

Los juegos de guerra como una disciplina en constante desarrollo. ¿Cómo garantizar que los nuevos formatos están mejorando la práctica?

Natalia Zwarts

<https://orcid.org/0009-0009-8322-0167>

Es expositora en el programa de gestión de seguridad de la Información y Comunicaciones para la Universidad de La Haya. Especialista en diseño de juegos de guerra y formatos experimentales para la educación y la investigación. Entre los cargos previamente asumidos se encuentran la sincronización de tecnologías avanzadas en la OTAN y la cooperación civil-militar. Actualmente desarrolla juegos de guerra educativos destinados a mejorar la ciberseguridad.

Universidad de Ciencias Aplicadas de la Haya

Email: N.H.Zwarts@hhs.nl

24

Resumen: Este artículo presenta las características de los juegos de guerra, que pueden ser usados para apoyar el aprendizaje, análisis y desarrollo dentro de las organizaciones. A pesar de ser una disciplina antigua reconocida por profesionales, es a menudo criticado por un posible sesgo en su diseño, subjetividad de los jugadores y producir principalmente evidencia anecdótica en sus resultados. Para contrarrestar estos problemas, los juegos de guerra pueden ser evaluados a través de una escala de fiabilidad, comparando sus fuentes de datos, número de interacciones, grupo de jugadores y madurez científica. Una vez que la evaluación haya sido completada, es posible reconocer cuan dignas de confianza pueden ser las conclusiones del juego de guerra. Si hay la necesidad de mejorar el diseño existente, la autora propone pasos que pueden ayudar en el incremento de la fiabilidad.

Palabras clave: Juego de guerra, entrenamiento, fiabilidad, diseño, educación.

1. INTRODUCCIÓN

Los juegos de guerra son una herramienta capaz de llevar a cabo tres funciones principales: formación, análisis, y desarrollo (Bourguilleau, Wojtowicz, Lépinard, 2020). La paradoja de los juegos de guerra resulta tanto antigua como innovadora. Matthew Caffrey recopiló la más completa historia de juegos de guerra que representaba sus múltiples facetas. Propuso dividir los juegos de guerra en cuatro generaciones: (1) desde materiales didácticos que se enfocan en habilidades de supervivencia hasta (2) juegos abstractos sobre estrategia; (3) simulaciones de combate para la optimización, y (4) proliferación mundial a través de distintas organizaciones, así como los métodos en evolución (Caffrey, 2019). Si la historia nos señala a cualquier dirección, observamos un catálogo creciente de soluciones de juegos de guerra.

Si se organiza el método para alcanzar la solución según los principios científicos, los juegos de guerra pueden actuar como un laboratorio proxy, cuya finalidad es medir la efectividad de los planes seleccionados (Liu, Ding, Hu y Wang, 2023). Los juegos de guerra tienen la ventaja de recrear un ambiente sintético inmersivo que permite la interacción entre los jugadores, el escenario y los mecanismos del sistema (Wojtowicz, 2020). La presencia de las consecuencias y la habilidad de medir los resultados proporcionan una ventaja analítica en comparación con otros métodos de investigación. (Lin-Greenberg, Pauly y Schneider, 2021).

John Curry denominó a los juegos de guerra como “una herramienta imperfecta, pero útil” (Curry, 2020). El razonamiento que usó como mirada crítica para los resultados de los juegos de guerra estructurados fue: la posibilidad de discrepancias, el modelado incorrecto y la falta de evidencia en la vida real que apoyen los resultados recopilados. Esta crítica inicial nos lleva al tema de este artículo: ¿cómo asegurar que los juegos de guerra sean ciertos? ¿Cómo verificar la precisión de los resultados? ¿Cómo pueden los comandantes confiar en los juegos de guerra como una herramienta de formación?

La urgencia en estos temas se relaciona con la práctica de los juegos de guerra: se puede confiar únicamente si transmite la información, las conclusiones y los objetivos de aprendizajes correctos. Si da continuidad a la ficción, o incluso la falsedad, la práctica se vuelve más peligrosa que útil. Stephen Downes-Martin lo expresó como un sinnúmero de métodos desarrollados, con escasez de pruebas y lógica (Downes-Martin, 2015). En muchos sentidos, la práctica de los juegos de guerra es mucho más prolífica que los juegos de guerra como disciplina científica, especialmente en el ámbito de la educación militar.

La Escuela Naval de Posgrado (NPS, por sus siglas en inglés) en Newport cultiva la tradición de los juegos de guerra con treinta y cinco juegos activos que se utilizan para formar a lo largo de los diferentes formatos y objetivos (Route, 2016).

FIGURA 1
Estudiantes de la NPS durante el curso “Wargaming Applications”, 2021 .



Fuente: <https://nps.edu/-/game-on-nps-wargaming-week-ties-tactics-strategy-to-improve-defense-planning> [se accedió el 29.11.2023]

Ciertamente, pueden considerarse como una experiencia para los estudiantes a quienes se les desafía en pensamiento crítico, planificación y anticipación de las acciones del enemigo. A pesar de la comprobada trayectoria de aprendizaje, pueden plantearse preguntas sobre sus conclusiones. ¿Cómo pueden compararse y considerarse creíbles en sus conclusiones? ¿Cómo garantizar que los estudiantes están desarrollando las habilidades correspondientes? ¿Cómo evaluar de forma sistemática la efectividad de los juegos de guerra? El siguiente artículo ofrece una clasificación de las intervenciones de juegos de guerra al mismo tiempo que presenta formas de validar los resultados del diseño específico.

La evaluación de los juegos de guerra contribuye a la credibilidad de los resultados. Separa la necesidad de justificar el uso del método, en lugar de brindar evidencia de los efectos válidos; además de brindar un camino para comprobar si un juego de guerra cumple con su propósito y qué tan confiable es como herramienta. El alcance de este artículo está delimitado por el uso actual de

los juegos de guerra (2020-2023) y la categorización por tamaño de la muestra, validez y fiabilidad. Para reconocer estos puntos mediante recomendaciones aplicables, el autor presenta la Escala de Métodos Científicos de Maryland con su traducción a la práctica de los juegos de guerra.

2. MARCO TEÓRICO

En 1997, el Congreso encargó a un equipo compuesto por seis docentes de Maryland, Lawrence Sherman, Denise Gottfredson, Doris MacKenzie, John Eck, Peter Reuter y Shawn Bushway, llevar a cabo una minuciosa evaluación de programas de prevención del crimen en los Estados Unidos que llevó a la publicación del siguiente informe de investigación “Preventing Crime: What Works, What Doesn’t, What’s Promising” [Prevención del crimen: qué funcione, qué no, qué resulta prometedor]. El objetivo de la investigación era desarrollar una manera de distinguir entre estudios y trabajo fiables que pueden considerarse como no creíbles. Esta tarea proporcionó una de las herramientas cruciales para debatir sobre el valor de las pruebas presentadas por una investigación determinada.

Se desarrolló originalmente con la finalidad de medir la calidad metodológica de los estudios de prevención del crimen, la escala de métodos científicos de Maryland propone cinco niveles de evaluación, siendo uno el más básico y el menos fiable y cinco el más avanzado y fiable.

- Nivel 1: correlación entre un programa de prevención y una medición del crimen en un punto determinado.
- Nivel 2: mediciones del crimen antes y después del programa sin condiciones de control comparables.
- Nivel 3: mediciones del crimen antes y después del programa en unidades de control comparables y experimentales, controlando otras variables que influyen en el crimen.
- Nivel 4: mediciones del crimen antes y después del programa en múltiples unidades de control y experimentales, controlando otras variables que influyen en el crimen.
- Nivel 5: asignación aleatoria de las condiciones del programa y control a las unidades.

Los autores de la escala tomaron como referencia de excelencia los ensayos controlados aleatorizados (ECA). Los ECA son reconocidos debido a los ensayos clínicos que requieren aislar los efectos de un determinado fármaco de otros

factores. Esto significa que el investigador responsable examina minuciosamente a cada paciente asegurándose de que cumple los requisitos en cuanto a su estado (por ejemplo, no padecer otras enfermedades distintas de la que es objeto del estudio). En los ECA, la población se divide en grupos, algunos de los cuales reciben el tratamiento y otros no. Esto permite a los investigadores separar las causas de la intervención del efecto placebo.

En el caso de los juegos de guerra, la implementación a menudo la implementación tiene prioridad sobre el análisis científico. Cuanto mayor sea la necesidad, menor será la posibilidad de recopilar datos y crear múltiples grupos. Algunos de los conceptos prescritos podrían incluirse sin pérdida de tiempo, por ejemplo, mediante encuestas breves que midan cómo está cambiando el conocimiento de los jugadores.

La división y la medición arrojan un resultado definitivo sobre el efecto del tratamiento, incluidos los síntomas adversos y las mejoras con respecto a las referencias. Los efectos se miden tanto para el cambio positivo como para las reacciones negativas. Se trata de una visión más detallada del cambio de situación basado en los puntos de datos.

La escala permite la clasificación de la recopilación de datos, tamaño de la muestra, aleatorización y diseño experimental para reconocer la visión realista sobre los esfuerzos científicos: siendo siempre la perspectiva actual sobre el problema dado. Esta limitación es también una solución: ver el estudio a través de su posición en la escala de fiabilidad.

3. EL CONTEXTO

En los últimos 10 años, la ciencia pasó por una crisis de confianza debido a que estudios anteriores evidenciaban un bajo índice de replicabilidad. Surgió la cuestión crítica, principalmente, sobre si la evidencia de los experimentos divulgada en la literatura acreditada es correcta. Los esfuerzos por duplicar experimentos famosos llegaron a resultados negativos, refutando gran parte del conocimiento reconocido como estándar en disciplinas como la psicología, la economía e incluso el derecho (Świątkowski y Dompnier, 2017).

Se empezó a proponer herramientas que señalaban a una mejor revisión de los resultados anunciados. El núcleo del problema puede resumirse en “diseños que carecían de suficiente poder y resultados exagerados” (Ferraro y Shukla, 2020). La siguiente sección se centra en las consideraciones específicas de los juegos de guerra sobre el aumento de la precisión de los resultados y la fiabilidad de los estudios.

Argumento central

Los juegos de guerras son susceptibles a la subjetividad debido al alto grado de interacción humana. Se han hecho muchos reclamos sobre la elevada dependencia de los juegos de guerra en múltiples factores: primero, la narrativa utilizada para presentar sus desafíos; segundo, la mecánica que dirige el juego; tercero, los materiales utilizados para construir el juego; y, más importante aún, los jugadores involucrados cada iteración.

Los juegos de guerra deben verse como una doble vía, teniendo en cuenta también la influencia de los facilitadores en los resultados. Al dar un conjunto del mismo escenario, mismos jugadores y diferentes facilitadores, seguramente habrá un cambio en la retroalimentación dada. La forma en que se dan las instrucciones, se explican las reglas y se mueven los objetos influye ciertamente en la experiencia del jugador. Puede verse como una fortaleza de las diferentes experiencias o como un problema debido a la constante variación de los resultados.

A menudo, la percepción de la experiencia del jugador cambia los efectos de un juego de guerra. Por ejemplo, si no se disfruta del escenario o de las reglas, entonces se puede considerar el juego de guerra como una mala herramienta. La clave a la hora de evaluarlos como métodos científicos es estudiar los niveles de objetividad que pueden alcanzarse.

Si intentáramos establecer una escala de fiabilidad para juegos de guerra, puede proponerse una versión simplificada que permita la comparación de diferentes formatos según su credibilidad. Como se menciona anteriormente, no se trata de restar valor a la experimentación que conduce al descubrimiento, sino más bien de establecer una escala que muestra el avance de la fiabilidad en función de la fase de desarrollo.

- Nivel 1: funciona (iteración individual, opiniones de expertos y evidencia anecdótica).
- Nivel 2: probado en múltiples grupos y los resultados pueden analizarse sistemáticamente.
- Nivel 3: existe una medición de los valores de referencia y una medición de los cambios después del juego de guerra.
- Nivel 4: existen grupos de control, que muestran resultados con o sin juegos de guerra.
- Nivel 5: existen múltiples formatos que han sido probados, comparados y analizados. Se eligió la solución más efectiva y la recopilación de datos puede realizarse en un período extendido para informar sobre las mejoras en la vida real.

En el camino hacia la objetividad, no hay que olvidar que el objetivo no es llegar a un entorno estéril. Es importante captar el contexto, las observaciones adicionales y las hipótesis falsificadas. Es el equivalente a descubrir que el sistema de apoyo tiene la misma importancia que los suplementos en los ensayos clínicos. Dentro de la búsqueda de una mayor precisión, los estudios no deben evaluarse, sino situarse en el nivel que representa sus posibles conclusiones. Permite la experimentación en todos los niveles y un desarrollo integral dentro de la disciplina.

Análisis/Debate

Es difícil llevar a cabo una autoevaluación (Vatne, Guttelvik, Hennem y Malerus, 2022) para conocer el nivel de confianza de los resultados, pero es posible reconocer en qué peldaño de la escala se encuentra el juego de guerra. Para ello, deben señalarse los indicadores de las diferencias y el movimiento potencial hacia arriba en la escala. La siguiente sección ilustra un ejemplo de aumento del valor científico de las conclusiones mediante la mejora del diseño del experimento.

¿Cómo aplicar la escala de fiabilidad de los juegos de guerra en un diseño existente?

Los juegos de guerra son una práctica extendida, pero una representación científica escasa. Nuevos estudios proponen formas para mejorar las fallas tradicionales de formatos menos estructurados. Por ejemplo, comparando los diseños existentes y comprobando su precisión para representar la realidad de una batalla específica (Burden, 2023). No existe aún un marco que incremente sistemáticamente los enfoques básicos hacia una mayor fiabilidad. La siguiente tabla muestra los peldaños que se utilizan para llegar al nivel más alto de fiabilidad en los juegos de guerra.

Como ejemplo, una sola edición de un juego de guerra probablemente tenga un efecto en la formación, análisis y desarrollo dentro de una organización. Sin embargo, para ser capaz de determinar que aumenta el nivel de conocimiento sobre tecnología moderna, se debe contar con la comprensión inicial y más

Tabla 1. Peldaños de la escala de fiabilidad.

Nivel	Diseño existente	Peldaños para aumentar nivel
1: Diseño inicial	Funciona (iteración individual, opiniones de expertos y evidencia anecdótica).	<ul style="list-style-type: none"> • Planificar una línea de tiempo para Volver a jugar el juego de guerra. • Comparar opiniones de expertos a partir de la política y el campo. • Evidencia recopilada de distintas fuentes (literatura, estudios de caso, datos estadísticos disponibles).
2: Diseño estructurado	Probado en múltiples grupos y los resultados pueden analizarse sistemáticamente.	<ul style="list-style-type: none"> • Planificar una línea de tiempo para Volver a jugar el juego de guerra. • Comparar opiniones de expertos a partir de la política y el campo. • Evidencia recopilada de distintas fuentes (literatura, estudios de caso, datos estadísticos disponibles).
3: Diseño del efecto medible	Existe una medición de los valores de referencia y una medición de los cambios después del juego de guerra.	<ul style="list-style-type: none"> • Método para establecer valores antes y después del juego de guerra • Incluir mediciones durante el juego de Guerra para reconocer los efectos.
4: Diseño transferible	Existen grupos de control que muestran resultados con o sin juegos de guerra.	<ul style="list-style-type: none"> • Juegos de guerra en paralelo programados con un grupo de control e intervención.

<p>5: Diseño de ventaja comparativa</p>	<p>Existen múltiples formatos probados, comparados y analizados. Se selecciona la solución más efectiva y se puede llevar a cabo la recopilación de datos en un período con el fin de informar las mejoras en la vida real.</p>	<ul style="list-style-type: none"> • Selección de formatos múltiples para realizar un análisis estructurado. • Línea de tiempo que permite la recopilación de datos a largo plazo.
--	---	--

Elaboración propia.

actual para verificar si hay progreso. A este punto, podría mostrarse que el 60% de los participantes si adquieren conocimientos durante el juego de guerra. Comparándolo con una lectura u otra intervención educativa se podría demostrar aún más que el juego de guerra es realmente la mejor solución a nuestro problema de preparación deficiente a una determinada tarea. Podría extenderse a más grupos para conocer si un juego de guerra específico funciona mejor y como este resulta útil para el grupo.

4. CONCLUSIONES

El uso de juegos de guerra como una herramienta conlleva muchos beneficios que van desde realizar preguntas exploratorias, pasando por la recopilación de resultados, hasta llegar a repercutir en la realidad. La introducción de jugadores humanos proporciona una parte significativa de la lógica en el proceso de toma de decisiones, sopesando la información según su relevancia. Existe ciertas limitaciones que deben considerarse para separar el modelo mental de los diseñadores de juegos de guerra y la realidad de la situación (Wojtowicz, 2020b). Como contrapunto, la suspensión de incredulidad y el entorno experimental pueden promover formas innovadoras de pensamiento, que no surgirían sin escenarios hipotéticos. La clave para el desarrollo es conocer qué pueden lograr los juegos de guerra y qué permanece en el nivel de efectos limitados.

Los ciclos de juegos de guerra en periodos de actividad elevada, disminuye de la práctica y el renacer del interés (Bae and Brown, 2021). Trabajar a través de los niveles básicos del diseño de juegos de guerra y utilizarlos en la organización representa un gran paso para establecer un desarrollo a largo plazo. La escala de fiabilidad que ha sido presentada en este artículo busca aumentar la ambición de comparar, ampliar la audiencia de los diseños existentes y aumentar la confianza en los resultados.

REFERENCIAS

- Bae, S.J., & Brown, I.T., USMC (2021). Promise Unfulfilled: A Brief History of Educational Wargaming in the Marine Corps. *Journal of Advanced Military Studies* 12(2), 45-80. <https://www.muse.jhu.edu/article/805918>.
- Bourguilleau, A., Wojtowicz, N., & Lépinard, P. (Déc 2020). Wargames for training future managers. *Management et Datascience*, 5(1). <https://doi.org/10.36863/mds.a.14547>.
- Burden, David J.H. "The Battles of Hue: Understanding Urban Conflicts through Wargaming." *Journal of Strategic Security* 16, no. 3 (2023): 128-140. DOI: <https://doi.org/10.5038/1944-0472.16.3.2138>
Available at: <https://digitalcommons.usf.edu/jss/vol16/iss3/9>
- Caffrey, M. (2019). on Wargaming: How Wargames Have Shaped History and How They May Shape the Future, Naval War College, Newport Papers, 2019.
- Curry, J. (2020). Professional Wargaming: A Flawed but Useful Tool. *Simulation & Gaming*, 51(5), 612-631. <https://doi.org/10.1177/1046878120901852>
- Downes-Martin (2015). Wargaming as a Catalyst for Innovation. *Connections 2015 Conference*, National Defense University.
- Ferraro, P., Shukla, P. (2020) Is a Replicability Crisis on the Horizon for Environmental and Resource Economics? *Review of Environmental Economics and Policy*, 14(2) 2022.
- Lin-Greenberg, E., Pauly, R. B. C., & Schneider, J. G. (2022). Wargaming for International Relations research. *European Journal of International Relations*, 28(1), 83-109. <https://doi.org/10.1177/13540661211064090>
- Liu, H., Ding, Y., Hu, H., Ou, H., Wang, J. (2023). Thinking of War Game Lab Construction. In: Long, S., Dhillon, B.S. (eds) *Man-Machine-Environment System Engineering. MMESE 2023. Lecture Notes in Electrical Engineering*, vol 1069. Springer, Singapore. https://doi.org/10.1007/978-981-99-4882-6_90
- Route, R. (2016). *Wargaming is Alive and Well at Naval Postgraduate School*, The NPS Institutional Archive. Monterey, California.
- Świątkowski, W. and Dompnier, B. (2017). Replicability Crisis in Social Psychology: Looking at the Past to Find New Pathways for the Future. *International Review of Social Psychology*, 30(1), 111-124, DOI: <https://doi.org/10.5334/irsp.66>
- Wojtowicz, N. (2020a) *Wargaming Experiences I: Soldiers, Scientists and Civilians*. J10 Gaming, Delft.
- Wojtowicz, N. (2020b) *Wargaming of the Black Sea Security: Exploratory Study into the Strategies for the Region*. *Ukraine Analytica*, 2020 (19).
- Vatne, D., Guttelvik, M., Hennum, A.C., Malerud, S., *Wargaming for the Purpose of Knowledge Development: Lessons Learned from Studying Allied Courses Of Action*, *Scandinavian Journal of Military Studies*, 2022 (5), 297-308.