WARGAMING AND AGENT-BASED MODELLING AS HYBRID EPISTEMIC INSTRUMENTS: ONTOLOGY, EPISTEMOLOGY, AND A RESEARCH AGENDA

INTRODUCTION

Wargaming has again become central to how scholars and practitioners think about conflict, crisis, emerging technology, and decision-making under uncertainty. In military institutions, it is used to test concepts, examine adversary reactions, and support planning. In international relations, it is increasingly treated as a method for studying rare events, microfoundations of theory, escalation, and strategic interaction. In policy analysis, it is often valued precisely because it appears to offer a way to think through futures that are consequential, uncertain, and not yet empirically observable (Banks, 2024; Lin-Greenberg, Pauly, and Schneider, 2022; Valeriano and Jensen, 2021).

At the same time, computational methods are pressing into this same space. Artificial intelligence, agent-based modelling (ABM), modelling and simulation (M&S), and design-of-experiments approaches are increasingly proposed as ways to boost or transform wargaming. Recent work suggests that AI could generate adversary courses of action, ABMs could represent civilian and human-domain dynamics, and integrated modelling-wargaming architectures could improve the analysis of complex operational environments (Chance et al., 2026; Davis and Bracken, 2025; James et al., 2017). This convergence is often described in practical terms: computational tools are faster, can run more scenarios, can represent more actors, and can reduce burdens on human adjudicators. But the convergence is also philosophical. The incorporation of computational methods changes the character of the wargame itself. It also changes what counts as evidence, where expertise sits, and what kind of conclusions users can reasonably draw.

The literature has not fully caught up with this shift. Wargaming scholarship has made substantial progress on its ontology and epistemology. Authors such as Perla, Rubel, Banks, Bartels, Lin-Greenberg and colleagues, Enstad, and Hulterström have shown that wargaming cannot be treated as a simple prediction device or reduced to a narrow positivist model of science. Yet this literature remains focused mainly on human-played or professionally facilitated games. Conversely, the literature on ABM and computational military simulation often makes strong claims about complexity, emergence, distributed interaction, or AI support, but usually without integrating those claims into the deeper philosophical questions already raised by the wargaming literature. What is missing is a framework for thinking about what happens when the two are combined. We have one body of work that asks what knowledge wargaming yields, and another that asks what computational models can do, but far less work that asks how these questions relate to one another.

This article addresses that gap. Its central argument is not simply that agent-based modelling can improve wargaming, but that combining wargaming and ABM produces a different kind of epistemic arrangement. Wargaming and ABM should be understood as complementary but distinct practices that assign different jobs to human players, computational agents, adjudication, and analysis. Wargaming is especially good at bringing out human judgment, strategic interaction, and the lived experience of decision-making under uncertainty. Yet it remains limited by opaque adjudication, weak mechanism specification, difficulties in representing distributed background dynamics, and the challenge of systematic variation and robustness testing. ABM can help address some of these weaknesses by formalizing heterogeneous interaction and representing emergent effects, while enabling structured exploration of assumptions. But ABM does not automatically solve wargaming's philosophical problems. Instead, it introduces its own issues concerning abstraction, representation, validity, causal interpretation, and policy use. So the challenge is not just to combine methods technically. It is to explain what kind of inquiry this hybrid actually is, and what kind of knowledge it can produce.

For practitioners, the immediate implication is that hybrid game design should begin by deciding which parts of a problem are best represented through human play, which through formal models, and how adjudication links the two. Strategic bargaining, role conflict, interpretation, and improvisation may be better elicited through players, while civilian response, information diffusion, or cascading infrastructure effects may be better represented computationally. The design problem is therefore not simply whether to add modelling, but where to place it and what kind of inference the resulting arrangement can support.

The argument proceeds in five steps. The first section reconstructs the ontology and epistemology of wargaming, arguing that wargames are hybrid objects and practices whose outputs are best understood as situated and conditional forms of knowledge. The second section identifies unresolved gaps in wargaming that make computational augmentation attractive, especially problems of mechanism, background dynamics, and systematic exploration. The third section develops a philosophical account of ABM and related computational methods, arguing that they are best understood as generative and exploratory instruments rather than forecasting machines. The fourth section proposes an integrated framework for combining wargaming and ABM as hybrid epistemic instruments. The fifth section develops a research agenda for such combinations across philosophical, methodological, applied, and institutional dimensions.

WARGAMING AS A HYBRID EPISTEMIC PRACTICE

Ontology: What Kind of Thing Is a Wargame?

The first philosophical challenge is ontological. Before asking what wargames tell us, one must ask what kind of thing a wargame is. This question is more difficult than it appears, because wargames are often described in several different and only partially compatible ways. They are called models, simulations, games, exercises, narratives, laboratories, and methods. Some of these descriptions are enlightening while others blur important distinctions.

Perla's classic formulation remains a strong starting point. For him, a wargame is a model of warfare in which the flow of events is shaped by the decisions of players representing opposing sides (Perla, 1985). This definition matters because it distinguishes wargaming both from real operations and from purely analytical or computational modelling. A wargame is not war itself but is a designed representation of warfare. But neither is it just a static model, because player decisions actively shape events. Decision and opposition are not added features. They are part of what makes the activity a wargame rather than another kind of simulation.

Even so, Perla's formulation is only the beginning. Wargames are not merely models. They are also games in a stronger sense: they are structured fields of action in which actors pursue partly contradictory objectives under rules and constraints. They are also narrative environments. Perla and McGrady's influential account of "story-living" emphasizes that wargames are not just stories about conflict, but participatory settings in which players inhabit a scenario and take responsibility for choices within it (Perla and McGrady, 2011). This narrative dimension is not ornamental. It shapes the way players understand stakes, causality, and consequence.

Banks sharpens the ontological picture by arguing that analytic wargames contain a distinctive "methodological machinery": representation, consequential decision-making, adjudication, immersion, and bespoke design (Banks, 2024). Banks's list matters because it breaks the wargame down into its operational parts: representation, decision, adjudication, immersion, and design. Wargames are representational, but not merely descriptive. They are immersive, but not simply theatrical. They involve decisions, but those decisions matter because they are adjudicated into consequences. And they are bespoke, meaning they are tailored to specific problems, rather than generated from one universal template.

This synthesis suggests that wargames are ontologically hybrid in at least four respects: First, they are *representational artefacts*. They stand in for actual or possible events in war, crisis, or competition. This entails abstraction, simplification, and selective focus. Like other models, they are not replicas of reality. They are purposeful simplifications. Second, they are *interactive games*. Their internal logic depends on the presence of opposition, choice, uncertainty, and consequence. If all relevant dynamics are fixed in advance and no actor exercises meaningful agency, one no longer has a wargame in the strong sense. Third, they are *narrative environments of action*. They shape understanding not merely by showing a scenario, but by placing actors inside a scenario where meaning unfolds through participation. This is part of why they can have educational and cognitive power beyond abstract analysis. Fourth, they are *institutional and professional practices*. Wargames do not exist in a vacuum. They are embedded in military organizations, research programs, policy communities, and educational institutions. Their meaning depends partly on who plays them, why they are organized, how results are interpreted, and what larger organizational purposes they serve. Enstad's treatment of wargames and exercises as contributors to professional knowledge makes this point especially clearly: wargaming is not just an instrument external to practice; it helps shape practice itself (Enstad, 2022). A wargame is a model enacted through consequential play.

The educational literature reinforces this plural ontology. Hulterström's philosophical account of educational wargaming treats games as abstract portrayals of armed conflict in which players engage with fictional facts, fictional problems, and fictional challenges. The point is not that wargames perfectly reproduce combat, but that they create disciplined environments of make-believe in which meaningful forms of judgment, adaptation, and skill may be cultivated (Hulterström, 2024). This moves the ontological discussion further. Wargames are not just about what they represent but they are also about what they enable participants to do.

Thus, wargaming is best understood as a *hybrid, designed, interactive representation of conflict embedded in practice*.

Epistemology: What Kind of Thing Does Wargaming Produce?

If that is what a wargame is, what kind of knowledge does it yield? This question has long haunted the method. Practitioners often celebrate wargaming as insightful or useful, but the epistemic status of those insights has remained controversial.

The strongest consensus in the literature is negative: wargames do not primarily produce prediction or proof in any strong scientific sense. Rubel is explicit that wargames are not forecasting machines. Their outputs are at best indicative rather than definitive. They can reveal potential relationships and possibilities, but they do not supply certain knowledge of future outcomes (Rubel, 2006). This is not a weakness unique to wargaming; it reflects the nature of the problems to which wargaming is often applied. War and strategic competition are weakly structured domains, saturated with contingency, adaptation, and hidden variables. In such environments, demand for exact prediction is often misplaced.

Perla makes a complementary point. Wargames are most valuable when used to investigate processes rather than calculate results. What matters is not simply who wins in a game, or what losses occur, but why events unfolded as they did, what decisions mattered, what uncertainties shaped action, and what hidden assumptions were exposed (Perla, 1985). In that sense, the core epistemic value of wargaming lies in process knowledge.

From the literature, at least five kinds of knowledge claims associated with wargaming can be distinguished: The first is *exploratory knowledge*. Wargames can surface issues, possibilities, vulnerabilities, and interaction effects that were not already visible to participants. Mason emphasizes this exploratory role in his discussion of wargames as tools for extending horizons, testing ideas, and generating innovation under uncertainty (Mason, 2018). Exploratory knowledge is not trivial. In many policy settings, discovering a plausible but previously ignored pathway or problem is already valuable. The second is *process knowledge*. Wargames can reveal how choices interact over time, how actors interpret signals, how escalation pathways emerge, and how responses feed back into future options. This is one reason wargames have become increasingly attractive in international relations research. Lin-Greenberg, Pauly, and Schneider argue that wargames can illuminate the microfoundations of theory by showing how decisions actually emerge through group deliberation, interpretation, and consequence-based response (Lin-Greenberg, Pauly, and Schneider, 2022). Their point is not that games perfectly mirror real-world decision-making, but that they can provide access to mechanisms and processes otherwise hard to observe. The third is *experiential or practical knowledge*. Perla and McGrady argue that wargames work partly because they create synthetic experiences that affect how participants think, feel, and behave. Games do not merely communicate propositions; they create a form of lived engagement with choice and consequence (Perla and McGrady, 2011). Enstad pushes the point further by arguing that wargaming shapes professional knowledge in a broader sense: it helps form how officers interpret situations, what they attend to, and what responses come to seem natural or appropriate (Enstad, 2022). This suggests that wargaming can yield more than propositional knowledge. It can contribute to know-how, judgment, and professional orientation. The fourth is *counterfactual or mechanism-sensitive knowledge*. Valeriano and Jensen argue that wargames can be useful to social science because they allow scholars to study alternative realities in structured ways, especially where direct observation is impossible or dangerous (Valeriano and Jensen, 2021). This does not mean they deliver universal law-like claims. It means they can help examine how different choices, assumptions, or interactions might matter under plausible conditions. The fifth is *ecologically informed understanding*. Lin-Greenberg and colleagues argue that one of the distinctive strengths of wargaming as a research method may lie in ecological validity: because wargames place participants in interactive, consequence-laden, role-based environments, they may sometimes mirror relevant features of actual decision settings more effectively than other methods (Lin-Greenberg, Pauly, and Schneider, 2022). This is not a blanket claim that wargames are always more realistic. Rather, it is a claim that under some conditions they may better capture the social and cognitive texture of real strategic decision-making.

Taken together, these points suggest that wargaming's epistemology is plural but not unlimited. Wargames can be highly valuable, yet the kind of value they offer is usually more modest and more situated than users sometimes assume. They are strong at generating exploratory, processual, experiential, and mechanism-sensitive understanding. They are weak at delivering high-confidence prediction. They are often better at asking "what could happen, why, and through what interactions?" than "what will happen?"

This is why the art-versus-science debate has been increasingly challenged. Bartels argues that the real issue is not whether gaming is art or science, but what kind of science it can be under different philosophies of science (Bartels, 2019). Banks makes a related argument from another direction: rather than importing one epistemic standard from outside, the field should investigate the recurring methodological features of wargaming and ask how those features generate knowledge on their own terms (Banks, 2024). Both authors are pushing back against the old choice between calling wargaming either art or science. Wargaming does not need to mimic laboratory experimentation to be rigorous. But neither is it exempt from methodological scrutiny because it involves craft or artistry.

The Inferential Limits of Wargaming

The wargaming literature has become much clearer about what games can do well. It is equally important to be clear about what they do poorly, because those weaknesses create the space in which computational methods become attractive.

The first recurring weakness is *adjudicative opacity*. Since adjudication determines how actions translate into consequences, it is one of the central epistemic engines of any wargame. Yet in many games it remains partly subjective, weakly documented, or dependent on expert judgment that is difficult to audit after the fact. Even when expert adjudication is defensible, it can obscure causal interpretation. One may know that a certain game move led to a certain effect, yet remain unsure whether that outcome arose from player choice, scenario structure, adjudicator intuition, or hidden rule logic. The second weakness is *weak mechanism specification*. Wargames often surface plausible sequences and important interactions, but the mechanisms producing those outcomes are not always explicit. If a red team behaves aggressively, if civilians react unexpectedly, or if escalation takes a surprising course, one still must ask: what exactly generated that behavior? Was it role interpretation, cultural assumption, specific scenario prompts, group dynamics, or a more general mechanism that might matter elsewhere? Wargames can reveal mechanisms in an informal sense, but often do so without fully formalizing them. The third weakness is *limited systematic variation*. Many professional and academic wargames are one-off or small-N events. This restricts the ability to explore parameter space, compare repeated runs, or assess robustness across different assumptions and participant pools. Even well-designed research games rarely achieve the kind of structured sensitivity analysis available in formal modelling. The fourth weakness is *dependence on player selection and expertise*. Ecological validity may be an advantage when representative or expert participants are available, but it becomes a liability when they are not. Player background, role commitment, incentives, and interpretive style can shape outcomes heavily. The literature increasingly recognizes this, but the problem remains structurally difficult to solve. The fifth weakness is *difficulty representing background dynamics*. Many games foreground commanders, negotiators, staffs, or elite decision-makers. This is often appropriate. But it can also mean that broader social, civilian, informational, infrastructural, or institutional processes are treated simplistically. Civilian populations may appear only as injects or as static scenario conditions. Information flows may be compressed into a few pre-scripted events. Logistics, legitimacy, rumor, local adaptation, or humanitarian disruption may be underdeveloped. Yet precisely these dynamics often shape real outcomes. The sixth weakness is *institutional and sponsorship bias*. Rubel warns that wargames can generate "valid-looking garbage" when poor design or organizational politics distort their outputs (Rubel, 2006). Curry's historical analysis likewise shows that professional wargames have often been wrong, even when taken seriously by their institutions, because they encoded their period's assumptions more than future reality (Curry, 2020). Wargames can be dangerously persuasive, precisely because they feel realistic and consequential.

These weaknesses do not refute wargaming. They define its inferential limits. Once those limits are seen clearly, the issue is not merely whether computational methods can be added to wargaming as technical supplements. It is whether the introduction of computational components changes how representation, adjudication, mechanism, and inference are distributed within the inquiry itself. The relevant question is therefore not only what computational augmentation can do, but what kind of epistemic arrangement it creates.

OPEN QUESTIONS IN WARGAMING AND THE SPACE FOR COMPUTATIONAL METHODS

What Exactly is Learned in a Wargame?

A first unresolved question is how to classify wargaming outputs more precisely. The literature offers several answers: insight, process understanding, synthetic experience, practical know-how, and professional judgment. However, existing studies do not fully harmonize them.

Part of the difficulty is that different kinds of games aim at different outputs. Analytical games, educational games, professional military education games, exploratory policy games, and concept development games all seek somewhat different epistemic goods. Hulterström's distinction between educational and analytical wargames is helpful here. Educational games are aimed primarily at developing know-how and battlefield acumen while analytical games are more concerned with understanding specific situations, interactions, and possibilities (Hulterström, 2024). Lin-Greenberg and colleagues likewise show that research games may seek both outcome data and deliberative data, each serving different inferential purposes (Lin-Greenberg, Pauly, and Schneider, 2022).

Still, a deeper issue remains. Even when a game produces "insight," what is the epistemic status of that insight? Is it a refined hypothesis? A contextual warning? A demonstration of mechanism? A prompt for further analysis? A realistic rehearsal of elite reasoning? The field is better at defending wargaming's value than at sharply classifying its outputs. This matters because without such classification, discussions of validity and integration remain vague. Computational methods become attractive partly because they seem to promise sharper traceability between assumptions and outputs. But whether that promise is realized depends on one's account of what counts as a wargame output in the first place.

How Should Background Dynamics be Represented?

A second unresolved question concerns the representation of background dynamics. Many wargames focus, understandably, on strategic or operational decision-makers. Yet real conflict environments are shaped by far more than the choices of formal leaders. Civilian populations, NGOs, local authorities, logistics systems, information ecologies, cultural norms, infrastructures, and second-order effects all matter. They matter especially in irregular conflict, stabilization, urban warfare, cyber operations, humanitarian crises, and civil-military interaction.

The challenge is that many of these processes are difficult to represent convincingly in traditional seminar or manual games. They are often converted into control-cell interventions, event cards, or scripted assumptions. This allows them to be included, but often in a highly compressed and under-mechanized way. James and colleagues make this point forcefully in their argument that Army wargaming often struggles to represent the Human Domain with sufficient rigor, especially under tactical time constraints and when dealing with nonstate actors or civilian populations (James et al., 2017). Van Lieburg, Petiet, and Le Grand similarly argue that modern conflict environments involve active noncombatants and complex socio-cultural dynamics that many existing wargames handle poorly (van Lieburg, Petiet, and Le Grand, 2007).

This problem is not merely additive. It is not solved by including "more civilians" in a scenario. The issue is how to represent distributed interactions, local adaptation, indirect effects, and nonlinearity without either overwhelming the game or reducing the social world to a few arbitrary knobs. This is one area in which ABM is especially relevant, because ABM is specifically designed to represent heterogeneous agents interacting locally and generating macro-patterns that are not specified from above.

Causal Grounding in Wargaming

A third unresolved question concerns causal grounding. Wargames often reveal plausible pathways, but users can too easily treat those pathways as robust causal findings.

The causal problem in wargaming has several layers. One must distinguish at least four possible sources of an observed effect: player reasoning, rule structure, scenario assumptions, and adjudication practice. If an outcome recurs across repeated games, one may suspect a deeper mechanism. But in small-N or bespoke settings, causal grounding often remains ambiguous. This is one reason Banks insists that understanding wargaming's epistemology requires close attention to its recurrent methodological features, including adjudication and representation, rather than borrowing the language of "findings" too casually from other methods (Banks, 2024).

Computational models, especially ABM, appear attractive here because they can make behavioral rules, interaction structures, and environmental conditions more explicit. But explicitness alone is not enough. A formally specified mechanism is not thereby a true or validated mechanism. Computational methods can clarify causal assumptions, but they can also lend false precision to assumptions that remain deeply contestable. The open question, then, is not whether computational methods can improve causal clarity in some sense. They can. The real question is what sort of causal claims hybrid systems can legitimately support.

How Can Exploration Be Made More Systematic?

A fourth unresolved question is how to test robustness beyond case-based, episodic play. Goodman, Risi, and Lucas make this point from the AI side when they argue that tractable AI use in wargaming depends on appropriate platforms, structured state representations, and repeatable computational environments rather than one-off human play alone (Goodman, Risi, and Lucas, 2020). Hotchkiss and colleagues show a similar logic in analytical wargaming via design-of-experiments, demonstrating that systematic exploration of a noisy agent-based simulation can reveal key drivers and nonlinear interactions more effectively than limited one-factor-at-a-time approaches (Hotchkiss et al., 2026).

The broader problem is that many wargames generate important but fragile results. A certain game may reveal a vulnerability, but how much confidence should one place in it? Was it a contingent product of one particular group, one particular scenario, one adjudicator, or one rule configuration? Computational methods promise larger and more systematic exploration of possibility spaces. Again, however, that promise depends on philosophical discipline. Repeated variation is only valuable if the varied parameters are meaningful and if users understand what the resulting distributions do and do not show.

Realism as Fitness to Epistemic Purpose

A final open question concerns realism. Wargame practitioners often speak of realism, fidelity, plausibility, or immersion, but these terms can obscure more than they clarify. Rubel argues that fidelity should be tied to the purpose of the game, not maximized indiscriminately (Rubel, 2006). Hulterström likewise argues that educational value does not depend on literal truth or perfect reproduction, but on whether the game makes relevant affordances and problems perceptible in ways that can cultivate battlefield acumen (Hulterström, 2024). This points toward a more subtle understanding of realism: not resemblance in every detail, but fitness to epistemic purpose.

Computational methods complicate this further. An ABM may be highly explicit but still unrealistic in a deeper sense. A wargame may be low-fidelity in surface detail but highly realistic in the sense that it reproduces meaningful strategic dilemmas. The integration of wargaming and ABM therefore depends on moving beyond the notion that computationalization automatically increases realism. Sometimes formalization increases clarity at the cost of richness while sometimes it enables a better representation of distributed dynamics. The open question is what forms of realism matter for which epistemic purposes.

THE PHILOSOPHY OF AGENT-BASED MODELLING AND RELATED COMPUTATIONAL METHODS

The Ontology of ABM

Agent-based modelling is often introduced pragmatically: agents are autonomous entities that follow rules, interact with one another and with environments, and generate larger patterns. This is correct, but philosophically incomplete. To understand how ABM may complement wargaming, one must ask what sort of model an ABM is.

The most important feature of ABM is that it is a bottom-up, generative, rule-based representation of systems composed of heterogeneous interacting entities. Rather than modelling a system through aggregate equations alone, ABM specifies the properties, goals, states, and interaction rules of individual agents and then examines what emerges from their behavior. In the military and social-scientific literature summarized here, this is repeatedly connected to complexity, adaptation, emergence, nonlinearity, and distributed interaction (Bose and Gadeock, 2024; Chaudhary, 2015; James et al., 2017).

This choice matters because it assumes that conflict outcomes emerge from interaction among many actors, rather than from aggregate variables alone. It implies a distinctive ontology of the target domain. ABM is especially attractive when the world is understood not as an equilibrium-seeking aggregate system but as a complex adaptive environment in which macro-level outcomes depend on local interactions, heterogeneous behaviors, and feedback loops. Bose and Gadeock argue explicitly that warfare should be understood as a complex adaptive system, and that analogue or manual wargames can themselves be viewed as qualitative agent-based models because they represent interacting agents, discrete time, behavioural variation, and emergent outcomes (Bose and Gadeock, 2024). Van Lieburg and colleagues make a related point when they treat socio-cultural environments as self-organizing systems whose behavior can be explored with simplified agent-based models (van Lieburg, Petiet, and Le Grand, 2007).

To say this philosophically: ABMs are not just formalized stories about agents. They are generative ontologies of process. They do not merely describe outcomes but they specify how outcomes can be produced from the interaction of parts. That makes them appealing wherever distributed adaptation and emergence matter.

What Kind of Knowledge do ABMs Produce?

ABMs are most convincing as instruments of generative explanation and exploratory analysis rather than strong prediction.

The key epistemic idea is simple: if one can specify agents, interactions, environments, and rules such that a pattern of interest emerges, then one has shown one plausible way that such a pattern could be generated. This does not prove that the real world works exactly that way. But it does provide a mechanism-rich explanation or at least a candidate mechanism. In this sense, ABM is often strongest not as prophecy but as structured "how-possibly" analysis.

The studies cited here mostly support that reading. Hotchkiss and colleagues use design-of-experiments techniques on a noisy nonviolent-resistance agent-based model to identify which factors and combinations matter most for successful outcomes. Their contribution is not prediction of any specific campaign, but structured identification of major drivers and nonlinear relations in a complex model (Hotchkiss et al., 2026). James and colleagues present ABM as a way to improve rigor in tactical wargaming, especially for the Human Domain, by offering a repeatable and structured way to represent human interaction beyond intuition alone (James et al., 2017). Goodman and coauthors, from the AI side, stress that wargame-related computational approaches need to be designed around the structural features of the problem, not copied blindly from benchmark game successes (Goodman, Risi, and Lucas, 2020). Davis and Bracken make a parallel point in arguing that AI for wargaming should be embedded in an iterative modelling-gaming-analysis ecosystem rather than treated as an isolated technical solution (Davis and Bracken, 2025).

Thus ABM tends to produce at least four kinds of epistemic value.

First, it produces *generative explanations*. It can show how a macro-pattern might arise from micro-level interaction. Second, it produces *exploratory knowledge* about sensitivities and thresholds. Because models can be rerun across many conditions, they can reveal when outcomes change sharply, which assumptions matter most, and where local interactions produce disproportionate effects. Third, it produces *structured possibility-space knowledge*. ABMs can help examine a broader set of plausible trajectories than a single human-played game can. Fourth, it can produce *decision-supporting background knowledge* by representing environmental or social dynamics that shape but do not coincide with elite human choice.

The critical limitation is that ABM usually does not justify strong prediction on its own. This is especially true in strategic, social, or conflict environments where data are sparse, actors adapt, and future contexts differ significantly from calibration cases. The most responsible account of ABM, therefore, is that it yields formalized, exploratory, and mechanism-rich insight rather than exact forecasts.

The Philosophical Tensions with ABM

ABM is relevant to wargaming not only because of what it clarifies, but also because of the new epistemic risks it introduces.

The first tension is *abstraction versus realism*. All models simplify, but ABMs do so in characteristic ways. They must define agent types, internal states, interaction rules, and environments. These choices are always theory-laden. They may isolate important mechanisms, but they may also suppress aspects of social reality that matter greatly. A model that captures movement, protest thresholds, and local influence may miss ideology, institutional memory, fear, or collective meaning. The philosophical question is therefore not whether ABM is realistic in some absolute sense but whether its abstractions are appropriate to the problem and whether the mechanisms omitted are tolerable given the purpose.

The second tension is *generative sufficiency versus empirical adequacy*. A model may successfully generate a pattern of interest, but this does not prove that the model's mechanisms are the real mechanisms. Multiple very different models can generate similar patterns. This is a classic issue in generative modelling. It means that ABM is often better at showing that something could happen by a certain process than that it did or will happen that way.

The third tension is *explanation versus prediction*. Analysts and institutions often want models that predict, but many ABMs are epistemically strongest when used to explore, diagnose, and understand. The pressure to turn them into forecasting tools can tempt users to overclaim.

The fourth tension is *formal transparency versus hidden assumptions*. ABM can seem transparent because rules are explicit. Yet formal explicitness can conceal deeper assumptions in ontology choice, parameter setting, calibration logic, aggregation decisions, or boundary conditions. A rule-based model can be easier to inspect than an opaque seminar-game judgment, but this does not make it assumption-neutral.

The fifth tension is *tractability versus richness*. To model a conflict environment computationally, one must render it computable. This often requires reducing open-ended social meaning into states, variables, and rules. Such reduction is not always inappropriate. But it can distort social phenomena precisely when it becomes easiest to compute them.

The sixth tension is *policy relevance versus methodological modesty*. Policy audiences understandably want actionable outputs. Yet ABMs are often most honest when they say: here are plausible mechanisms, sensitivities, and conditions, not a direct prescription. Bridging that gap requires interpretive work, not just computational output.

Taken together, these problems mean that ABM does not simply fix what is weak in wargaming. It changes the terms of the problem. It changes the form of those weaknesses. Where wargaming struggles with informal adjudication, ABM may struggle with overformalization. Where wargaming struggles with repeatability, ABM may struggle with validation of repeated but stylized worlds. The point of hybridization is not to eliminate these epistemic risks, but to redistribute and manage it more intelligently.

Even with these tensions, ABM is especially relevant to certain classes of wargaming problem. These are cases where distributed interaction, heterogeneous behavior, local adaptation, and nonlinearity matter so much that purely elite-focused, manually adjudicated, or one-shot games are likely to underrepresent the problem. ABM fits wargaming especially well in cases involving civilian and human-domain dynamics, information and perception dynamics, infrastructure and cascading effects, large possibility spaces, and hybrid decision-support architectures.

The philosophical lesson is that ABM is most useful where wargaming's strengths and weaknesses are asymmetrical. It is not needed merely because a game is "too simple" or because computation is fashionable. It is useful where the game's central epistemic burden includes background interaction effects that humans alone cannot easily track or vary systematically.

TOWARD AN INTEGRATED PHILOSOPHY OF WARGAMING AND ABM

The main claim here is that hybrid wargaming systems are not just wargames with computational add-ons. It makes more sense to treat them as arrangements in which different parts do different jobs. The point of integration is therefore not simply to combine tools, but to specify how epistemic labor is allocated across human judgment, formalized agents, adjudication, and post hoc analysis.

Different Epistemic Questions, Different Strengths

The most useful way to clarify their relationship is to ask what kinds of questions each method is especially good at addressing.

Wargaming is strongest when the central question is: What do strategic or operational actors do when they confront one another under uncertainty, consequence, time pressure, and interpretation? Its comparative advantage lies in eliciting human judgment, deliberation, bargaining, misperception, adaptation, and process insight. It shines when conflict is treated as a field of contested meaning and decision.

ABM is strongest when the central question is: What macro-patterns can emerge when heterogeneous actors interact under specified rules and environmental conditions? Its comparative advantage lies in generative mechanism, distributed dynamics, threshold effects, sensitivity analysis, and repeated exploration. It shines when conflict is treated as an interacting system whose outcomes are not reducible to a few explicit decisions.

A hybrid framework can therefore address a third kind of question: How do strategic decisions and distributed social dynamics co-produce outcomes? This is precisely the space where many contemporary policy and military problems lie. Civil-military operations, stabilization campaigns, information confrontation, cyber spillovers, infrastructure disruption, urban operations, and humanitarian coordination all involve elite decisions unfolding within broader environments populated by many other actors.

This three-part distinction is useful because it prevents simplistic claims. Wargaming is not “subjective” while ABM is “objective.” Both are constructed and assumption-laden. Nor is ABM simply “more rigorous” while gaming is “more realistic.” Their epistemic virtues are different. Wargaming foregrounds situated human interaction while ABM foregrounds generative system dynamics. A hybrid design can combine them, but only if their roles are kept conceptually distinct.

A Hybrid Ontology

An integrated ontology of wargaming and ABM begins from the idea that a hybrid system is neither merely a wargame with software added nor merely an ABM with occasional human players. It is a layered epistemic arrangement in which different components represent different parts of the conflict world, perform different inferential functions, and supply different forms of epistemic leverage.

At one layer, the system remains a wargame. Human players represent decision-makers, staffs, commanders, or negotiators who face meaningful choices and interpretive uncertainty. Their agency is valuable precisely because it is not fully reducible to formal rules. It captures strategic reasoning, institutional habit, bargaining, and improvisation.

At another layer, the system includes computationally represented dynamics: civilian populations, infrastructure effects, local adaptation, rumor spread, logistics disruptions, or AI-generated adversary behavior. These are not “background” in a trivial sense but are part of the environment within which human decision-makers act. However, they are often too distributed or numerous to be represented convincingly by free play alone.

At a third layer, the system includes analytical tooling for repetition, variation, sensitivity analysis, or post-game replay. This layer allows investigation of how robust an observed outcome was, whether it depended on narrow assumptions, and what other trajectories might plausibly have followed from similar starting conditions.

Under this ontology, the hybrid system is not one homogeneous method but a designed epistemic arrangement in which human players provide situated agency, computational agents provide distributed mechanism, analytical tools provide structured exploration, and adjudication links these elements into a single inquiry.

What Integration Should Not Mean

This hybrid view makes it easier to say what integration should not mean.

First, integration should not mean replacing human play wherever possible. Some of the most important knowledge in wargaming comes from human interpretation, role conflict, bounded reasoning, and negotiation. Fully automating these dimensions would often destroy the very phenomena of interest.

Second, integration should not mean treating computational outputs as inherently more objective. Explicit rules and repeatable runs are valuable, but they do not remove assumption-ladenness. Computational systems can create new forms of opacity and misplaced confidence.

Third, integration should not mean collapsing wargaming into simulation. A wargame is not just a simulator with a user interface. It remains distinct because human decision, opposition, consequence, and interpretive agency are central to its logic.

Fourth, integration should not mean maximizing formalization for its own sake. The goal is not to model everything, but to model what matters for the question. Overformalization can produce a tractable but philosophically impoverished environment.

Fifth, integration should not mean equating more runs with stronger knowledge. Large-scale exploration is useful only when the underlying model space is meaningful and when users understand the inferential status of the results.

Principles for Hybrid Design

A more productive synthesis requires positive principles. At least six are central.

- *Purpose first*: Hybrid systems must begin from a clear statement of purpose. Is the aim educational, analytical, exploratory, policy-supporting, concept-testing, or decision-assisting? Different aims justify different kinds of human-computational division of labor.
- *Fit-for-purpose fidelity*: Fidelity should not be maximized indiscriminately. Both games and models should include enough detail to represent the causal or practical relationships relevant to the inquiry, but not so much that the system becomes opaque or brittle.
- *Division of epistemic labor*: Human players should be used where interpretation, strategic adaptation, bargaining, or role-based choice is central. Computational methods should be used where distributed dynamics, repeated variation, or background mechanism need to be represented.
- *Transparency of assumptions*: Rules, agent behaviors, boundary conditions, adjudication logic, and analytical procedures should be made inspectable. Hybrid systems need documentation standards stronger than those often used in either informal gaming or black-box modelling alone.
- *Exploratory modesty*: Outputs should be treated as structured possibilities, sensitivities, and pathways rather than direct forecasts. This is especially important in high-uncertainty policy settings.
- *Iterative integration*: Hybrid systems work best when modelling, gaming, and analysis inform one another iteratively. Games can reveal unmodelled mechanisms or unrealistic agent behavior; models can reveal underexplored pathways or hidden sensitivities; analysis can identify where further play or model refinement is needed. Davis and Bracken’s integrated architecture points in this direction (Davis and Bracken, 2025).

These principles do not constitute a final theory but are a minimal framework for a philosophically defensible hybrid inquiry.

A RESEARCH AGENDA FOR COMBINING WARGAMING, ABM, AND RELATED SIMULATION METHODS

If hybrid wargaming systems are best understood not merely as mixed tools but as epistemic arrangements, then the next step is not simply to build more hybrid platforms. It is to articulate a research agenda for how such arrangements should be designed, interpreted, and institutionalized.

Philosophical Agenda

The first line of future work concerns the philosophy of hybrid inquiry itself.

A central question is: What counts as valid knowledge in a hybrid human-computational wargame? Existing wargaming literature offers concepts such as insight, indicative knowledge, ecological validity, and professional knowledge. ABM literature offers concepts such as generative explanation, emergence, and sensitivity analysis. A major task is to clarify how these vocabularies relate. Does a hybrid system produce stronger causal claims than a traditional wargame? Or does it merely produce a different blend of situated judgment and formalized possibility-space exploration?

A second question is: How should realism be judged in hybrid systems? Wargaming literature increasingly treats realism as fit to purpose rather than literal reproduction. ABM literature often treats realism through the lenses of mechanism, calibration, or pattern generation. A hybrid philosophy must distinguish among several kinds of realism: strategic realism, behavioral realism, mechanism realism, and background-environment realism. These may pull in different directions.

A third question is: How should explanation and understanding be related? Wargames often yield understanding through immersion and choice; ABMs often yield explanation through formalized interaction. A hybrid system may therefore combine epistemic modes that are too often separated in philosophy of science. Future work should explore whether such combinations can produce a richer and more accountable form of practical explanation.

A fourth question is: What is the status of computational agents in wargaming ontology? If a red force is partly controlled by reinforcement learning, by scripted heuristics, or by cognitive AI, what kind of “player” is it? Is it an adjudicative tool, an environmental mechanism, an adversary representation, or something else? This matters for how results are interpreted and how agency is attributed within the game.

A fifth question is: How should normativity and responsibility be understood in hybrid decision-support systems? When a game plus model combination informs policy or military planning, responsibility does not rest in one place. It is distributed across designers, modelers, facilitators, participants, and sponsors. The philosophy of hybrid wargaming should therefore include not only epistemology but also epistemic ethics.

Methodological Agenda

The second line of work concerns method design.

A foundational question is: Where in the wargaming cycle should ABM be inserted? There are at least three possibilities:

1. It may be used before the game for scenario generation, environmental shaping, or identifying critical uncertainties.
2. It may be used during the game to represent civilian behavior, logistics, information flows, or other distributed dynamics that respond to player decisions.

3. It may be used after the game to replay key decisions, explore counterfactuals, and test robustness across assumptions.

These are not equivalent uses. Each implies a different epistemic role for the model.

A second question is: How should deliberative data and model-generated data be combined? Research games increasingly collect transcripts, choices, timing, and observational notes, while models can generate distributions, traces, and sensitivity maps. Hybrid systems create an opportunity to link these forms of evidence, but also a challenge of integration. Future work should examine how causal interpretation changes when one combines human reasoning data with computational outcome spaces.

A third question is: How can adjudication be made more transparent in hybrid systems? If models are used during play, the line between adjudication and environmental simulation becomes blurred. Future work should distinguish several modes of adjudication: human judgment, model-assisted judgment, model-constrained judgment, and autonomous computational adjudication. Each raises different issues of auditability and trust.

A fourth question is: How should robustness be assessed? Hotchkiss and colleagues suggest one path via design-of-experiments on noisy agent-based models (Hotchkiss et al., 2026). Related work is needed to examine how repeated gaming and repeated simulation can be combined meaningfully. For example, one could examine whether a finding persists across varied agent parameters and varied human participant groups, or whether certain findings are robust in models but fragile in deliberation.

A fifth question is: What level of ontology is needed for model design? Ouriques and colleagues show that formal ontology-driven conceptual modelling may help structure wargame design characteristics and elements (Ouriques et al., 2025). A useful future line of work would connect such ontological formalisms to hybrid design, asking how explicit domain ontologies might improve interoperability, transparency, and cumulative learning without turning design into empty bureaucracy.

Applied Agenda

The most immediate opportunities for hybrid wargaming and ABM lie in applied research domains where elite decisions interact with distributed social or infrastructural dynamics. These domains are especially important because they combine elite decision-making with distributed civilian, informational, or infrastructural dynamics that traditional wargames often struggle to represent. The following cases illustrate the point most directly.

The first is *civil-military interaction and civilian response*. Hybrid systems could represent how local populations react to movement restrictions, damage, information campaigns, aid distribution, or troop behavior. This is highly relevant to stabilization operations, urban conflict, humanitarian access, and civil-military coordination.

The second is *information confrontation and rumor dynamics*. Many operations are shaped by narratives, social trust, misinformation, and rapid shifts in perceived legitimacy. ABM is well-suited to model local information diffusion and behavioral thresholds, while wargaming can represent how decision-makers respond to such dynamics in real time.

The third is *critical infrastructure and cyber effects*. Pournelle's emphasis on cyber's whole-of-society consequences points toward hybrid systems in which human players make strategic choices while computational models represent cascading service disruption, perception effects, or unintended spillovers (Pournelle, 2024).

The fourth is *gray-zone and hybrid competition*. Many such problems hinge on thresholds, ambiguity, proxies, perception, and cumulative interaction below formal war. Hybrid systems may be especially useful for exploring these intermediate spaces where distributed dynamics and elite signaling interact.

The fifth is *force design and concept development*. Goodman and colleagues note that AI and structured computational environments may be most tractable in bounded concept-development contexts rather than fully open strategic games (Goodman, Risi, and Lucas, 2020). Hybrid architectures can support such uses without overstating predictive reach.

Institutional Agenda

Finally, any serious research program must address institutional conditions.

Hybrid systems require *interdisciplinary teams*. Wargame designers, military subject-matter experts, computational social scientists, AI engineers, human-factors specialists, and policy analysts each bring necessary but partial expertise. One future challenge is to determine how these teams should work together without allowing one epistemic culture to dominate the others.

A second institutional question concerns *documentation and review standards*. Traditional wargaming has often been underdocumented while computational modelling is sometimes documented technically but not interpretively. Hybrid inquiry needs both. Model assumptions, scenario construction, adjudication procedures, player roles, and analysis choices should be recorded in forms that permit critique and replication where appropriate.

A third question concerns *classification, access, and cumulative learning*. Valeriano and Jensen note that wargaming as social science faces obstacles of peer review, classification, and transparency (Valeriano and Jensen, 2021). Hybrid systems may intensify these obstacles if computational components are proprietary, classified, or institutionally siloed. The field will need ways to build cumulative knowledge despite these constraints.

A fourth question concerns *epistemic governance*. Who has authority to interpret outputs? How are disagreements between human players, facilitators, and model results resolved? What counts as an acceptable black box? What level of explainability is needed before outputs can inform real planning or policy? These are not peripheral questions. They determine whether hybrid systems support judgment or simply relocate authority into obscure technical components.

CONCLUSION

This article has argued that the combination of wargaming and ABM should not be understood merely as a technical improvement in analytic capability. It should be understood as the construction of a hybrid epistemic arrangement. Once human players, computational agents, adjudicative mechanisms, and analytical tooling are combined, the central question is no longer simply whether one method compensates for the weaknesses of another. It is how representation, judgment, mechanism, and inference are distributed across the system as a whole.

The argument developed here points toward hybrid human-computational inquiry. In such systems, wargaming contributes situated strategic agency, role-based interaction, and deliberative process insight, while ABM and related simulation methods contribute formalized mechanisms, distributed background dynamics, and repeatable exploration. Hybrid systems, done well, do not replace human judgment. They support and discipline it. They do not transform wargaming into prediction science. They broaden what can be explored and clarify what is being assumed.

A mature field will therefore not simply ask whether AI, ABM, or simulation can be “used in wargaming.” It will ask a harder question: what kinds of epistemic labor should be assigned to human players, to formal agents, to adjudicators, and to analytical tooling, and on what philosophical grounds? Answering that question is a precondition for treating hybrid systems as intellectually defensible rather than merely technologically impressive.

The future of wargaming is neither purely human nor purely computational. It is hybrid. But whether that hybrid future yields better knowledge will depend less on computational power than on conceptual clarity, methodological modesty, and disciplined design.

REFERENCES

Banks, David E. 2024. “The Methodological Machinery of Wargaming: A Path toward Discovering Wargaming’s Epistemological Foundations.” *International Studies Review* 26 (1): viae002. https://doi.org/10.1093/isr/viae002.

Bartels, Elizabeth M. 2019. “The Science of Wargames: A Discussion of Philosophies of Science for Research Games.” Paper presented at the workshop on War Gaming and Implications for International Relations Research, July.

Bose, Saikat K., and S. K. Gadeock. 2024. *Optimisation of Professional Wargaming with Board and Tabletop Wargames Which Really Are Qualitative Agent-Based Models.* Occasional Paper 1-2024. New Delhi: United Service Institution of India.

Chance, Greg, Callum Pender, and Chris Halliday. 2026. “Using AI in Wargaming Simulation as a Multi-Domain Decision Support Tool.” *STO Review*, March 9, 2026.

Chaudhary, Gaurav. 2015. “Agent-Based Approaches for Behavioural Modelling in Military Simulations.” *International Journal of Computer Technology & Applications* 6 (6): 966–974.

Curry, John. 2020. “Professional Wargaming: A Flawed but Useful Tool.” *Simulation & Gaming* 51 (5): 612–631. https://doi.org/10.1177/1046878120901852.

Davis, Paul K., and Paul Bracken. 2025. “Artificial Intelligence for Wargaming and Modeling.” *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 22 (1): 25–40. https://doi.org/10.1177/15485129211073126.

Enstad, Kjetil. 2022. “Professional Knowledge through Wargames and Exercises.” *Scandinavian Journal of Military Studies* 5 (1): 233–243. https://doi.org/10.31374/sjms.130.

Goodman, John, Sebastian Risi, and Simon Lucas. 2020. "AI and Wargaming." arXiv preprint arXiv:2009.08922. https://arxiv.org/abs/2009.08922.

Hotchkiss, Mindy, Hang Dinh, Brendan Thibault, and Lance Fiondella. 2026. "Design of Experiments in Analytical Wargaming: Exploring Agent-Based Simulation Models with Nonlinear Dynamics." *Quality and Reliability Engineering International* 42 (3): 1023–1039.

Hulterström, Patrik. 2024. *From Play to Power: A Philosophical Inquiry into How Educational Wargaming Can Help Cultivate Battlefield Acumen.* Åbo: Åbo Akademi University Press.

James, Alex, James Pierowicz, Michael Moskal, Thomas Hanratty, David Tuttle, Brian Sensenig, and Brian Hedges. 2017. *Assessing Consequential Scenarios in a Complex Operational Environment Using Agent-Based Simulation.* ARL-TR-7954. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory.

Lin-Greenberg, Erik, Reid B. C. Pauly, and Jacquelyn G. Schneider. 2022. "Wargaming for International Relations Research." *European Journal of International Relations* 28 (1): 83–109. https://doi.org/10.1177/13540661211064090.

Mason, Roger C. 2018. "Wargaming: Its History and Future." *The International Journal of Intelligence, Security, and Public Affairs* 20 (2): 77–101. https://doi.org/10.1080/23800992.2018.1484238.

Ouriques, Leandro, Carlos Eduardo Barbosa, Joshua Kritz, and Geraldo Xexéo. 2025. "Towards an Ontology of Wargame Design." *IEEE Access* 13: 78928–78958. https://doi.org/10.1109/ACCESS.2025.3566249.

Perla, Peter P. 1985. "What Wargaming Is and Is Not." *Naval War College Review* 38 (5): 70–78.

Perla, Peter P., and Ed McGrady. 2011. "Why Wargaming Works." *Naval War College Review* 64 (3): 111–130.

Pournelle, Phillip. 2024. "The Need for Cooperation between Wargaming and Modeling & Simulation for Examining Cyber, Space, Electronic Warfare, and Other Topics." *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 21 (4): 359–362. https://doi.org/10.1177/15485129221118100.

Rubel, Robert C. 2006. "The Epistemology of War Gaming." *Naval War College Review* 59 (2): 108–128.

Valeriano, Brandon G., and Benjamin Jensen. 2021. "Wargaming for Social Science." *SSRN Electronic Journal*, July 16, 2021. https://doi.org/10.2139/ssrn.3888395.

van Lieburg, A., P. J. Petiet, and N. P. Le Grand. 2007. *Improving Wargames Using Complex System Practices.* Technical paper. The Hague: TNO Defence and Security.