

# Assessment Strategies for Educational Wargames

Kate Kuehn

---

**Abstract:** Purposeful integration of assessment within educational wargame design is increasingly essential as military education expands those activities within its curriculum. This multimethod case study examines key challenges and strategies for assessment within educational wargaming practice. Drawing insights from faculty interviews, academic documents, and faculty meeting observations, the study identifies six key assessment challenges: gamesmanship, lack of control, multiple faculty roles, receptiveness to feedback, evaluation of individuals in teams, and fairness of evaluation. It then discusses how experienced faculty mitigate these challenges throughout the assessment design process from identifying outcomes to ensuring the quality of evaluation.

**Keywords:** wargaming, assessment, professional military education, PME, authentic learning, case study research

The chairman of the Joint Chiefs of Staff has directed an increase in authentic assessment in Joint professional military education (JPME) curriculum, with particular emphasis on activities like wargaming.<sup>1</sup> At the same time, Service-level leaders, including the Commandant of the Marine

---

Kate Kuehn is the director of Institutional Research, Assessment, and Planning at Marine Corps University. In addition to managing the university's institutional effectiveness process, she supports the evaluation of all professional military education programs and directorates. Kuehn has spent 11 years working on evaluation and assessment of military education programs, providing advice on the design of learning assessments at the classroom, program, and institutional level. She is a member of the Military Education Assessment Advisory Group and has frequently served on military accreditation teams. Her research focuses on assessment and performance evaluation in complex contexts, currently focused on team-based simulated learning environments. Kuehn has an MA from the Georgetown University School of Foreign Service and a BA from the College of William and Mary. She is currently finishing coursework for a PhD in education research methodology at George Mason University.

*Journal of Advanced Military Studies* vol. 12, no. 2

Fall 2021

[www.usmcu.edu/mcupress](http://www.usmcu.edu/mcupress)

<https://doi.org/10.21140/mcu.j.20211202005>

Corps and secretary of the Navy, have directed expansion of wargaming in professional military education curriculum more broadly.<sup>2</sup> Wargames (i.e., artificial competitive environments in which individuals or teams develop and then test the effectiveness of their solutions to complex problems) are often considered authentic learning tools in a PME context.<sup>3</sup> The newest *Officer Professional Military Education Policy* places a heavy emphasis on authentic assessments that simulate real-world applications, with an intentional focus on documenting and evaluating student mastery of key learning outcomes.<sup>4</sup> The purpose of professional military education is to facilitate a student's transition from one career stage to the next by synthesizing their experience with new knowledge and skills. PME institutions balance educational and professional imperatives, seeking to both foster higher-order thinking skills and ensure that those skills transfer into each student's ability to perform certain concrete competencies (i.e., job tasks) after graduation.

There is a long history of wargaming in educational contexts, but Yuna Huh Wong and Garrett Heath specifically highlight a gap in connecting wargaming practice to teaching and learning theory.<sup>5</sup> Their article contributed to an ongoing debate in the Department of Defense (DOD) community about the expansion of wargaming within the military education enterprise, raising questions about the quality of wargaming, organizational and workforce capacity to support the mandated growth, and, important to this article's discussion, understanding wargaming as a learning activity. Meaningful and valid assessment of learning activities aligns theory, task, and evaluation criteria and is essential to high-quality educational practice.<sup>6</sup> To capture and assess learning in wargaming, this gap must be filled. There is also a dearth of published research on the assessment of educational wargaming. Some literature addresses approaches to facilitation within and shortly after game play, but not in the context of accomplishing high-quality learning assessment. Compounding that challenge to assessment design is the complexity of the wargaming environment with its many possible outcomes and data sources.

The author's research seeks to better understand challenges and strategies for assessment within educational wargaming, employing an exploratory case study approach integrating information using multiple methods to develop a rich picture of wargaming assessment practices within the selected context.<sup>7</sup> It draws insights from three major sources: faculty interviews, academic documents, and faculty meeting observations.<sup>8</sup> By examining the perspectives and enacted practices of experienced faculty within wargaming, this study seeks to identify strategies that can serve as useful teaching tools for other faculty as well as contribute to broader theory about designing assessment in such spaces. The article begins with a discussion of considerations for assessment design and implementation within a wargaming context. After outlining the research

question and method, the article then explores assessment challenges in this complex learning context. It concludes with key strategies for mitigating these challenges and implementing effective assessment of wargaming.

## **Assessment Principles in a Wargame Context**

This section provides a framework for designing and dissecting assessment within learning activities, briefly outlining some fundamental principles of assessment design and implementation that are important to consider in any learning context. It then extends that lens to the literature on wargaming and its educational functions within PME.

### **Principles of Assessment Design**

Grant P. Wiggins and Jay McTighe argue that the form and function of learning activities and assessments should be driven by the desired learning outcomes or results (i.e., backward design).<sup>9</sup> In essence, a learning activity should be designed to produce the results and desired evidence of learning that is being sought. Each assessment serves its own function within the curriculum: formative assessments focus on feedback and improving future performance, while summative assessments often produce a score or grade for an academic record and document to what degree students did or did not achieve the desired level of mastery.<sup>10</sup>

Building on the principles of backward design, learning assessment has its own design rules that provide a framework for designing new assessments and for understanding (or improving) how assessment functions within an existing learning activity. Systematic analysis of an assessment design should consider:

1. The desired learning outcome and associated performance expectations;
2. Each activity where that outcome and its associated behaviors are best observed;
3. The tool(s) most appropriate for documenting observations in reference to performance criteria and consistent with the assessment purpose; and
4. The quality, or validity, of each assessment within the activity context.

While the first three are relatively straightforward, the quality of assessment requires elaboration. Assessment quality is first and foremost judged by its content validity. John Gardner defines a high-quality assessment as one that has a clearly defined outcome, which each student has the opportunity to demonstrate.<sup>11</sup> Additionally, he emphasizes that the assessment must have clear criteria or standards for student performance. Finally, the end product of that assess-

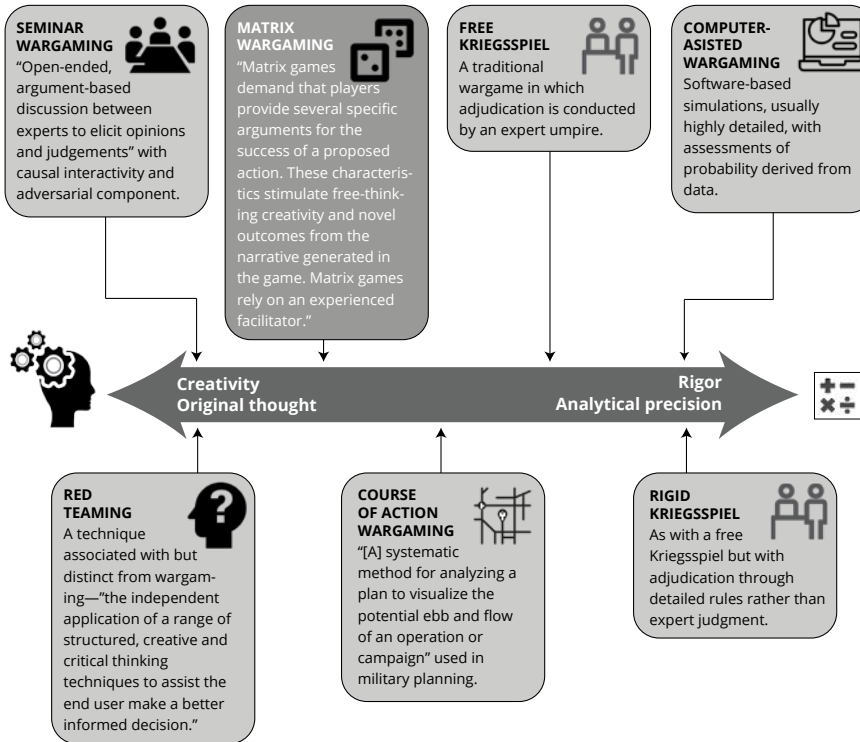
ment must be meaningful and actionable. For a formative assessment, meaningful feedback informs future student improvement and teaching strategies. For summative assessment, actionable feedback informs program evaluation and external understanding of a student's level of mastery. The ability to reproduce the same rating using the same assessment instrument to evaluate the same performance constitutes reliability.<sup>12</sup>

### **Assessment of Wargames**

Resources and guidance on military wargaming emphasize the variety of purposes and values of wargaming that fall within three broad categories: analytical, educational, and experiential.<sup>13</sup> Some characterize wargaming as the basis for testing and refining military concepts and capabilities, some as a tool for structured thinking, and some as a means of experiencing new or hypothetical environments to examine motivations, actions, and consequences.<sup>14</sup> In other words, a wargame can provide a whole host of different functions: examining the likelihood of a plan's success based on the probabilities built into the simulated environment, revealing the gaps and strengths of a planning process and its assumptions, forcing perspective taking to better understand an adversary's thinking, creating a shared experience of cause and effect within a complex environment, etc. A game can, and likely will, provide many of these functions; however, the game and its assessment will be designed differently based on which function has primacy. When conducted for learning, educational wargaming must be aligned to learning outcomes and connected to the broader curriculum.<sup>15</sup>

Literature on wargaming in education falls into two major categories: the first, often published on popular military blogs, captures reflections from faculty who are using wargaming in the classroom.<sup>16</sup> The second focuses on design and implementation of wargames, looking at different game types and structures.<sup>17</sup> Often this second category addresses both educational and analytical wargames. Assessment discussion often has an analytical rather than an educational focus, examining the feasibility of a plan rather than the particulars of learning or student performance in the activity. As an educational tool, these authors often reference serious games and/or game-based learning, but discussions focus more on the concrete activity than on how learning occurs within the environment (i.e., game-based learning).

Wargaming encompasses a wide variety of different game types for which there are many different taxonomies based on level of war, modality of game (e.g., digital), player freedom (rigid to free), level of abstraction or fidelity, type of system (open or closed), and how game outcomes are determined (i.e., adjudication style).<sup>18</sup> These structural features shape how players behave and,

**Figure 1.** The relationship between wargame structure and game outcomes

Source: courtesy of author, adapted by MCUP.

therefore, the types of outcomes that they demonstrate. Focused on the level of creativity expected in the final outcome of a game, Neil Ashdown's adaptation of the United Kingdom (UK) Ministry of Defence *Wargaming and Red Teaming Handbooks* in *Jane's Intelligence Review* makes an explicit connection between different game styles and the type of solution that a game designer is looking for, with the rigid end of the spectrum seen as more precise and objective (figure 1). Game-based learning literature draws similar conclusions about the relationship between structure and learning outcomes, with greater rigidity or directedness in terms of a solution to a problem (or game outcome) better suited to conveying rather than creating knowledge.<sup>19</sup> An organizing principle of selecting a good game for education is that the game should align with the content, level of warfare, and learning outcomes targeted by the activity.<sup>20</sup>

These game features also shape what assessment tools can be integrated into the activity and what data will be available to inform that process.<sup>21</sup> Assessment in a game environment must adapt to the different rules and tools available

in that activity setting. A competitive game has multiple, dynamic sources of data, garnered from game outcomes, within team interactions, between team actions and reactions, faculty facilitation, team decisions, and/or products, etc. Depending on the targeted outcome, assessment can integrate data from any of these sources, preferably while minimizing disruption to the teaching and learning process.<sup>22</sup>

## **Research Question and Method**

Wargaming, with all its varying manifestations, offers a complex task environment with unique challenges for the design and implementation of high-quality assessment. This multi-method case study examines faculty perspectives of assessment in wargaming and associated curriculum, with a particular focus on the challenges and strategies for assessment of team-based (i.e., collaborative) adversarial wargames.

Faculty interviews served as the primary sources for insights into challenges and strategies for conducting assessments of wargames. Interviewees were invited to participate based on their role and experience designing and conducting educational wargames.<sup>23</sup> They had to be directly involved with the design and/or facilitation of an adversarial module for an upcoming educational wargame and had to have taught at the school for at least one full year. Participants included the game leads for two of the departments and two administrators with responsibility for integrating games across the curriculum, representing the diversity of professional backgrounds within the college faculty, including three military Services (U.S. Army, U.S. Air Force, and U.S. Marine Corps) and one PhD civilian faculty member. Each had a diverse background in terms of participating in and conducting games, analytical and educational, in support of DOD requirements. Interviews were transcribed and analyzed for themes regarding assessment challenges and strategies using open coding and then developing axial categories in *MAXQDA*, a qualitative data analysis software.<sup>24</sup>

The analysis also included review of relevant academic documents, including lesson cards (i.e., syllabi), instructor guides, rubrics, and any other game play materials or resources provided to students and faculty for conducting or evaluating each activity. Additionally, the author observed an all-hands “Faculty Wargaming Day,” which set the scene for the program’s approach to wargaming for the academic year. Adapting Zina O’Leary’s document analysis techniques, the author reviewed documents and observation notes to enrich their understanding of the types of games being played, their purpose, and how they were being assessed.<sup>25</sup>

Qualitative studies focus on examining meaning in context, making the importance of gathering the right data (i.e., credibility) and drawing meaning-

ful inferences from that data (i.e., confirmability) essential.<sup>26</sup> As an assessment practitioner at a military education institution as well as a researcher, the author's own professional experiences advising on, and in many cases implementing, assessment activities invariably colored the expectations and interpretations during this research. That said, the focus of this discussion is on the perspectives of the faculty participants, with the analysis trying to make clear the distinction between faculty strategies and the author's own recommended strategies drawn from assessment literature. The use of multiple methods also strengthens the richness and quality of the data collected and can mitigate confirmation bias.<sup>27</sup>

## **Context**

This is a case study of an intermediate-level Service school responding, as many institutions are, to the PME directive to expand wargaming. Following a 10-month program of study, the school confers an accredited master's degree and Joint professional military education-level one (JPME I) credential upon its graduates. At this stage, most students have between 15 and 17 years of military experience and have attained the rank of major, a career transition point from company to field grade officer responsibilities. This school already had a problem-based learning focus that leveraged Socratic-style seminar discussions and learning by doing in the classroom.

Game-based learning activities are also not new; each department has employed gaming in its own way, ranging from individual decision games to informal, scenario-based debates to multiweek planning exercises. Games ranged in length from single seminar discussions (approximately 90 minutes) to multiday activities. Within this school's context, wargames were specifically linked to the development and demonstration of higher-order thinking skills: critical thinking, creative problem solving, decision making, and communication skills. Additionally, games were embedded and scaffolded across the curriculum. In some cases, games themselves were scaffolded, with multiple games in the same format scheduled across the year that address increasingly complex problems. In other cases, games were sequenced within a particular curriculum topic for a seminar, building and connecting on other readings, discussions, and assessments.

This school sought to expand competitive wargaming in particular, where the game must involve thinking players making decisions on both sides of the contested environment. This adversarial wargaming environment is characterized by a dynamic interaction between opposing players/teams in which both sides shape the environment through their actions and reactions. This unstructured wargaming environment falls at the extreme of the spectrum shown in figure 1, presenting unique challenges for observing and evaluating learning outcomes. When played in teams, another frequent and authentic feature of the

program's games, the collaborative aspect of the game adds an additional layer of complexity for both students and faculty assessors.

## **Challenges for Assessment in Wargames**

This analysis identified six key challenges for faculty as they approached assessment in this unscripted learning context. Each of these challenges had particular impacts on the ability to observe, record, and evaluate student performance in a fair and meaningful way. This section provides a review of these challenges.

### **#1: Gamesmanship**

Not unique to wargaming within the game-based learning field, faculty highlighted students' tendency to both game and fight the game rather than focus on the learning process as the first challenge.<sup>28</sup> Students might "game the game" by focusing on how to manipulate the activity rules to maximize points rather than the logic or intellectual reasoning behind a decision. While dissecting a game to detect loopholes does show critical thinking, such lusory focus disrupts both the learning and the faculty members' ability to observe it. All faculty emphasized in the interviews that winning, while the team's desired intent, is not the primary point of the learning activity. In fact, the faculty highlighted that winning can create a faulty assumption that there was one right answer rather than creating a broader realization that decisions are rooted in each context and what is "right" changes. They saw overcoming this challenge as a key faculty responsibility during game facilitation.

Another aspect of gamesmanship occurs when students dismiss game outcomes as erroneous rather than treating them as data to be analyzed or "fight the game." For example, a student team might sustain greater losses than anticipated and attribute it to a flaw in the game. In the extreme, students might dismiss any lessons derived from the gaming experience because they are seen as fixed. This dimension of gamesmanship creates challenges for meaningful feedback and, according to faculty, can also be enhanced by the competitive nature of adversarial wargaming.

### **#2: Lack of Control**

Within the competitive wargaming activity, the two teams playing against one another shape, and are shaped by, the actions of their peer adversaries. Unlike more structured wargames, the games can go in many different directions, bound only by the prespecified objectives for each team and the resources that they have available. As one faculty member observed, there is no reason to recreate history in these particular games; the desire is rather to surprise as well as anticipate your opponent. While this fosters challenge and engagement for students, it means there is less clarity for faculty in terms of achieving the learning



outcome. This puts the onus on the faculty member to continually pull back to the learning outcome without constraining creativity or team dynamics.

With this in mind, faculty must identify outcomes and assessment criteria that are a reasonable expectation of what can be observed across multiple pathways. For example, critical thinking and quality of argumentation can be observed regardless of which course of action a team might select. In a sense, what is assessed is a characteristic of each action or decision rather than the need for a team to select one “correct” answer. At first glance, the clear definition of assessment criteria seems to have more of an impact on summative assessment design, where the desire is to ensure fairness in the final grade and evaluation of each student. But we must also consider the consistency of what will be addressed through formative feedback mechanisms. Formative feedback must provide actionable advice for how students can succeed more effectively in future performances. Such consistency for formative feedback might be established, for example, by establishing a common understanding of how and what is important to emphasize during postgame discussions to ensure all students benefit from emphasis on what school faculty feel is most important.

### **#3: Multiple Faculty Roles**

As an additional complication, during longer wargaming activities, faculty are not present in the room for all team interactions, as they move between each team’s room and the adjudication space as they facilitate the game. In essence, faculty can wear up to three hats as they administer, facilitate, and assess. As administrators, they input, gather, and extract game outcomes and gather documentation of each team’s decisions. As facilitators, they provide feedback and guidance to each team, interpreting game results after each turn and running postgame discussions to recap the game outcomes. Finally, as assessors, they must also observe individual student mastery of learning outcomes to provide advice for future performance and, when summative, a record of student performance. Compounding this fragmentation of responsibility, many faculty (approximately one-third) will be in their first year of teaching and are newly learning the game rules and expectations. As a result, focus may get pulled to making the game work and ensuring students stay engaged, with assessment as a more secondary consideration. This poses a challenge for assessment in terms of faculty ability to observe, document, and then ultimately evaluate outcomes.

### **#4: Receptiveness to Feedback**

Both the complexity and interactive nature of the competitive wargaming activity can lead to problems when the wargaming ends and the faculty must transition into assessment of that activity. Often students must switch from a high-tempo and relatively autonomous stance to a more reflective and facilitat-

ed learning environment. At the same time, they also have to be open to self, peer, and faculty critiques of their performance.

Varying the length of the activity, the faculty member and students may also become overwhelmed by the amount of data available to them. It becomes challenging to prioritize and zoom in on key learning points within this context. Yet, all faculty interviews emphasized that this formal processing session was the most important part for learning through wargames. Several faculty commented that there was never enough time, but that even with unlimited time, students would eventually reach oversaturation with the amount of feedback.

### **#5 Evaluating Individuals within Teams**

Faculty interviews highlighted the complexity of identifying individual performance within a team setting, which was a core design characteristic of the wargames that were examined in this study. Often, faculty formative feedback would focus on how well an individual contributed to their team rather than on mastery of particular knowledge and skills. In a team, an individual's contribution is shaped by the group's dynamics, decision-making structure, communication style, and, if applicable, the decisions of the team's leader. One faculty member even commented that they included team-based aspects in games to reinforce the challenges of collaborative decision making. Ultimately, group dynamics can obscure individual performance that looks for knowledge or skills beyond their contribution to the team. Students may also be constrained by their assigned roles, with potential impact to what and when to contribute. If evaluating higher-order thinking such as decision making at the individual level, one must see the thinking process of each participant or else make a contentious assumption that the final team decision and observed team conversation reflects each individual's thinking skills.

### **#6: Fairness of Evaluation**

The previously mentioned challenges lead to a larger question raised by faculty about ensuring fairness in summative evaluation, as each of those issues mentioned can complicate the ability to observe each student's mastery. Additionally, in early game iterations, students are themselves learning the rules of the game and, depending on the complexity of those games, their performance may reflect more about their ability to quickly understand game rules than their understanding of key concepts. Evaluation can also impact motivation, particularly if students do not feel that they have a fair chance to succeed, potentially impacting reception of feedback. Every faculty member raised the point about evaluation criteria. Partly in the context of not making expectations too specific and granular, but instead focusing on overarching skills. This recommendation

was reflected in the rubrics used by two of the departments. The exercise-based game, for example, was evaluated by a rubric that examines planning (planning process, theory, and doctrine), problem framing (critical and innovative thinking), problem solving, risk management, and leadership (leadership and communication).

Some faculty comments also indicated that interviewees were not entirely satisfied with the evaluation approach, seeing the need to continue evolving what is evaluated and when. In particular, they raised the challenge of ensuring that rubric criteria connect to what is most important during the new adversarial module and adapts as the game changes.

## **Strategies for Assessment of Adversarial Wargames**

As experienced faculty members, those interviewed shared their approaches to overcoming the previously mentioned challenges as they approached design and implementation of assessments. The following sections group their strategies by key components of the assessment process in order to give a more holistic picture of how the strategies contribute to a preliminary framework for wargaming assessment design.

### **#1: Identify the Outcome**

Faculty underscored the importance of selecting the right outcomes, previously noted as integral to the backward design process. Both the interviews and the school's wargaming rubrics focused on processes rather than concrete knowledge or information that one might assess using more traditional assessment like a test. Wargaming is a process, and faculty emphasized the importance of using a complex activity to capture something similarly complex such as application and use. An outcome seeking assessment of more specific knowledge might require greater scripting in the game, in-game documentation, or incorporation of a pre-/post-assessment to reliably ensure students have the opportunity to demonstrate that particular element.

The outcomes selected often connected to program-level learning outcomes indicative of the role these activities played in synthesizing the curriculum content. Not surprisingly, one department used the final wargaming activity as a capstone to their curriculum. Two departments developed a department-specific exercise or wargaming-related rubric to assess all gaming activities across the course of the curriculum, allowing students and faculty to track performance improvement across activities as well as see any areas to target for improvement in other aspects of instruction. At the same time, faculty interviewed expressed dissatisfaction with the rubric either in terms of how well it captured the observable performances or in terms of how readily other faculty could grasp and apply that rubric in context.

## **#2: Observe the Outcome**

Faculty members play a critical role in keeping the outcome in focus during the wargaming activity. Faculty can prime students to look for outcomes during the game introduction and link back to them at key intervals. For example, faculty interviews highlighted the importance of the faculty members' role in adjudication to add meaningful interpretation to game results and even provide informal or formal scenario injects (i.e., scripted game events) that help guide the team for the next turn—a form of in-game formative assessment. The post-game discussion is a critical opportunity to recenter thinking on key outcomes while analyzing game data. Faculty emphasized avoiding the tendency to focus on winning and instead focused on key decision points or events, why they happened, and what the implications were for dealing with future problems. The faculty interviewed were still conceptualizing how to best prepare and develop other faculty to conduct such facilitation and assessment. Within the medical education sector, such preparation is often done through a formal training program, which requires faculty to both observe students in the activity and then participate as a student in a full activity run-through complete with an evaluation from experienced faculty. In some cases, the burden on faculty can be reduced by providing additional personnel that allows division of responsibilities for facilitation, game implementation, and assessment.

A faculty member may be unable to observe all team interactions or to elicit what each individual is thinking during gameplay; however, there are natural built-in opportunities for faculty to incorporate such assessment into the game. More specifically, this includes the decision points where a team issues its instructions, the end of a turn where the faculty member briefs the turn's results, and the postgame discussion. Mid-game opportunities take advantage of the flow of information to examine in-stride thinking without significantly changing the pace of the game. Some games occur in the same room, in which case there is little separation between the team decision and feedback stages, but the same naturally occurring opportunities for assessment exist. Faculty may respond with targeted questioning tied to outcomes or even eliciting individual student input. Faculty can even stagger focus on outcome or individual performances across the turns of the wargaming activity by varying questioning or incorporating different documentation requirements. For example, a team's action sheet might be adapted to capture information about reasoning or risk. The faculty member might also take a strategic pause (a.k.a. operational pause), if needed, to clarify understanding or deepen thinking about a key point.

## **#3: Select the Assessment Tool**

There were two principal types of assessment tools seen in interviews and curriculum documents: rubrics and facilitated dialogue. Assessment was conduct-

ed during and after wargaming activities with both formative and summative purpose. No specific tools or guidance existed for faculty notetaking during the activity or for dialogue-based feedback at the turn or postgame stages; however, faculty interviews associated the approach to good Socratic seminar management. The rubrics used for summative evaluation could be used to inform faculty note-taking but were not formally designed for that purpose. Instead, the school used rubrics to enable summative evaluation and grading. At the same time, rubric criteria were used consistently across activities so that each activity's evaluation was summative but also relevant to performance in the next activity.

There were also tools used within the wargaming activities that could be expanded or adapted to provide more in-stride documentation for assessment. Teams fill out turn sheets and set up internal tracking tools to determine courses of action and track key decision points. These could be adapted to align with less observed outcomes.

#### **#4: Quality of Evaluation**

Faculty emphasized phasing in summative evaluation across the course of the year to allow students and faculty time to adjust to game-based learning. For games with complex rules that will not be repeated, build in an opportunity to learn the rules prior to measuring performance. The practice of scaffolding games or sequencing games as modules within larger curriculum topics also provides an opportunity for multiple assessment points. This opens up the possibility of having some games play an exclusively formative role, which might allow failure and risk-taking while the subsequent assessment examines individual learning from those mistakes. In later iterations, faculty would expect not to see the same mistakes repeated.

While the faculty used rubrics, each emphasized the importance of looking carefully at each rubric's evaluation criteria within the faculty community to ensure clear and continued linkages between professional expectations and performance evaluation standards. Additionally, faculty highlighted the importance of incorporating the rich data produced in the game environment into the assessment design, leveraging the evidence from that joint experience. One faculty interviewee called it "real time feedback to their decision making" when highlighting the advantage of concrete evidence that games can provide to the learning process as students see cause and effect.

#### **Conclusion**

This research, focused on assessment practice and challenges, examined the educational purpose and functions of team-based adversarial wargaming at an intermediate-level PME school. These challenges and strategies were rooted in the context of the complex activities and captured the exploits of experienced

faculty dealing with their design and implementation. The case study method is valuable for capturing experience in context as a particular model for others seeking to address similar challenges within their own contexts. As further research continues this conversation about assessment in complex contexts, additional research would benefit from expanding to include the perspectives of more faculty with different experience levels, wargame designers, as well as students' perspectives.

The faculty members contributing to this research provided important insights into the range of challenges that occur in such complex learning contexts and how those might be mitigated. In particular, faculty highlighted the types of outcomes that are most appropriately assessed by these unstructured spaces and how to maintain focus on them during this kind of activity. Additionally, they highlighted natural inject points for assessment during the wargaming activity, taking advantage of natural seams and feedback intervals within the experience. Finally, they highlighted the need to be mindful of summative assessment and individual performance evaluation within complex group settings with a reminder not to undervalue the formative learning gains accomplished in these spaces.

---

## Endnotes

1. *Chairman of the Joint Chiefs of Staff Instruction (CJCSI)1800.01F, Officer Professional Military Education Policy* (Washington, DC: Joint Chiefs of Staff, 15 May 2020).
2. Gen David H. Berger, *Commandant's Planning Guidance: 38th Commandant of the Marine Corps* (Washington, DC: Headquarters Marine Corps, July 2019); and *Education for Seapower Strategy 2020* (Washington, DC: Department of the Navy, 2020).
3. Definitions of *wargaming* vary across the literature on the topic and are often differentiated by their association with particular disciplines (e.g., modeling and simulation), incorporation of technology, purpose, or specific design features. This article's definition seeks to encompass wargaming practice more broadly, establishing some commonality for the practice of assessment. For more, see Peter P. Perla, *Peter Perla's the Art of Wargaming: A Guide for Professionals and Hobbyists*, ed. John Curry, 2d ed. (Bristol, UK: Lulu.com, 2011); and Yuna Huh Wong et al., *Next-Generation Wargaming for the U.S. Marine Corps: Recommended Courses of Action* (Santa Monica, CA: Rand, 2019), <https://doi.org/10.7249/RR2227>.
4. *CJCSI 1800.01F, Officer Professional Military Education Policy*.
5. Yuna Wong and Garrett Heath, "Is the Department of Defense Making Enough Progress in Wargaming?," *War on the Rocks*, 17 February 2021. For more on this topic, also see Yuna Wong, "Developing an Academic Discipline of Wargaming," King's Warfighting Network, London, 16 January 2019, YouTube video, 1:29:27 min.
6. Noel M. Meyers and Duncan D. Nulty, "How to Use (Five) Curriculum Design Principles to Align Authentic Learning Environments, Assessment, Students' Approaches to Thinking and Learning Outcomes," *Assessment & Evaluation in Higher Education* 34, no. 5 (October 2009): 565–77, <https://doi.org/10.1080/02602930802226502>.
7. Maggi Savin-Baden and Claire Howell Major, *Qualitative Research: The Essential Guide to Theory and Practice* (New York: Routledge, 2012).
8. This study was approved by the George Mason University Institutional Review Board (Reference IRBNet number: 1589661-1). As the study involves research with DOD

- employees in their official capacities, the project was also reviewed by the Human Research Protection Program and Survey Office.
9. Grant Wiggins and Jay McTighe, *Understanding by Design*, 2d ed. (Alexandria, VA: Association for Supervision and Curriculum Development, 2005).
  10. Bruce B. Frey, *Modern Classroom Assessment* (Los Angeles, CA: Sage, 2014).
  11. John Gardner, ed., *Assessment and Learning*, 2d ed. (Los Angeles, CA: Sage, 2012), <http://dx.doi.org/10.4135/9781446250808>.
  12. For observed performances that are rated by a faculty observers, reliability is also often assessed by examining inter-rater reliability. Linda M. Crocker and James Algina, *Introduction to Classical and Modern Test Theory* (Mason, OH: Cengage Learning, 2008).
  13. Pat Harrigan and Matthew G. Kirschenbaum, eds., *Zones of Control: Perspectives on Wargaming* (Cambridge, MA: MIT Press, 2016); and Graham Longley-Brown, *Successful Professional Wargames: A Practitioner's Handbook*, ed. John Curry (Bristol, UK: History of Wargaming Project, 2019).
  14. Longley-Brown, *Successful Professional Wargames*; Ministry of Defence, *Wargaming Handbook* (Wiltshire, UK: Development, Concepts and Doctrine Centre, 2017); and Harrigan and Kirschenbaum, *Zones of Control*.
  15. Wong et al., *Next-Generation Wargaming for the U.S. Marine Corps*.
  16. James Cook, "How We Do Strategy as Performance up at Newport," *War on the Rocks*, 18 March 2019; and James Lacy, "Wargaming in the Classroom: An Odyssey," *War on the Rocks*, 19 April 2016.
  17. Harrigan and Kirschenbaum, *Zones of Control*; and Philip Sabin, *Simulating War: Studying Conflict through Simulation Games* (London: Bloomsbury Academic, 2014).
  18. Harrigan and Kirschenbaum, *Zones of Control*; Longley-Brown, *Successful Professional Wargames*; and Wong et al., *Next-Generation Wargaming for the U.S. Marine Corps*.
  19. Elizabeth Bartels, "Gaming: Learning at Play," *ORMS Today* 41, no. 4 (August 2014), <https://doi.org/10.1287/orms.2014.04.13>; and Dirk Ifenthaler, Deniz Eseryel, and Xun Ge, "Assessment for Game-Based Learning," in *Assessment in Game-Based Learning: Foundations, Innovations, and Perspectives*, ed. Dirk Ifenthaler, Deniz Eseryel, and Xun Ge (New York: Springer, 2012), 1–8, [https://doi.org/10.1007/978-1-4614-3546-4\\_1](https://doi.org/10.1007/978-1-4614-3546-4_1).
  20. Ian T. Brown and Benjamin M. Herbold, "Make It Stick," *Marine Corps Gazette* 105, no. 6 (June 2021): 22–31.
  21. P. G. Schrader and Michael McCreery, "Are All Games the Same?," in *Assessment in Game-Based Learning*, 11–28.
  22. Ifenthaler, Eseryel, and Ge, "Assessment for Game-Based Learning."
  23. Joseph A. Maxwell, *Qualitative Research Design: An Interactive Approach*, 3d ed., Applied Social Research Methods 41 (Los Angeles, CA: Sage, 2013).
  24. Johnny Saldaña, *The Coding Manual for Qualitative Researchers*, 3d ed. (Los Angeles, CA: Sage, 2016).
  25. Zina O'Leary, *The Essential Guide to Doing Your Research Project*, 3d ed. (Washington, DC: Sage, 2017).
  26. Andrew K. Shenton, "Strategies for Ensuring Trustworthiness in Qualitative Research Projects," *Education for Information* 22, no. 2 (2004): 63–75, <https://doi.org/10.3233/EFI-2004-22201>.
  27. Joseph A. Maxwell, "Using Numbers in Qualitative Research," *Qualitative Inquiry* 16, no. 6 (July 2010): 475–82, <https://doi.org/10.1177/1077800410364740>.
  28. Jan L. Plass, Richard E. Mayer, and Bruce D. Homer, eds., *Handbook of Game-Based Learning* (Cambridge, MA: MIT Press, 2020).