research, Avant-technogarde, and moral, the last of which included:

> Design For Evil. There are a lot of dangerous people out there. Some have access to their own police state and army. A designer who does not plan for this is naive. *Compare to: Internet utopias, street finding its own use, dual-use tech*.

A lot of DAC's work has been around product design, chairs for bars and restaurants that have a cut out so you sit on your bag handle and can feel if someone is trying to take it, work around bike theft, better designs of bike stands and better communications on how to better secure your bike (two locks, off the ground etc), as well as work on Graffiti and how to work with communities and create spaces for expression. I was working on ideas around data and privacy, especially looking at the wearable tech space.

I developed a simple 'character' to discuss how one might exist in a highly connected digital space, a Smart City, and how one could maintain a semblance of control but still operate in such a surveilled environment. The Elegant Cyborg was a digital flaneur in this space, one who uses all the tech (so by definition is well to do) but plays the games with data capture companies, they mask and hide behind VPNs and other anonymising tech when they can, they use disposable email addresses (something which has been encoded into Apples approach), maybe they play the game when they 'swap' data patterns with other users, thinking back to when Tesco introduced the loyalty card in the UK, some people arranged swaps of the cards, so that the user profile was 'muddied', i would consider the Elegant Cyborg a cheat in how they operate, but outwardly. To me where this falls down is that this figure is a self centred operator, willing to cheat for their own use of the systems, but unwilling to share for a social benefit.

The other piece of work I briefly wanted to mention was work I did at Royal Holloway, with Prof. Lizzie Coles Kemp, she is pioneering methods of working with groups to discuss risk (especially cyber) in groups / organisations using creative tools. One we developed was using lego models and then I took it into creating Actor Network Maps as a model co-created by the group of an organisation. Here the benefit (as I saw it) was that the org model resembled the everyday experience of people doing the work, not the job descriptions and lines of managerial reporting. Whilst no direct cheating, you can see how organisations 'lie' to themselves in how they operate by hiding such experiences away from each other.

Some references:

Design Against Crime - http://www.designagainstcrime.com/

Viridian Design Archive - http://www.viridiandesign.org/

Viridian Design Note 3 - Viridian Design Principles - http://www.viridiandesign.org/notes/1-25/Note%2000003.txt

(Re)introducing the Viridian Design Movement -
(This is the article that reminded me of the Design for Evil principle)
https://medium.com/5-viridian-years/re-introducing-the-viridian-design-movement-aade6358664b

Elegant Cyborg Posts (from my blog) - http://www.geekyoto.com/?s=elegant+cyborg

There are some PDFs from the Holloway work, around Creative Security, which are also available here as print and make booklets
https://bookleteer.com/publication.html?id=3277,
https://bookleteer.com/publication.html?id=3268,
https://bookleteer.com/publication.html?id=3272

Rex Brynen

A few years back [a Western intelligence agency] proposed to [a commercial digital game company] if they could develop a game to counter recruitment and radicalization efforts by [a non-state actor]. I was informally asked by the company to comment on the initiative -- and I strongly advised against it, because I felt there was too high a risk that the target could repurpose the game, or respond to in ways that would undermine or even reverse the intended effect. In other words, the combination of innovative gamers and innovative bad guys (and the overlapping Venn diagram of innovative bad guy gamers) simply made it impossible to fully anticipate how it might be used and what effects it might have once released into the group's social/digital ecosystem. As far as I know the project never went ahead.

Mark Simpkins

Some of us looked at YouTube, and what you have there is basically the same thing, innovative bad guys and a platform that was just calling for creative content. Pat Kane, who wrote a book called "The Play Ethic", gave a presentation to some of the creative teams at the BBC a long time ago (2006), the slides are here
https://www.slideshare.net/theplayethic/pat-kanes-presentation-to-bbc-digital-futures-2006-on-the-ambiguity-of-play
and he touched on the use of hip-hop videos as a recruitment tool and distributed them on YouTube (he referenced a video 'Dirty Kuffar' by Sheikh Terra and the Soul Salah Crew, 2004). I'm afraid I can't remember a lot of the details of the talk and there are no other notes online but the gist was around the idea that everyone was getting media savvy, and the tools means anyone can do complex media but unlike good old TV, no regulators to keep you 'Socially Acceptable'. From that I started getting interested in the 'design' or non-design of military power point presentations (as it basically felt like the 'enemy' was far more media savvy).

Amna Greaves

Of interest to me is Grayzone Warfare aka Influence Operations. Rob Seater, Joel Kurucar, and I had spent a short while pondering how to simulate the mechanisms of Social Media to identify how bias is, or could be, manipulated. This led us down a dark path of "well if we can simulate bias, could we be harming our players". Ethics, of course, plays a key part but ... how to study it via the Systems Analysis methods that gaming presents without a) harming your players and b) creating more potent mechanisms for inducing bias should bring most people to a halt. At least long enough to decide if the ends justify the means.

If the Achilles heel of democratic nations in faithfully adhering to free speech and freedom of sharing information leaves the population open to manipulation, would games designed to expose subliminal intent, and elevate social media consumers into a liminal space where intent and bias are exposed, be helpful or only further weaponize online social structures? (Let's ignore for the moment the fact that this infrastructure is managed by corporations who make money from clicks, without regard for the vulnerability of their users or collective impact on our society.)

Very simply we have 3 objective in creating a game to simulate Influence Ops on Social Media:

1. Data collection - can we collect enough data on how humans react to misinformation to create computational models of behavior
2. Training - can we develop methods to trigger a thoughtful response, the same way a vaccine can trigger an immune response

3. Countering - can we learn from how humans try and deceive each other to create AI tools which can alert the end user when they're being targeted

Of course the unethical risks would be:

a. by exposing players to misinformation you might inadvertently lower their resilience
b. instead of utilizing this data to protect end users, malicious actors would use it to sway the undecided more effectively

I'm hoping there is a transition/liminal space where people become more inherently aware, have their acuity enhanced, to the actual messages within the content. I think training is the way to go, but the ethos in developing a training game in this space is ill-defined. At least for now.

Stephen Downes-Martin

We should expect any game "designed to expose subliminal intent, and elevate social media consumers into a liminal space where intent and bias are exposed" will be both helpful and further weaponize online social structures. Since online social structures are already weaponized and those weapons appear currently superior to any defenses we have, we should perhaps not be shy in designing such games with the purpose of developing defenses.

This will be a never ending battle in which the weapons and defenses co-evolve. An example is the co-evolution of anti-tank warheads and tank armor. Which dominates depends on the date, and whoever stops development first permanently loses. Another example is medical research which cures diseases and also is weaponized in the form of chem/bio warfare.

Perhaps exposing the bias-creating activities of an actor would trigger antipathy towards that actor as an unintended OR intended effect. If the latter does this constitute weaponizing by encouraging actions against that actor?

A User Guide for Unethical Wargaming

William "Wilf" Owen

1. It's all about reputation and nothing to do with expertise. There is no minimum standard in Wargaming. The community will never say out loud that you don't know what you are talking about and have no clue what you are doing because defending Wargaming is more important than doing good Wargaming. All gaming is good. Never admit otherwise!
2. Make sure you have all the "right people" in "the room" so as no one actually asks awkward questions about your competence or understanding because, for example, you cannot explain your own electronic warfare modelling.
3. Citing irrelevant credentials that have nothing to do with Wargaming is an excellent way to promote expertise in Wargaming.
4. Tell sponsors you can absolutely run a wargame that will give them actually evidence to support a decision. Make sure you understand what evidence they actually want so that your 'specially designed game' can provide the right answer.
5. Never allow your 'specially designed game' (SDG) to have its validity tested by using it to look at other questions. That undermines the premise of a "special design." Remember every new problem requires a completely new game system and comparing two different systems that in a way you don't control, will lead to some very uncomfortable questions.
6. The best way to develop an SDG is to copy a commercial game and then make just enough alterations so as you can claim it's a new game.
7. Remember to constantly over-state and over-complicate any form of combat resolution modelling and stay clear of all and any Practical Military Historians and/or Operational Analysts who will question it or even its relevance and value.
8. Give the sponsor exactly what they want in terms of an enjoyable, entertaining experience that can be done quickly and cheaply without causing him too many days away from his desk.
9. Avoid having participants who have a track record in real-world decision making. They always create problems by applying their real-world understanding.
10. Most sponsors do not realise that the more time you take to train the participants in the game system, the less useful it is, so make sure you disguise this fact somehow. It doesn't matter if reality is entirely different from the game system you are promoting.
11. The best way to promote enjoyment to a professional sponsor is to convince him to focus on abstract intangibles which tell him that Wargaming is really complicated and needs real expertise. Since you don't have any, it's all a bluff anyway.

12. Only by bluffing will you distract folks away from the fact that the game system promotes a decision-making process far removed from anything done in reality.
13. Write the game report in outline before the game is conducted and then prepare staff to look for events that would support the conclusion you want to sell to get more business. The more faux-complicated the game is, the better you can disguise events that never happened.
14. Above all else, make sure you promote all Wargaming and gaming to create a community unable and unwilling to critique itself so as it may collectively advance individuals' reputations.
15. So, when you see these points in print, make sure to attack the author's motivation and the validity of the observations. Remember if anyone breaks rank, it's over for Wargaming!!

Stephen Downes-Martin

Wilf, great list -- many thanks. I mostly agree with you. Here's some comments for each of your items based on my personal experience:

Items 1, 14, 15:

Professions have three characteristics. A documented knowledge base about the topic, a documented set of ethical and performance standards, and a licensing body that not only determines the content of the theory base but also has the authority to grant and remove membership in the profession. For example the "profession of arms". It has a documented base of knowledge about warfare, it has a documented set of standards, and a system for promotion, demotion and eviction. The legal and medical professions have similar characteristics.

Wargaming only has the first of those items (lots of literature) so we are NOT a profession according to accepted usage of the word. This state of affairs does not bother me (nor I suspect anyone else) but it does mean we do not have an agreed on set of minimum standards for a wargame that can be enforced. Therefore criticism if it occurs has to be based on some high reputation wargamer's opinion about wargaming. So it all comes back to your point about personal reputation of the wargame designer being the measure of effectiveness of a wargame, and any criticism is brushed aside by pointing at the wargame designer's reputation and blame for any lack of effectiveness at dealing with the sponsor's problem is thus directed away from wargaming!

NOT being a profession does NOT mean we do NOT behave professionally. Most of the wargamers I communicate and work with are knowledgeable and experienced, and are skilled and ethical. They choose to behave professionally. But there is no agreed set of standards, and thus it is all too easy for the charlatans and inexperienced to proliferate and thereby give

wargaming a bad name. Unfortunately many sponsors turn a blind eye to poor wargaming to avoid being embarrassed or to justify a preferred decision.

A way to exploit this is to make sure you include in the game someone from the sponsor's staff whom the sponsor trusts and find out as much as possible from your contact about what the sponsor really wants. Allow that person to think they are collaborating with you on game design and give that person center stage in explaining the implications of the wargame to the sponsor after the game. The sponsor will be motivated to accept the game, and any perceived problems will be ignored if the game outcome supports the sponsor's ideas.

Items 2, 9:

Using classification issues is a great way to do this. I've seen games derailed when a player/observer in the room claims that information at a higher level of classification (than the game is being played at) completely debunks the adjudication or conflicts with critical game assumptions. I have also seen challenges to a game's assumptions, adjudication or scenario successfully fought off by appealing to "higher classification information which I cannot provide, just trust me, the game is properly designed". We are back to reputation. A way to exploit this is to state up front that the game has been designed to take into account, while simultaneously not revealing it, classified information not included in the game. This cannot be challenged, and since it is untrue your game is not at risk of revealing information at a higher level of classification than the game permits. BUT make sure you exclude from the room people who can challenge you based on their own knowledge and reputation.

Most senior people much prefer to talk to their peers and prefer to talk than "play a game". Exploit this. If participants who have a track record in real-world decision making are imposed on you, design a special senior VIP cell of officers and civilians who will discuss and debate various products from the game "at a higher level" than the game is being played. So, for example, for an operational game have a senior Strategic cell. You control what information from the game is passed to them, and because they are discussing at a higher level, with a longer time horizon than the game, you also choose what to send back to the game (if anything) from their deliberations. You also control how to write up the report of their deliberations.

Item 3:

Wargaming is not a profession, and I argue also it is not a discipline. To be good at it requires knowledge about a wide range of disciplines rarely discussed by wargamers, for example psychology, group dynamics, decision analysis, leadership styles, history, weapon systems design, physics, the list goes on. Most wargamers have a top credential in one or more of these, but then that is also true of many non-wargamers working in national security. At least having an advanced research degree in one of the many fields that wargaming uses demonstrates an

ability to think at a high level and deep knowledge in a field that might be relevant to the current game.

Items 4, 13:

I believe the best way to do this is to use inductive adjudication, where the adjudicators decide on the outcome of an interaction based on "the sponsor's objectives". They have to make sure the outcome is possible (i.e. not a zero probability in the deductive adjudication combat results table) and provide a story of how it might have come about. The "That's very unlikely!" challenge is countered with "Agreed, but not impossible, and it happened. So how are you going to deal with it, Sir?"

Items 5, 6, 10, 11:

"Our boutique wargaming facility tailors every game (expensively, of course) to your incredibly important and unique problem using our proven MSU-2000 proprietary wargaming system!" Or words to that effect. New game every time, no one can look under the hood to challenge it. I was amazed decades ago when I first discovered how willing senior officers were to accept "black box" designs. They still are.

The boutique proprietary system means the sponsor does not need to "train to play"! The players just describe their decisions and the wargamers do the playing behind the scenes (call it "adjudication"). Remember, most officers don't like playing games anyway, they are good at "doing their jobs" and "talking". Give them a game in which that is what they do. The intangible complicated stuff is dealt with for them by the MSU-2000 system provided by the high reputatio contractor-consulting practice.

Item 7:

I saw a new adjudication method at a DoD wargame about an important subject actually called "The Gonculator". It used an enormous spreadsheet to do bogus arithmetic on ordinal numbers to a precision that would have been the envy of a nuclear physics lab. What was extraordinary was the players and sponsor were not warned by the name of the adjudication device!

Item 8:

Fast, Good, Cheap. Pick two. So provide fast and cheap in the game, and invite the sponsor and staff at their location to participate in a debrief session that is actually a discussion of how the wargame outcomes can be argued to support the sponsor's "right answer" -- see item 4.

Item 12:

Bluffing ... cuts across all the techniques. Bluffing has high payoff in most arenas. An obvious example is Poker.

Items 14, 15:

When does the house of cards collapse? When do sponsors realize en masse that they are not getting enough from their wargaming time and dollars?

Ed McGrady

I have to quibble (being a wargamer) about Wilf's focus on self criticism. I am not sure that there is even a structure that would allow for criticism of games, or game designers, much less anyone who would pay attention to it. The impression I get is that you feel the community is self satisfied and unwilling to critique itself -- I have not seen that, rather I've seen a willingness to criticize virtually anything that moves, but very few vehicles that allow this criticism to have an effect. I do agree that there is a reluctance to criticize other designers, but I'm pretty sure that comes from an assessment that a) they wouldn't listen and b) calling people out in public is probably not going to make you a lot of friends. So cost/benefit trade offs are made, as they are in any area of endeavor. (And I am speaking as someone who has had evaluation/hiring authority and experienced not listening first hand)

The point about having real operators in your game is spot on. If your game cannot survive contact with people who have been forward, then it's not a very good game.

Stephen Downes-Martin

Ed claims that wargamers are willing to "criticize virtually anything that moves". I agree that we criticize what Peter Perla calls "the instrumentality" of the wargame design, but I agree with Wilf that we rarely criticize the efficacy of the game at addressing the sponsor's objectives for the game. We spend far too much time congratulating ourselves on clever game mechanics, and not enough solving sponsors' problems.

Graham Longley Brown

The conversations to date seem to concentrate on the ***demand*** from sponsors for wargames and wargamers. My point relates to the ***supply*** of wargames and wargamers -- it is people being offered by 'prime' firms as experienced wargame practitioners when these folk lack the necessary skills and experience. The motivation to supply these 'pretender wargamers' is profit and/or reputational gain ("Everyone's wargaming, so our firm needs a wargaming capability"). It is quite easy to 'talk a good wargame', especially to inexperienced customers, and there is a lack of good wargamers so wargames are too-often run by unethical and/or inexperienced folk. I have too often witnessed such people fail to deliver the key 'pillars' of successful wargames:

- design to a purpose (sometimes not knowing what a wargame is and is not, let alone how to design one);
- robust analysis (lacking even a basic knowledge of that crucial and all-pervading function);
- and facilitation (for example simply reading out the headings of a too-rigid process).

This might play into our wider conversations because an experienced and ethical wargame practitioner will:

1. Emphasise to a sponsor what a wargame can't do or when it might not be appropriate (other techniques might be better, or used in conjunction with wargaming).
2. Directly resist a sponsor who tries to skew wargame design or execution to deliver pre-ordained outcomes - even walking away if necessary.
3. Shape the wargame design and delivery to mitigate any sponsor (and all other) biases, covertly if necessary.

'Pretender wargamers' won't do 1 or 2, and can't do 3. The damage done to individual wargames can be considerable - but so too is the reputational harm to wargaming generally: anyone witnessing a poor wargame is likely to henceforth discount the technique.

Stephen Downes-Martin

Graham is describing that part of the "outer game" where the sponsor and game provider negotiate over what skills are required to design and implement the game. The question generated by Graham is "*how* would one successfully persuade a sponsor to accept unskilled and inexperienced wargame designers and providers for a wargame project" (preferably without the sponsor realising this was happening). Note that simply lying about the wargamers' skills is an unskilled level of unethical behavior -- too easily found out with consequential pain for the 'prime' firm. No, one must persuade the sponsor that the skills provided by the 'firm' are indeed necessary and sufficient even though they are not. Graham's laydown of the problem provides insights into how to do this (and therefore how to spot when it is being done to one).

First, put ourselves in the role of the unethical marketing or sales manager responsible for winning wargame projects despite a lack of skill and experience in wargame design and execution, and perhaps despite the sponsor's problem, question or objective not being best served by a wargame. For the purpose of the thought experiment at the heart of this working group "You" are now the unethical sales or marketing manager. Initially "you" are interested in winning the contract, and do not care what the wargame reveals or implies. If you also want to use the wargame to influence the sponsor to support a specific decision post-game, you will use additional techniques to those described below. First, however, you must win the contract.

You might consider two types of intelligent and honest sponsor; those familiar vs unfamiliar with Wargaming. For dishonest sponsors the problem has to be solved by the sponsor's chain of

command or larger organization. Note that honest people are still subject to wishful thinking, psychological and cognitive biases, logical fallacies and character flaws, all of which can be exploited.

First, for either type of honest sponsor, lay the foundation:

- Do your IPME (Intelligence Preparation of the Marketing Environment -- that's only partly a joke!) on the sponsor(s). Identify overworked individuals with high stress jobs who are short of time, they will be grateful for assistance.
- Develop relationships with independent wargamers by discussing with them the overlaps and complementary nature of their skill sets with your organization's. Make them sign non-disclosure agreements.
- Develop a relationship with the sponsor based on what you are good at and veer the discussions into wargaming areas you want to expand into without proposing solutions (yet! You don't want to give away how shallow your bench is). So far so good, it's what any honest organization would do anyway.

Second, get sneaky:

A. You are faced with a sponsor who is familiar with wargaming:

- Run a workshop/conference about the sponsor's problem (sanitized to remove the sponsor's identity or sensitive/classified information) inviting your independent wargamer contacts to contribute.
- Build a credible wargame response that addresses the sponsor's objectives based on results from the workshop/conference -- your wargame participants write the proposal. Include the names of key wargame contacts in the proposal.
- If you win, during execution use your (small number of) key wargame consultants to keep the game on track.
- Have your wargames consultant(s) write the game report.
- Having developed a wargame reputation with the sponsor, propose the next contract without your wargame consultants to increase your profits. Your first sponsor has paid for your organization to enter the market.

Much of this looks like what an honest organization with good wargaming skills would do. What are the indications and warnings that bamboozlery is going on? I suggest possible danger signals include the following (these are not proof of malfeasance, they are indicators of vulnerability):

- The proposal does not include named project talent, and if it does the organization the talent works for is not explicitly identified -- a highly visible indicator.

- The proposal comes from an organization that has not had staff publish wargame papers or give wargame presentations at wargame conferences -- requires the sponsor to do work. Most honest people are loath to assume dishonesty on the part of charismatic colleagues, and if they are very busty will avoid deep investigation. (If you don't think this will work, ask yourself, how many managers when hiring do more than read the letters of recommendations and resume the candidate has provided and then interview the candidate? I.e. accept everything credible looking and charismatic candidates say about themselves). Remember the candidate has had more time to research you than you have had to research the candidate. The same is true of you (the "candidate") and the sponsor (the "hirer").
- No institutionalized process for checking credentials of contractors.
- The sponsor is spending a disproportionate amount of personal time with a single wargaming organization compared to other wargaming organizations.
- Several low cost sole source wargaming contracts are with a single company instead of a single larger valued competitively bid contract. (US Government may let sole source contracts if the value of the contract is less than a certain amount).

B. You are faced with an honest sponsor who is unfamiliar with wargaming

Some senior people suffer from the Dunning-Kruger effect. Also research shows a tendency among senior people to overestimate their knowledge and skills at topics with which they are in fact unfamiliar. Their past successes that resulted in their promotion makes them over-confident about the current situation. Others are self-aware about their own shortcomings. Your IPME will give you clues about which type you are facing. In either case start with the techniques described in A above.

If the sponsor suffers from the Dunning-Kruger effect you are in a position to manipulate the game outcome to support what the sponsor would really like the game to say. This places you in an excellent position to win the next wargame contract from that sponsor. Danger signals (again, these are indications and warning of a vulnerability) -- which will probably be ignored by the overconfident sponsor anyway, include:

- Same as A above
- The wargame design is such that "Blue" (the sponsor's pet rock) can't lose? Blue's victory is baked into the design.
- The wargame design allows the sponsor to play as part of the Blue cell or to be part of the adjudication
- The sponsor is allowed to do the post-game analysis

The sponsor's chain of command has to institutionalize procedures for looking for and dealing with these danger signals.

Use the Denialist Playbook

Stephen Downes-Martin

An interesting article in Scientific American (Nov 8, 2020) describes the denialist playbook, an approach that is similar between different groups denying different things (in the article this includes anti-vaxers and creationists). (https://www.scientificamerican.com/article/the-denialist-playbook/)

In short the playbook recommends:

1. Doubt the Science
2. Question Scientists' Motives and Integrity
3. Magnify Disagreements among Scientists and Cite Gadflies as Authorities
4. Exaggerate Potential Harm
5. Appeal to Personal Freedom
6. Reject Whatever Would Repudiate A Key Philosophy

(In some ways these overlap the "Tyrant's playbook" in "How to Become a Tyrant" (Netfflix, https://www.netflix.com/title/80989772), specifically session 4 "Control the Truth"),

I suspect this can be adapted to the discussions that occur during the outer-game as well as inner-game design. For example, if a sponsor objective is to explore a novel concept (a new approach to C2 between joint forces for example, and to refer back to a real-world example) then the unethical stakeholder (designer, player, sponsor, whoever) could use this playbook:

1. Doubt the doctrinal background of the objective
2. Question the motives and integrity of the service, community, contractor, whoever ... (supports the concept)
3. Magnify disagreements among supporters and cite gadflies as authorities.
4. Exaggerate the risks of the new versus the known and familiar
5. Point out how this might interfere with unity of command and commander's freedom of maneuver.

We know the denialists playbook works within the general population. Can it work when used within a wargame? I suspect yes -- having seen parts of it used successfully, but how about as an integrated whole? If it can work, what are the mitigations?

We have from research that the only known effective method of dealing with the "dishonesty shift" that occurs during small group discussions is to remind participants of the ethical norms of the group at the beginning of every discussion. This implies a possible mitigation approach, being to follow the initial discussion with the sponsor with an analysis of how the playbook could be applied, and then include in the follow on discussions this analysis, what to look for in case anyone is using it, and to call it out when seen. Basically "challenge the

challenges", make challengers explain their challenge every step of the way, including cunningly implied challenges.

Some References:

Mazar, N., Amir, O. & Ariely, D. (2008). "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance", Journal of Marketing Research: December 2008, Vol. 45, No. 6, pp. 633-644, https://journals.sagepub.com/doi/pdf/10.1509/jmkr.45.6.633

Kocher, M., Schudy, S. & Spantig, L. (2016). "I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups", University of Munich Discussion Paper 2016-8, https://doi.org/10.5282/ubm/epub.28966

Carol, S., (2020), "The Denialist Playbook: On vaccines, evolution, and more, rejection of science followed a similar pattern", Scientific American November 8, https://www.scientificamerican.com/article/the-denialist-playbook/

Daniel Tyler Brooks

I counter with Thomas Kuhn and Paul Feyerabend. Science is the consensus theory of justification within a community of experts using inductive methods of falsification. And because Science requires a dialectic of ideas within a social group of elites, any outside critique can be countered with an appeal to authority fallacy. This is important because a social group of elites is just as susceptible to groupthink and bias as a group of laypersons. This is particularly true the farther away you get from "hard" physical sciences. Paradigm shifts occur when rebels within a discipline, eclectic polymaths from outside the discipline, or reality smashing into bad theory forces an evolution or revolution in theory. So, scientists might be less wrong than the general joe, or the guy "using common sense," but the critique you critique may actually be the only reason we ever make any real scientific progress. If you need an example, look at the scientific consensus of COVID-19 response and how scientists advising public policy have perverse incentives to engage in face-saving behavior at the expense of public health. If there is any doubt in this, ask this question: Were masks ineffective at protecting the general population in the early stages of outbreak, or were public health officials covering for the fact that the government failed to replenish the strategic national stockpile and outsourced medical equipment production to China (who then used all of it)? "Follow the science" is meaningless. We follow the scientists, and the scientists are people with agendas and incentives.

So, more to the question, how does consensus theory of justification open the door to the unethical scientist designing a wargame for research purposes? When I've looked at this question before, the answer brought me back to the reddit thread "Explaining D&D with Tomatoes". I opened this paper talking about scientific pragmatism discussing the analogy, but the point is that consensus can be gained through charisma as much as wisdom or intelligence.

We like to think that science is special because of experimentation, but some ideas are so attractive (read: elegant) that they can waste the time of the world's greatest minds for decades without producing any empirical prediction.
https://drive.google.com/file/d/0BwELB9c3gW5BdV85MElKYmd3NzQ/view?usp=sharing

Stephen Downes-Martin

"Being unethical" is not the same as "being wrong". The scientists opposing external critiques of their beliefs can do so for good or for unethical reasons independent of whether their beliefs are justified or not. Nevertheless, the anti-vaxxers are wrong as are the creationists, and yet their playbook works with large numbers of intelligent and honest people.

Daniel Tyler Brooks

The difference between "being wrong" and "being unethical" is a question of intent, of course, acting in "good" or "bad" faith. But, you could also argue that being wrong is unethical from a consequentialist perspective. For example, if you predict a public health strategy will be what's best, and then it kills millions of people. You were wrong... were you unethical? Depends on how many people you silenced who disagreed with you, and were more right.

So, how is this useful? If I was designing an unethical wargame, I'd attempt to maximize groupthink by putting all adjudication decisions to popular vote, and refuse to record parking lot discussions, and have no "Minority Report" in the wargame report. I also wouldn't invite people who would be contentious.

Ed McGrady

Daniel brings up a good point. Why not, if I'm unethical, simply create a total fiction saying what I want it to say? If I were doing an unethical game I would simply write the game report, not identify the players, and not hold the game in the first place, or hold an event that everyone would declaim as the game, but was simply window dressing. How many people actually check? In fact, now that I'm going down this path, I'd generate a ton of fictitious data, and then do a lot of high-powered analysis on the data. In fact you could push even further and create a set of statistically valid game results derived from agent-based players conducted over a series of years in different scenarios. The agent-based player software is proprietary. You can't see it. If necessary could show people agent-based algorithms I threw together in Python, or do what I said above and hold a false game that had nothing to do with the actual game report. I'd have to modulate my overreach so as not to generate enough of a signature to create skepticism. But that is manageable. Again, how many people check?

Daniel Tyler Brooks

To McGradye's point, because a wargame is a model... presumably of reality, start with the conclusion and then reverse engineer a game that carries the players to that conclusion. Then show off your data as evidence of the conclusion. As long as the audience of the report is not academic or versed in the method, they'll never know.

Stephen Downes-Martin

I add a requirement to Ed's important one about avoiding overreach -- Try not to have your own fingerprints all over the unethical practice in case it is detected. One way to do that is to have a credible scapegoat, preferably one who has moved on and is no longer subject to investigation. So work behind the scenes as the assistant to a senior manager, preferably a military officer since they rotate every couple of years, who is unskilled at wargaming and analysis and who will lean heavily on your advice.

Ed's scenario involving a multi-pronged multi-event and multi-method campaign of unethical behavior might, to the uninitiated, appear far-fetched. Unfortunately (or fortunately for our working group!) it is anything but. What makes his approach effective (an implicit requirement for our working group) is that military sponsors rotate frequently and they want fast results from each event (the hotwash) to support their programs and do not go back and read the final analytic report. None of them are going to "check the paperwork and claims" if they can't be bothered to read reports about their own games. So a technique to support Ed's scenario is to ensure the last day of the game has a "hotwash" and "outbrief" to the sponsor to convince the senior stakeholders they have the game results and do not need to read any sort of "analysis" that might follow.

Ed McGrady

I am now encouraged to continue down this dark path. I believe there is an implicit assumption in several of the comments that there has to be some sort of thread between what happens in the "game" and the results presented to the world. In the world I inhabit that absolutely does not have to be the case. I do a "game" on the next generation whatsits that consists of a bunch of experts sitting around for a dulsatory 4 hours talking about scenarios and whatsits and running around "playing" a game. I thank them very sincerely for their contribution to national security. I then go off and write a game report, that will be widely circulated, that says whatever I want it to say, using whatever comes from the game, or not, that I can throw in for a veneer of a fig leaf in case I am caught by a shrewd fellow. As long as the game was not on predicting the latest Pantone color set and the subsequent report only talked about intermediate range ballistic missiles I'm probably good. The players will assume the insights came from some other cell, if they bother with it in the first place. I certainly won't identify the

players in the game report, I'll just assert that a game was played, this happened, and here's what it means. The only thread linking the game to the report is any threads of modesty I need to deflect the curious or skeptical. I contend that since most leaders don't get that far, I don't need those threads. But if you are less bold then the threads are there. But they don't at all ever have to be some clever ruse to get the players to do what you want them to do, or a well connived scenario that biases the game. Just say what you want, and use the game for cover. Done. (Realistically this is what happens to a lot of games).

Unethical Wargaming Bibliography

This section contains all bibliographic references contained in the Working Papers and the Discussion. It also contains additional references to deception, fraud, propaganda and psychology relevant to perpetrating successful unethical behaviors.

Activision. (2021). Cheating Cheaters: Malware Delivered as Call of Duty Cheats. https://www.activision.com/cdn/research/cheating_cheaters_final.pdf

Axelrod, R. (1984). The Evolution of Cooperation. *Basic Books Reprint 2009*

Axelrod R. (1997). The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. *Princeton University Press*

Bailey, R. (2021). How Much Scientific Research Is Actually Fraudulent? It may be more than you think. Reason, 7.9.2021 https://reason.com/2021/07/09/how-much-scientific-research-is-actually-fraudulent/

Batteux, E., Ferguson, E. & Tunney, R. (2017). Risk Preferences in Surrogate Decision Making. *Experimental Psychology*, 64(4), 290-297 https://doi.org/10.1027/1618-3169/a000371

Beach, E. (2016). A Mighty Fortress Is Our God: When Military Action Meets Religious Strife. In Zones of Control: Perspectives on Wargaming, Harrigan, P. (ed.), Matthew G. Kirschenbaum, M. (ed.), chapter 45, pages 539–544. *MIT Press*

Bekoff, M., Pierce, J. (2009). Wild Justice: Honor and Fairness among Beasts at Play. *American Journal of Play*. 1(4), 451-475, Spring https://www.wellbeingintlstudiesrepository.org/soccog/4/

Bell, B., Whaley, B. (1991). Cheating and Deception. *Transaction Publishers*

Berkowitz, R. (2020). A Game Designer's Analysis of QAnon. *curiouserinstitute*, September https://medium.com/curiouserinstitute/a-game-designers-analysis-of-qanon-580972548be5

Bernays, E. (1928). Propaganda. *Ig Publishing*

Betteridge, I. (2009). TechCrunch: Irresponsible journalism. *Technovia* February 23 https://web.archive.org/web/20090226202006/http://www.technovia.co.uk/2009/02/techcrunch-irresponsible-journalism.html

Boda, Mihály. (2020). The Hungarian Theory of Just War Based on the Idea of the Holy Crown. *Journal of Military Ethics*, 2020/4 http://real.mtak.hu/112918/

Bond, S., (2021). Wargaming at the Academy: Why Invest in NHWL and Building a Culture of Wargaming. *National Security Innovation Network: Defense Innovation Network*, January https://cimsec.org/wargaming-at-the-academy-why-invest-in-nhwl-and-building-a-culture-of-wargaming/

Born, J. Cheating At Texas Holdem. *Born Innovations* https://www.vanishingincmagic.com/magic-books/cheating-at-texas-holdem

Brandes, D. (2020). Docu-Fictions of War -- Review. *RealClear Defense*, November 11 https://www.realcleardefense.com/articles/2020/11/11/docu-fictions_of_war_583691.html

Brightman, H. (2007). Nash in Najaf: Game Theory and Its Applicability to the Iraqi Conflict. *Air & Space Power Journal*, Fall 2007, Vol. XX1, No. 3, Pp 35-41 https://www.files.ethz.ch/isn/120990/fal07.pdf

Britain, J., Sitkin, S. (1989). Facts, Figures, and Organizational Decisions: Carter Racing and Quantitative Analysis in the Organizational Behavior Classroom. *Journal of Management Education*, Vol 14, #1

Brooks, D. (2020). Propaganda Games. *YouTube*, October https://www.youtube.com/watch?v=UP4_bMhZ4gA

Brown, Anthony Cave (1975). Bodyguard of Lies. (Two volumes), *Harper & Row*

Card, O. (1991). Ender's Game. *Tom Doherty Associates*

Carol, S. (2020). The Denialist Playbook: On vaccines, evolution, and more, rejection of science followed a similar pattern. *Scientific American* November 8 https://www.scientificamerican.com/article/the-denialist-playbook/

Caroland, J., Conti, G. (2012). Lessons of the Kobayashi Maru: Cheating is Fundamental. *ShmooCon* https://www.youtube.com/watch?v=7902m6MTu0E

Case, N. (2018). How to Become a Centaur. *Journal of Design and Science* https://jods.mitpress.mit.edu/pub/issue3-case/release/6

Chambers, C. (2017). The Seven Deadly Sins of Psychology: A Manifesto for Reforming the Culture of Scientific Practice. *Princeton University Press*, April

Cohen, E., Gooch, J. (1990). Military Misfortunes: The Anatomy of Failure in War. *Free Press*

Cooney, G., Mastroianni, A., Abi-Esber, N., & Brooks, A. (2020). The many minds problem: disclosure in dyadic versus group conversation. *Current Opinion in Psychology*, Volume 31, February, Pages 22-27 https://doi.org/10.1016/j.copsyc.2019.06.032

Curry, J., Price, T. (2014). Matrix Games for Modern Wargaming. *Developments in Professional and Educational Wargames, vol. 2. History of Wargaming Project*. https://lulu.com/shop/product-21760442.html

Czeglédi, Mihály. (2021). A küldetésorientált vezetési szemlélet kialakulását befolyásoló tényezők. (Factors with Influence on the Shaping of Mission-Oriented Leadership Approach). *Honvédségi Szemle* January, pp 75-88 https://doi.org/10.35926/HSZ.2021.1.6

Czibor, Andrea. (2014). Döntések társas dilemma helyzetekben: személyiségjellemzők és szituációs faktorok hatása (Decisions in Social Dilemmas: The Impact of Personal Traits and Situational Factors). *PhD Thesis, University of Pécs* https://pea.lib.pte.hu/bitstream/handle/pea/15180/czibor-andrea-phd-2014.pdf?sequence=4&isAllowed=y

Danczer, Alfons (ed.) (1889). A mi hadseregünk. Az Osztrák-Magyar Monarchia népei fegyverben és a zászlók alatt (Our Army. The Peoples of the Austrian-Hungarian Monarchy in Arms and under the Flag), *Franklin Társulat*, pp 76-85, Budapest

Daniel, D., Herbig, K. (Eds). (1981), Strategic Military Deception. *Pergamon Press*

Davies, M. F. (1997). Belief Persistence after Evidential Discrediting: The Impact of Generated versus Provided Explanations on the Likelihood of Discredited Outcomes. *Journal of Experimental Social Psychology* 33, no. 6, November, pp. 561–78

Davis, P., Bigelow, J. (1998). Experiments in Multiresolution Modeling (MRM). *RAND National Defense Research Institute*, MR-1004-DARPA https://www.rand.org/pubs/monograph_reports/MR1004.html

Design Against Crime Research Centre http://www.designagainstcrime.com/

DIB. (2019a). Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense. Defense Innovation Board, October 31 https://media.defense.gov/2019/Oct/31/2002204458/-1/-1/0/DIB_AI_PRINCIPLES_PRIMARY_DOCUMENT.PDF

DIB. (2019b). Recommendations on the Ethical Use of Artificial Intelligence -- Supporting Document. Defense Innovation Board, October 31 https://media.defense.gov/2019/Oct/31/2002204459/-1/-1/0/DIB_AI_PRINCIPLES_SUPPORTING_DOCUMENT.PDF

Dinklage, P. (2021). How to Become a Tyrant. *Netflix* https://www.netflix.com/title/80989772)

Dodoiu, G., Leenders, R. & Dijk, H. (2017). A meta-analysis of whether groups make more risky or more cautious decisions than individuals. *Academy of Management Proceedings*, Vol 2016, No 1, 30 November https://doi.org/10.5465/ambpp.2016.194

Dorner, D. (1996). The Logic of Failure: Recognizing and Avoiding Error in Complex Environments. *Addison-Wesley*

Downes-Martin, S. (2004). Pathologies of War Gaming and What to Do About It. *Naval War College Briefing* February

Downes-Martin, S. (2013). Adjudication: The Diabolus in Machina of War Gaming. *Naval War College Review*, Vol 66, No. 3, pp 67 – 80
https://digital-commons.usnwc.edu/nwc-review/vol66/iss3/6/

Downes-Martin, S. (2014). Your Boss, Players, and Sponsor: The Three Witches of War Gaming. *Naval War College Review*, Vol. 67 (No. 1, Article 5), 2014
Your Boss, Players, and Sponsor: The Three Witches of War Gaming

Downes-Martin, S. (2015). Stress, Paranioia and Cheating: The Three Furies of Innovative Wargaming. *Connections US 2015 Wargaming Conference*, National Defense University Washington DC, July
https://paxsims.files.wordpress.com/2015/07/wargaming-as-a-catalyst-for-innovation-20150828.pdf

Downes-Martin, S. (2016). Wargaming to Deceive the Sponsor: Why and How? *Connections UK 2016 Wargaming Conference*, King's College, London, September
https://www.professionalwargaming.co.uk/WargamingToDeceivePaper.pdf

Downes-Martin, S. (2019a). Malign Wargaming I: Group Dynamics in Wargames and How to Exploit Them. *Connections North Wargaming Conference*, McGill University, Canada, February.
https://paxsims.files.wordpress.com/2019/02/group-dynamics-in-wargames-sdm-20190216.pdf

Downes-Martin, S. (2019b). How an opponent wargames is an intelligence collection requirement. *Intelligence Community Forum Mercyhurst University*, June
https://www.mercyhurst.edu/sites/default/files/icf_2019_downes-martin.methods_for_decision-makers.pdf

Downes-Martin, S. (2020). Preference Reversal Effects and Wargaming. *Connections North 2020 Wargaming Conference*, McGill University, Canada, February
https://paxsims.files.wordpress.com/2020/02/2020-preference-reversal-effects-and-wargaming.pdf

Downes-Martin, S. et al. (2017). Validity and Utility of Wargaming. Working Group 2 Report *MORS Wargaming Special Meeting*, Washington DC, October
https://paxsims.wordpress.com/2017/12/09/mors-validity-and-utility-of-wargaming/

Downes-Martin, S., Phillips, K., Shogren, K., Stetson, S., Tropeano, R. (2018). How Can We Avoid Risky and Dishonesty Shifts in Seminar Wargames? In *Connections: US Wargaming Conference: Game Lab*, National Defense University, Washington DC, July
https://paxsims.wordpress.com/2018/08/13/how-can-we-avoid-risky-and-dishonesty-shifts-in-seminar-wargames/

Ebner, J. (2021). The gamification of terrorism. *YouTube*, March
https://www.youtube.com/watch?v=rPC_Zf16Nts

Edelson, M., Polania, R., Ruff, C., Fehr, E. & Hare, T. (2018). Computational and neurobiological foundations of leadership decisions. *Science*, 03 Aug, Vol. 361, Issue 6401 https://www.zora.uzh.ch/id/eprint/153022/19/Computional_Edelson_akz.pdf

Eder, M. (2020). The Information Apocalypse, Part VII: Competence and Ethics. US Army War College War Room, September 10 https://warroom.armywarcollege.edu/articles/the-information-apocalypse-pt-vii/

Ellenberg, J. (2017). How Computers Turned Gerrymandering Into a Science. *The New York Times*, Oct 6 https://www.nytimes.com/2017/10/06/opinion/sunday/computers-gerrymandering-wisconsin.html

Eysenbach, B., Gupta, A., Ibarz, J., Levine, S. (2018). Diversity is all you need: Learning skills without a reward function. *Computing Research Repository* https://arxiv.org/abs/1802.06070

Fagan, G. (2005). Great battles of the ancient world. *Lecture Series in The Great Courses*. The Teaching Company https://www.thegreatcourses.com/courses/great-battles-of-the-ancient-world

Fallacy Man. (2016). Most scientific studies are wrong, but that doesn't mean what you think it means. *The Logic of Science*, May https://thelogicofscience.com/2016/05/10/most-scientific-studies-are-wrong-but-that-doesnt-mean what-you-think-it-means/

Fay, N., Garrod, S., Carletta, J. (2000). Group Discussion as Interactive Dialogue or as Serial Monologue: The Influence of Group Size. *Psychological Science*, November 1 https://doi.org/10.1111/1467-9280.00292

Fiala, Otto (2020). Resistance Operating Concept. *US Joint Special Operations University*, April https://www.hsdl.org/?abstract&did=838149

Floreano, D., Mitri, S., Magnenat, S., Keller, L. (2007). Evolutionary Conditions for the Emergence of Communication in Robots. *Current Biology* 17:514-9, April https://www.sciencedirect.com/science/article/pii/S0960982207009281

Friedman, R. (2012). Why Too Much Data Disables Your Decision Making. FastCompany, August 24 https://www.fastcompany.com/3000676/why-too-much-data-disables-your-decision-making

Geller, J. (2019). Does Call of Duty Believe in Anything? *YouTube* https://www.youtube.com/watch?v=FtCV421T52s

Gladwell, M. (2009). Cocksure: Banks, Battles, and the Psychology of Overconfidence. *New Yorker*, July www.newyorker.com/magazine/2009/07/27/cocksure

Goldman, S. (2021). The problem with academic funding, in 1 ironic scandal. *The Week*. Aug 23. https://theweek.com/talking-points/1004044/why-scholars-of-cheating-faked-their-own-results

Gooch, J., Perlmutter, A. (1982). Military Deception and Strategic Surprise. *Psychology Press*

Goodstein, D. (2010). On Fact and Fraud: Cautionary Tales from the Front Lines of Science. *Princeton University Press*

Graf, Jędrzej. (2021). Obrona Polski musi się opierać na sprawdzonych scenariuszach, a nie na dogmatach (The defense of Poland must be based on proven scenarios, not dogmas). *Defence24*, February 12 https://www.defence24.pl/obrona-polski-musi-sie-opierac-na-sprawdzonych-scenariuszach-a-nie-na-dogmatach?fbclid=IwAR0wiPbxmfEHte2QuqqbK-9N6rKD9RxsBhQXXuhFrLJDQE7L8RarcA9Gjsk

Greaves, A., Kurucar, J., Seater, R. (2020). Serious Games and Artificial Intelligence. *MIT Lincoln Laboratory and the Beaver Works Summer Institute* https://beaverworks.ll.mit.edu/CMS/bw/BWSI_Course_Serious_Game

Greenfieldboyce, N. (2018). What makes a leader. *NPR* August 2 https://www.npr.org/sections/health-shots/2018/08/02/634639437/what-makes-a-leader

Gupta, A. (2020). AI ethics groups are repeating one of society's classic mistakes. *MIT Technology Review*, 14 September https://www.technologyreview.com/2020/09/14/1008323/ai-ethics-representation-artificial-intelligence-opinion/

Hall, B., Wisniewski, K. (2021). Nashville bomber's bizarre writings reveal belief in aliens and lizard people. *News Channel 5*. January 2nd, 4th https://www.newschannel5.com/news/newschannel-5-investigates/nashville-bombers-bizarre-writings-reveal-belief-in-aliens-and-lizard-people

Handel, Michael I. (Ed) (1987). Strategic and Operational Deception in the Second World War. *Frank Cass*

Harangi-Tóth, Zoltán. (2020). Nemzetközi hadijáték trendek (International Wargaming Trends). *Honvédségi Szemle* February pp 113-124 https://doi.org/10.35926%2FHSZ.2020.2.11

Harrigan, P., Kirschenbaum, M (eds.) (2016). Zones of Control: Perspectives on Wargaming. *MIT Press*

Haselton, M. G., Nettle, D., Murray, D. (2016). The Evolution of Cognitive Bias. *The Handbook of Evolutionary Psychology*, (2nd edition) David M. Buss (ed.), pages 724–746, John Wiley & Sons Inc.

Hattori, Takushiro. (1949). On the Torago Wargame of 1943. *National Institute for Defense Studies Military History Archive, Tokyo* July 21

Heronyányi, József. (1987). A hadijátékok tervezésérõl és levezetésérõl. (About the planning and conduct of wargames). *Honvédelem*, 1987/11 pp 39–48

Heuer, R. (1999). Psychology of Intelligence Analysis. *CIA* https://www.cia.gov/resources/csi/books-monographs/psychology-of-intelligence-analysis-2

Hoffman, R. M. (General der Infanterie). (1952). War Gaming. U*S Army Foreign Military Studies Branch* Document P-094

Holowczak, A. (2020). Anti-Cheating in Online Chess. *Chess Magazine*, August https://www.newinchess.com/media/wysiwyg/product_pdf/8312.pdf

Honneth, A. (2008). Reification: A New Look at an Old Idea. *Oxford University Press*

Icke, D. (2007). The David Icke Guide to the Global Conspiracy (and how to end it). *David Icke Books Ltd*. October https://en.wikipedia.org/wiki/Reptilian_conspiracy_theory https://www.reddit.com/domain/forum.davidicke.com/

Ignatius, D. (2013). More Chatter than Needed. Washington Post, November 1 https://www.washingtonpost.com/opinions/david-ignatius-more-chatter-than-needed/2013/11/01/1194a984-425a-11e3-a624-41d661b0bb78_story.html?utm_term=.faca9cdfba1

IJA. (unknown). A Wargame Study on Northern Manchuria. *National Institute for Defense Studies Military History Archive, Tokyo*

Inose, Naoki. (1983). Showa 16-nen Natsu no Haisen (Defeated in War in the Summer of 1941). *Chuko Bunko, Tokyo*

Jackson, S. et al. (2013). Illuminati: The Game of Conspiracy. *Steve Jackson Games* http://www.sjgames.com/illuminati/img/illuminati_rules.pdf

Jervis, R. (2006). Understanding Beliefs. *Political Psychology* 27, 5, October https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9221.2006.00527.x

Johnson, D., Wrangham, R., Rosen, S. (2002). Is Military Incompetence Adaptive? An Empirical Test with Risk-Taking Behaviour in Modern Warfare. *Evolution and Human Behavior*, 23, pp. 245–64 https://www.sciencedirect.com/science/article/abs/pii/S1090513801000873

Karpowitz, C., Mendelberg, T. (2014). The Silent Sex. *Princeton University Press*.

Kasprak, A. (2020). How To 'Weaponize' Misleading Narratives About Voting: Lessons from Trump-Linked PILF Playbook. *Snopes* 15 October https://www.snopes.com/news/2020/10/15/the-fog-of-war/

Kelly J., Sadeghieh T., & Adeli K. (2014). Peer Review in Scientific Publications: Benefits, Critiques, & A Survival Guide. *EJIFCC*. 2014;25(3):227-243. Oct 24 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4975196/

Kim, Tae-Hyo, (2000), A Simulation: Possibilities and Limits of ROK-Japanese Naval Cooperation. In *Korea-Japan Security Relations: Prescriptive Studies, Sang-Woo Rhee and Tae-Hyo Kim (eds), pp. 239-258*, Oruem Publishing House, Seoul, South Korea

Klein, A. (Ed) (1955). Grand Deception: The World's Most Spectacular and Successful Hoaxes, Impostures, Ruses and Frauds. *Faber & Faber*

Klein, A. (Ed) (1958). The Double Dealers: Adventures in Grand Deception. *Faber & Faber*

Kocher, M., Schudy, S. & Spantig, L. (2016). I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups. *University of Munich* Discussion Paper 2016-8 https://doi.org/10.5282/ubm/epub.28966

Kronzek, A. (2008). 52 Ways to Cheat at Poker. *Penguin Random House, Canada*

Kwiatkowski, K. (2016). Civilian Casualties: shifting perspective in This War of Mine. In *Zones of Control: Perspectives on Wargaming*, Pat Harrigan &Matthew G. Kirschenbaum (eds.) pages 691–701, MIT Press

Langer, E. (1975). The Illusion of Control. *Journal of Personality and Social Psychology*, 32, no. 2, August, pp. 311–28 https://psycnet.apa.org/record/1977-03333-001

Lazar, S. (2016). Jus in Bello. Stanford Encyclopedia of Philosophy, Section 4 of "War" https://plato.stanford.edu/entries/war/#JusBell

Lehrer, J. (2012). Groupthink: The brainstorming myth. *The New Yorker*, January 20 https://www.newyorker.com/magazine/2012/01/30/groupthink

Leveson, N. (2016). Engineering a Safer World: Systems Thinking Applied to Safety. *MIT Press*, reprint edition (originally 2012)

Littman. Michael L. (2021). Collusion Rings Threaten the Integrity of Computer Science Research. *Communications of the ACM*, Vol. 64 No. 6, pp 43-44, June https://dl.acm.org/doi/pdf/10.1145/3429776

Liulevicius, V. (2006). World War I: The "Great" War. *Lecture Series in The Great Courses*. The Teaching Company

Longley-Brown, G. (2019). Successful Professional Wargames: A Practitioner's Handbook. *History of Wargaming Project* www.wargaming.co/professional/details/professionalhandbook.htm

Lorenz, J., Rauhut, H., Schweitzer, F. & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, May https://www.pnas.org/content/pnas/108/22/9020.full.pdf

MacDougall, R., Faden, L. (2016). Simulation Literacy: The Case for Wargames in the History Classroom. In *Zones of Control: Perspectives on Wargaming*, Pat Harrigan & Matthew G. Kirschenbaum (eds.), pages 447-454. MIT Press

Marcus, R. (2006). Dirty Poker. *D&B Publishing*

Mazar, N., Amir, O. & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research*, December 2008, Vol. 45, No. 6, pp. 633-644 https://journals.sagepub.com/doi/pdf/10.1509/jmkr.45.6.633

McDermott, P., Dominguez, C., Kasdaglis, N., Ryan, M., Trahan, I., Nelson, A. (2018). Human-Machine Teaming Systems Engineering Guide. *MITRE Corporation* Technical Paper MP180941 https://www.mitre.org/publications/technical-papers/human-machine-teaming-systems-engineering-guide

McShane, L. (2020). A trout in the milk? *The Spectator*, 10 October https://www.spectator.co.uk/article/a-trout-in-the-milk-

Miller, B. (2014). Unequal demands on women for university service harm careers. *Riverside University of California*, August https://www.sciencedaily.com/releases/2014/08/140822083626.htm

Miller, B. J. (2016). The Application of Statistical and Forensics Validation to Simulation Modeling in Wargames. In *Zones of Control: Perspectives on Wargaming*, Pat Harrigan & Matthew G. Kirschenbaum (eds.) pages 183–200, MIT Press

MIT (2021). Committee on the Use of Humans as Experimental Subjects (COUHES). *Massachusetts Institute of Technology* https://couhes.mit.edu/

Mizokami, K. (2020). The U.S. Lost a (Fictional) War with Iran 18 Years Ago. *Popular Mechanics* https://www.popularmechanics.com/military/a30392654/millennium-challenge-qassem-soleimani/

Monteiro, M. (2019). Ruined by Design: How Designers Destroyed the World, and What We Can Do to Fix It. *Mule Books*

Morimatsu, Toshio. (1983). Souryokusen Kenkyusho (The Total War Research Institute). *Hakuteisha, Tokyo*

Morris, C. (2020). Don't Jump Into Unconscious Bias Training Until You Read This. *Forbes*, July https://www.forbes.com/sites/carmenmorris/2020/07/21/dont-jump-into-unconscious-bias-training-until-you-read-this/?sh=15d132983f41

Mullen, B., Johnson, C. & Salas, E. (1991). Productivity Loss in Brainstorming Groups: A Meta-Analytic Integration. *Journal of Basic and Applied Social Psychology*, Volume 12, 1991 - Issue 1, pp. 3-23, 7 June https://doi.org/10.1207/s15324834basp1201_1

Murray, S. (2016). Upending Militarized Masculinity in Spec Ops: The Line. In *Zones of Control: Perspectives on Wargaming*, Pat Harrigan & Matthew G. Kirschenbaum (eds.), pages 319–328, MIT Press

NIDS. (1967). Hawai Sakusen (Naval Operations in Hawaii). *Asagumo Shinbun, Tokyo* http://www.nids.mod.go.jp/military_history_search/SoshoView?,=010

NIDS (1971). Middowei Kaisen (Battle of Midway). *Asagumo Shinbun, Tokyo* http://www.nids.mod.go.jp/military_history_search/SoshoView?kanno=043#

Nijstad, B., Stroebe, W. & Lodesijkx, H. (2006). The illusion of group productivity: A reduction of failures explanation. *European Journal of Social Psychology*, 36, 31-48 https://doi.org/10.1002/ejsp.295

Nyhan, B., Reifler, J. (2010). When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior* 32, no. 2, June, pp. 303–30 https://link.springer.com/article/10.1007/s11109-010-9112-2

Odom, J. (2019). Understanding China's Legal Gamesmanship in the Rules-Based Global Order. *Daniel K. Inouye Asia-Pacific Center for Security Studies*, February https://apcss.org/wp-content/uploads/2019/09/12-Chinas_Legal_Gamesmanship-odom.pdf

Owen, W. (2019). Concerns about Professional Wargames: A "Red Team" view bigger than COTS/MOTS. *Connections UK 2019 Wargaming Conference* https://www.professionalwargaming.co.uk/2019.html

Papasavva, A. et al. (2020). "Is it a Qoincidence?": A First Step Towards Understanding and Characterizing the QAnon Movement on Voat.co. *Computers and Society* https://arxiv.org/abs/2009.04885v3

Perla, P. (2012). Peter Perla's The Art of Wargaming. *History of Wargaming Project*

Perla, P. (2016). Operations Research, Systems Analysis, and Wargaming: Riding the Cycle of Research. In *Zones of Control: Perspectives on Wargaming*, Pat Harrigan & Matthew G. Kirschenbaum (eds.) pages 159–182, MIT Press

Perrow, C. (1999). Normal Accidents: Living with High-Risk Technologies - Updated Edition. *Princeton University Press*, September (orig. 1984)

Pew. (2021). The Future of Gamification. *Pew Research Center*, May https://www.pewresearch.org/internet/2012/05/18/the-future-of-gamification/

Podder, S. (2021). Robinhood to scrap confetti animation on app after 'gamification' criticism. *Nasdaq*, March https://www.nasdaq.com/articles/robinhood-to-scrap-confetti-animation-on-app-after-gamification-criticism-2021-03-31

Pogácsás, Krisztián (2013), Terepasztalos Parancsnoki Hadijáték, avagy a parancsnoki döntéshozatal és harcászati szemlélet fejlesztésének eszköze. (Sand Table Wargaming as a development tool for the commander's decision making). *Seregszemle*, Feb/Mar, pp 83–89.

Press, W., Dyson, F. (2012). Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences of the United States of America*, June 26 https://www.pnas.org/content/109/26/10409

Price, T. (1985). Home Front 86. *Bulldog Publications,* May

Robinson, M. & Downes-Martin, S. (Eds) (2018). In-Stride Adjudication. *Connections US 2018 Wargaming Conference* https://paxsims.files.wordpress.com/2018/09/in-stride-adjudication-working-group-report-20180908.pdf

Roozembeek, J., Van Der Linden, S. (2020). Breaking Harmony Square: A game that "inoculates" against political misinformation. *Harvard Misinformation Review* November 6 https://misinforeview.hks.harvard.edu/article/breaking-harmony-square-a-game-that-inoculates-against-political-misinformation/

Ross, L., Anderson, C. (1982). Shortcomings in the Attribution Process: On the Origins and Maintenance of Erroneous Social Assessment. In *Judgment under Uncertainty: Heuristics and Biases*, Daniel Kahneman, Paul Slovic, & Amos Tversky (Eds), Cambridge University Press 1982, pp. 129–52

Salen, K., Zimmerman, E. (2003). Rules of Play: Game Design Fundamentals. *MIT Press*

Seater, R., Rose, C., Norige, A., McCarthy, J., Kozar, M., DiPastina, P. (2016). Skill Transfer and Virtual Training for IND Response Decision Making: Game Design for Disaster Response Training. *MIT Lincoln Laboratory* Technical Report 1207 https://apps.dtic.mil/sti/pdfs/AD1033763.pdf

Sepinsky, J. (2021). Is it a Wargame? It Doesn't Matter: Rigorous Wargames Versus Effective Wargaming. *War on the Rock*s February 24 https://warontherocks.com/2021/02/is-it-a-wargame-it-doesnt-matter-rigorous-wargames-versus-effective-wargaming/

Setiya, K. (2018). Intention. *The Stanford Encyclopedia of Philosophy* https://plato.stanford.edu/entries/intention/

Shermer, M. (2002). Why People Believe Weird Things. *Henry Holt & Co*

Shermer, M. (2010). When scientists sin: Fraud, deception and lies in research reveal how science is (mostly) self-correcting. *Scientific American*, 303, no. 1 (July), p. 34 https://www.scientificamerican.com/article/when-scientists-sin/

Shultz, R., Godson, R. (1984). Dezinformatsia. *National Strategy Information Center*

Silver, D. et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529, 484–489, January https://www.nature.com/articles/nature16961

Silver, D. et al. (2017). Mastering the game of Go without human knowledge. *Nature*, 550, 354–359, October https://www.nature.com/articles/nature24270

Simmons, R. (2016). Goal-Driven Design and Napoleon's Triumph. In *Zones of Control: Perspectives on Wargaming*, Pat Harrigan & Matthew G. Kirschenbaum (eds.), pages 201–207, MIT Press

Simon, H. (1956). Rational Choice and the Structure of the Environment. *Psychological Review*, 63 (2): 129–138 https://uk.sagepub.com/sites/default/files/upm-binaries/25239_Chater~Vol_1~Ch_03.pdf

Singer, E. (2015). Game Theory Calls Cooperation Into Question. *Quanta Magazine*, February 12 https://www.quantamagazine.org/game-theory-explains-how-cooperation-evolved-20150212/

Somuchpoker. (2017). The five cheating scandals that shocked the live tournament world. *Youtube*, June 28 https://www.youtube.com/watch?v=CvA6zdEaT1g

Steger, I. (2017). Artificial intelligence has brought "doubt and suspicion" to the ancient world of Japanese chess. *Quartz Magazine*, February 16 https://qz.com/906447/artificial-intelligence-has-brought-doubt-and-suspicion-to-the-ancient-world-of-shogi-or-japanese-chess/

Sternberg, R. (2002). Why Smart People can be so Stupid. *Yale University Press*

Surowiecki, J. (2004). The Wisdom of Crowds. *Anchor Books*

Szűcs, Péter. (2020). Katonai vezetői gondolkodás fejlődését befolyásoló tényezők a rendszerváltástól napjainkig (Factors Influencing the Development of Military Leadership Thinking from the Change of Regime to Date). *Hadtudományi Szemle*, February pp 155-161 https://doi.org/10.32563/hsz.2020.2.12

Takahashi, D. (2013). Brenda Romero's Train board game will make you ponder. *VentureBeat*, May https://venturebeat.com/2013/05/11/brenda-romero-train-board-game-holocaust/

Tenner, E. (1997). Why Things Bite Back: Technology and the Revenge of Unintended Consequences. *Vintage*, 2nd edition, September (orig. 1996)

Tetlock, P. & Gardner, D. (2015). Superforecasting: The Art and Science of Prediction. *Crown Publishing*

Tetlock, P. (2005). Expert Political Judgment: How Good Is It? How Can We Know? *Princeton University Press*

Thaler, R., Tversky, A., Kahneman, D. & Schwartz, A. (1997). The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test. *The Quarterly Journal of Economics* Vol. 112, No. 2, May, pp. 647-661 https://doi.org/10.1162/003355397555226

Train, B. (2010). Guerrilla Checkers. *Board Game Geek* https://boardgamegeek.com/boardgame/71035/guerrilla-checkers

Train, B., Ruhnke, V. (2016). Chess, Go, and Vietnam: Gaming Modern Insurgency. In *Zones of Control: Perspectives on Wargaming*, Pat Harrigan & Matthew G. Kirschenbaum (eds.), pages 513–529, MIT Press

Tversky, A., Thaler, R. (2019). Anomalies: Preference Reversals. *Journal of Economic Perspectives*, Vol 4 No 2 Spring https://www.aeaweb.org/articles?id=10.1257/jep.4.2.201

UKMoD. (2017). Wargaming Handbook. *UK Ministry of Defence Development, Concepts and Doctrine Center* https://www.gov.uk/government/publications/defence-wargaming-handbook

Unknown Author. (circa 400 CE). Hnefatafl (Traditional tabletop board game). https://boardgamegeek.com/boardgame/2932/hnefatafl

Uri, S., Nelson, L., Simmons, J. (2021). Evidence of Fraud in an Influential Field Experiment About Dishonesty. *Data Colada*. August 17 https://datacolada.org/98

USAMRDC. US Army Medical Research and Development Command. Human Research Protection Office (HRPO). https://mrdc.amedd.army.mil/index.cfm/collaborate/research_protections/hrpo

USDOD. (2020). DOD Adopts Ethical Principles for Artificial Intelligence. *US DoD Newsroom,* February 24 https://www.defense.gov/Newsroom/Releases/Release/Article/2091996/dod-adopts-ethical-principles-for-artificial-intelligence/

Vane, R. (2016). Hypergaming. In *Zones of Control: Perspectives on Wargaming*, Pat Harrigan & Matthew G. Kirschenbaum (eds.), pages 241–24, MIT Press

Vaughan, D. (2016). The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA. *University of Chicago Press* (enlarged edition), January (orig. 1996)

Wallman, J. (2016). Cultural Wargaming: Understanding Cross-Cultural Communications Using Wargames. In *Zones of Control: Perspectives on Wargaming*, Pat Harrigan & Matthew G. Kirschenbaum (eds.), pages 545–551, MIT Press

Walzer, M. (1978). Just and Unjust Wars: A Moral Argument with Historical Illustrations. *Allen Lane*, London

Vego, M. (2012). German War Gaming. *Naval War College Review* Vol 65, No 4 https://digital-commons.usnwc.edu/nwc-review/vol65/iss4/10/

Weinstein, E. (2011). What Scientific Concept Would Improve Everybody's Cognitive Toolkit? *Edge* https://www.edge.org/response-detail/11783

Wells, H. G. (1913). Little Wars: With an Appendix on Kriegspiel. http://www.gutenberg.org/files/3691/3691-h/3691-h.htm

Weuve, C., Perla, P., Markowitz, M., Rubel, R., Downes-Martin, S., Martin, M., Vebber, P. (2004). Wargame Pathologies. *CNA* Report CRM D0010866.A1/Final September https://www.cna.org/CNA_files/PDF/D0010866.A1.pdf

Whaley, B. (2016). Practice to Deceive: Learning Curves of Military Deception Planners. *US Naval Institute Press*

Whaley, B. (2016). Turnabout and Deception: Crafting the Double Cross and the Theory of Outs. *US Naval Institute Press*

Wheaton, Kristan. The Mind's Lie. *Board Game Geek* https://boardgamegeek.com/boardgame/139711/minds-lie

Wills, G. (1994). Certain Trumpets: The Nature of Leadership. *Simon & Schuster*

Wong, Y. (2016). Irregular Warfare: The Kobayashi Maru of the Wargaming World. In *Zones of Control: Perspectives on Wargaming*, Pat Harrigan & Matthew G. Kirschenbaum (eds.), pages 531–537, MIT Press

Wong, Y., Bae, S., Bartels, E. & Smith B. (2019). Next-Generation Wargaming for the U.S. Marine Corps. *RAND* www.rand.org/pubs/research_reports/RR2227.html

Zeeuw, D. et al. (2020). Tracing normiefication: A cross-platform analysis of the QAnon conspiracy theory. *First Monday*, Volume 25, Number 11 https://firstmonday.org/ojs/index.php/fm/article/download/10643/9998 *https://dx.doi.org/10.5210/fm.v25i11.10643*

Zenko, M. (2015). Millennium challenge: The Real Story of a Corrupted Military Exercise and its Legacy. *War on the Rocks*, November https://warontherocks.com/2015/11/millennium-challenge-the-real-story-of-a-corrupted-military-exercise-and-its-legacy/

Zsolnai-Feher, Károly. (2018). 4 Experiments Where the AI Outsmarted Its Creators, *YouTube* https://www.youtube.com/watch?v=GdTBqBnqhaQ&list=RDCMUCbfYPyITQ-7l4upoX8nvctg&index=2&ab channel=TwoMinutePapers