SYM-AM-20-064

PROCEEDINGS
OF THE
SEVENTEENTH ANNUAL
ACQUISITION RESEARCH SYMPOSIUM

Acquisition Research:
Creating Synergy for Informed Change

May 13–14, 2020

Published: April 13, 2020

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

The research presented in this report was supported by the Acquisition Research Program of the Graduate School of Defense Management at the Naval Postgraduate School.

To request defense acquisition research, to become a research sponsor, or to print additional copies of reports, please contact any of the staff listed on the Acquisition Research Program website (www.acquisitionresearch.net).

Building a Broader Base for Evidence-Based Acquisition Policy-making

Elizabeth (Ellie) Bartels—is an Associate Policy Researcher and specialist in national security policy analysis gaming. Her work explores a wide range of strategic and operational concerns, with a focus on games on novel topics and those that integrate other analytical techniques. Prior to joining RAND, Bartels was a Senior Associate at Caerus Associates and a Research Analyst at the National Defense University's gaming center. She has a PhD from the Pardee RAND Graduate School, an MS in political science from the Massachusetts Institute of Technology, and a BA in political science from the University of Chicago. [ebartels@rand.org]

Jeffrey A. Drezner—is a Senior Policy Researcher at RAND. He has over 35 years of professional experience conducting policy analysis on a wide range of issues, including planning and program management, analyses of cost and schedule outcomes in complex system development programs, aerospace industrial policy, and defense acquisition policy and reform. His research continues to emphasize mixed qualitative and quantitative approaches to analyze issues associated with technology development, organizational behavior, and program management. Drezner received his PhD in political science from Claremont Graduate University. [zner@rand.org]

Joel B. Predd—is a Senior Engineer at the RAND Corporation and Director of the NSRD Acquisition and Technology Policy Center (ATP). Predd oversees RAND's portfolio of force development, technology, and acquisition research for Office of the Secretary of Defense (OSD), the Joint Staff, the Combatant Commands, and international sponsors. Predd's research addresses topics of force development, acquisition, and technology development, including recent projects on the long-term trends affecting U.S. power projection operations and concept options for multi-domain battle and land-based fires in power projection scenarios. Predd earned a PhD in electrical engineering from Princeton University. [jpredd@rand.org]

Abstract

One of the primary responsibilities of the under secretary of defense for acquisition and sustainment (USD[A&S]) is to ensure the health of the overall Defense Acquisition System (DAS)—as distinct from the health of any particular acquisition program, portfolio, or pathway. USD(A&S) can bolster the health of the DAS by developing and promulgating sound acquisition policy that improves the function and operation of the DAS at the enterprise level. The premise of this paper—and the premise of OUSD(A&S) leadership—is that acquisition policy-making should be data-driven. However, there are limitations of relying on empirical (e.g., historical) data to guide acquisition policy. In light of these limitations, we argue that acquisition policy-making more generally should be evidence-based in recognition of a wider range of analytic tools that can be brought to bear on acquisition policy questions. This report, intended for acquisition professionals, summarizes the case for a broader evidence base and then focuses on one specific tool that we suggest may add analytic value: policy gaming.

Summary

One of the primary responsibilities of the USD(A&S) is to ensure the health of the overall Defense Acquisition System (DAS)—as distinct from the health of any particular acquisition program, portfolio, or pathway. USD(A&S) can bolster the health of the DAS by developing and promulgating sound acquisition policy that improves the function and operation of the DAS at the enterprise level. The premise of this paper—and the premise of OUSD(A&S) leadership—is that acquisition policy-making should be data-driven. However, there are limitations of relying on empirical (e.g., historical) data to guide acquisition policy. In light of these limitations, we argue that acquisition policy-making more generally should be evidence-based in recognition of a wider range of analytic tools that can be brought to bear on acquisition policy questions. This report, intended for acquisition professionals, summarizes the case for a broader evidence base and then focuses on one specific tool that we suggest may add analytic value: policy gaming.

Policy gaming can be used to generate observations about how stakeholders may change their decision-making and behavior in light of changes in policy. Because the strengths and limitations of games differ from those of traditional tools for acquisition analysis, we argue that games complement the existing portfolio of analytic approaches. In this paper, we describe an acquisition policy game that RAND developed to enrich the available evidence base to support acquisition policy-making, summarize insights from a prototype game focused on middle-tier acquisition policy, and speculate about how such games may yield broader insights into how new acquisition policies could affect the DAS.[1]

Key findings include the following:

- Standard approaches to data-driven acquisition policy research may not be sufficient when seeking to anticipate new, substantial policy changes.
- Gaming is a tool used in other areas of defense policy analysis that has promise to help inform acquisition policy creation and implementation.
- A prototype gaming effort conducted by RAND was able to anticipate potential implementation problems for the middle-tier acquisition (MTA) policy, illustrating the potential utility of the approach.

The information that follows supports the key findings provided above.

The Strengths and Limitations of Empirical Data for Setting Acquisition Policy

One approach to setting acquisition policy is to use empirical data to assess the outcomes of the DAS. Under this approach, acquisition policy should be revised periodically on the basis of feedback on how well the DAS is actually performing under current conditions. Conceptually, there is no disputing the value of this approach, and this approach reflects traditional analytic approaches associated with performance evaluation of organizations. USD(A&S) has prioritized this approach, as evidenced by the establishment of the Office of Acquisition Enablers and ongoing work. However, historical performance evaluation faces challenges in providing insight into the impacts of changes in the policy or execution environment.

In practice, an empirical, data-driven approach to acquisition policy-making entails at least four assumptions that cannot always be met. First, to be relevant to evaluating the effect of a new acquisition policy, data on the historical performance of the DAS must be generalizable to the new policy under exploration. This assumption may be valid in many cases, especially for incremental changes to acquisition policy. For example, a change in the dollar-value threshold that triggers increased OSD oversight may be justified by historical data that demonstrates that programs that reach that threshold entail greater risk of bad outcomes. However, the generalizability assumption may break down for non-incremental changes in acquisition policy or technology that, in effect, change the paradigm of defense acquisition and render the historical data less relevant. For example, a new policy to govern the MTA pathway may not benefit from an empirical approach simply because historical program data do not reflect a middle-tier process. For another example, a policy to govern the acquisition of large commercial constellations of smaller satellites in low Earth orbit may not benefit from historical data on space acquisition, which has relied on designing and building small numbers of relatively expensive military satellites. An empirical approach to acquisition policy-making applies only in cases where the data generalize, which would seem to be cases involving relatively incremental changes to acquisition policy.

[1] This research was sponsored by the Office of the Under Secretary of Defense for Acquisition and Sustainment and conducted within the Acquisition and Technology Policy Center of the RAND National Defense Research Institute, a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the Unified Combatant Commands, the Navy, the Marine Corps, the defense agencies, and the defense Intelligence Community. For more information on the RAND Acquisition and Technology Policy Center, see www.rand.org/nsrd/ndri/centers/atp, or contact the director. (Contact information is provided on the webpage.)

Second, and relatedly, current conditions may not hold in the future, placing new pressures on the DAS. For example, potential funding shortfalls in either the near- or far-term can adversely affect program or portfolio performance. The emergence of new technology or new threats create opportunities and challenges that may not have a good analog in historical program data. An analytic approach that focuses only on current conditions risks missing external forces that have been shown to affect the health of the DAS.

Third, an empirical approach to developing and establishing acquisition policy assumes that the data are sufficiently rich to permit one to establish a causal effect of acquisition policy on acquisition outcomes. This discussion could devolve to a technical matter of statistics, but a few general points can be made. There are perhaps on the order of 300 ACAT-1 acquisition programs for which there is good historical data. In some cases, explanatory variables of interest may never have been collected, severely constraining analysis. Examples include technical maturity of specific subsystems critical to performance (affects mostly older historical programs), experience base and tenure of the program management team, and industrial base considerations (structure, number of firms with specific capabilities). These data gaps can sometimes be filled, but such efforts are labor-intensive and tend to take time to accomplish. This contrasts with the requirements of many modern methods of statistics and machine learning where data sets exceeding millions if not billions of items are needed to develop models involving dozens or hundreds of variables.

Even if the effect of acquisition policy were large enough to be measurable with data sets of smaller size, one may struggle to tease out the effect of acquisition policy from the effect of myriad other factors that are changing simultaneously. For example, large complex programs tend to have several program managers, so no single manager is with the program for more than one major milestone. Other senior officials may also have short tenures relative to the life cycle of a major acquisition program. Rapid turnover introduces the possibility that leadership changes, not policy, are driving outcomes. While such details can be traced in careful study of specific cases to determine the relative importance of different potential causes, such an approach further reduces the number of examples, increasing the generalizability problem discussed above. In short, limitations in empirical data reduce our confidence in what can be learned from those data.

There is also a fourth practical assumption—that the cost and risks of making a change to acquisition policy justify an empirical (e.g., historical data) analytic approach. Here too, incremental changes to acquisition policy seem to make this case; it is hard to imagine dire consequences from a change to the dollar-threshold that determines the level of oversight a program receives. However, less incremental changes to acquisition policy may come with unintended consequences that one may prefer to discover prior to implementation. Examples of such changes include the widespread use of other transaction authority, or the creation of new acquisition pathways like middle tier. If DoD adopts a “let’s just try” approach to acquisition policy—experimenting with new policies—those risks may be manifest.

In summary, the empirical approach to setting acquisition policy is important and valuable, but it faces fundamental and practical limits. The empirical approach appears to apply most naturally to relatively incremental changes to acquisition policy for which the historical data generalize and the effects of acquisition policy on acquisition outcomes are large enough to be detected with relatively small data sets and many explanatory variables.

Other Approaches to Setting Acquisition Policy

In addition to using historical empirical data, two other approaches are traditionally used to support acquisition analysis and policy-making, as evidenced by the Section 809 report (2019)—arguably the most comprehensive recent examination of the DAS. These are analyses using commercial data and expert judgment.

There’s no disputing the appeal of benchmarking DoD practices against the commercial

world, but significant differences between DoD and the private sector limit the applicability of those lessons. For instance, the DoD's experience with performance-based logistics, outsourcing, contracting strategy, and managing the vendor base all demonstrate that successful commercial practices often produce different results when applied to a government environment.[2]

Another alternative is expert judgment, which has proven critical to devising nontraditional strategies like the MTA pathway, the creation of the Strategic Capabilities Office, the use of Other Transaction Authority as an alternative to FAR-based contracting, and a host of other innovative approaches to acquisition. However, expert judgment applies best to cases where the experts have sufficient experience to render a judgment. Expert judgment can also be unstructured and untraceable—it often takes the form of heuristics that have grown from years of experience but are difficult to explain to an outsider unless the logic behind the decision is elicited explicitly. But is it possible to leverage expert judgment in ways that can mitigate some of these limitations? Many elements of the DoD are regularly charged with considering the future performance of new policies, strategies, and standard operating procedures, all of which might offer potential methods. One tool for evaluating potential policies that has been a mainstay of other defense communities is gaming.

In summary, while we applaud the current focus on improving the state of data-based policy-making, we are concerned that data is not always available that can answer key questions about the future health of the acquisition system. Put simply, historical data is not always sufficient to understand current and future conditions in which policies must be implemented. At the same time, traditional alternatives to empirical analysis, such as comparison to commercial examples and expert judgment, have long-recognized shortcomings. We suggest that tapping into a wider range of evidence-based analysis, starting with gaming, may complement these approaches.

Gaming and How It Informs Other Areas of Defense Policy

Looking to other areas of defense policy, we find that games can provide useful evidence about proposed policies by providing a sandbox to observe decision-making. In particular, games appear to be valuable in cases where relevant real-world data are not available because the new policy or other condition of interest has not yet occurred. Other conditions that make games particularly appropriate are policies that touch multiple stakeholders with distinct objectives and policies that would elicit revolutionary, rather than evolutionary, changes.[3] This feature suggests that a particularly relevant application of gaming to acquisition is to generate evidence when considering the implementation of a new policy that represents a break from historical practice across DoD. In other words, games are a means of generating evidence-based policies when traditional approaches fall short.

While games can vary in their appearance, ranging from groups of individuals sitting around a table discussing a briefing, to more structured boards and tokens, to computerized systems, games for policy share some common elements. Games consist of human players in a competitive environment making decisions using a defined process and experiencing projected consequences of their choices. In effect, games allow players to practice decisions in a potential future environment, using processes and tools that may not be in place today. As such, they are a way of eliciting expert judgment in a group format and structured by both a specific process and the need to make defined decisions. Former Deputy Secretary of Defense Robert Work and former Vice Chairman of the Joint Chiefs General Paul Selva summed up the value of the method when they stated: "Wargaming is one of the most effective means available to offer senior leaders a glimpse of future conflict, however incomplete. Wargames provide opportunities to test new ideas and explore

[2] See, for example, Drezner, Schmid, Grana, McKernan, and Ashby, 2020; Klitgaard and Light, 2005; Halliday and Adams, 2006; and Section 809 Panel, 2019.

[3] For key texts on policy games and their uses, see Levine, Schelling, and Jones, 1991; Parson, 1996; Perla, 2011; Perla and McGrady, 2011.

the art of the possible" (Work & Silva, 2015).

Looking across the defense enterprise, there are five purposes for which games are frequently used (Bartels, 2020):

- **Understanding a system.** The game allows different stakeholders to walk through the decision-making process in a low risk "sandbox" to better understand how it might work in practice.
- **Exploring decision-making under alternative conditions.** The game allows for structured comparison of decision-making under different contexts.
- **Sparking innovative ideas for new solutions to emerging problems.** Stakeholders might suggest new processes to speed and smooth decision-making across offices, which can then be further analyzed prior to being incorporated into new guidance.
- **Evaluating proposed policies.** While such analysis is unlikely to be directly predictive, it can help compare the strengths and weaknesses of different approaches, identifying potential pitfalls before a policy is implemented.
- **Educating and communicating.** Games can explain how decisions are made about a specific problem by providing players with the experience of making choices and seeing the potential outcomes of their actions.

Each of these purposes may require somewhat different designs to achieve the intended outcomes, and the designs might look quite different from each other in practice, thus requiring custom efforts rather than an off-the-shelf analytic product.

However, regardless of the purpose, all games share a common benefit—they allow researchers and policy-makers to watch how decisions are made under conditions that do not exist today. Furthermore, because groups of players are involved, games provide a forum for multiple experts to debate a concrete problem and make specific decisions in settings where researchers and policy-makers can observe and interrogate the decision-making process to unpack heuristics and weigh competing judgments.

Despite its many benefits, gaming has important limitations, many of which are also shared with other methods. The most fundamental limitation is that games are models, not reality. Thus, what is observed is inherently artificial—analysis of gaming results must always explore how the differences between the real world and the world of the game might skew results. While the traceability of games allows for some additional types of quality control compared to complex computerized models, care is still required in their analysis. In addition, games are not a good approach to test the effects of minor policy changes or estimating point solutions; instead, they are more appropriate for understanding broad implications of large changes. Importantly, the strengths and limitations of games differ from those of traditional tools for acquisition analysis, suggesting that games might complement the existing portfolio of approaches.

Prototyping an Acquisition Policy Game

To explore how games might inform our understanding of acquisition policy, RAND designed and prototyped a structured seminar-style game examining the recent innovation of the MTA pathway. The MTA pathway has several of the key attributes that make gaming an attractive option. The policy is new enough that we do not have much empirical data about its use for different types of programs or its relevance to a range of acquisition stakeholders with different responsibilities in the process, and it represents a sizable shift from current business practices.

The RAND game consisted of a highly simplified simulation of decision-making about how to use the middle tier. We convened a group of players who represented the interests of key staff in the acquisition decision-making process, including project managers, contracting officers, technical

specialists from the A&S staff, and warfighters. Players were asked to review a series of program descriptions using a structured assessment tool and to make a recommendation about whether the program should be acquired and, if so, what acquisition pathway should be used: middle-tier prototype, middle-tier fielding, the traditional 5000 series, or urgent capability. We designed the program descriptions to address operationally relevant capabilities. The programs varied by a number of factors we believed might shape acquisition pathway decisions, including projected cost and schedule, senior leader and warfighter demand, and technology maturity and complexity. After deciding on what pathway to pursue, players also assessed the risk of program failure in terms of performance, cost, and schedule.

Figure 1 shows the game board on which these votes were conducted. For each program, each player placed a chip in one of the top row of boxes to indicate their preferred pathway. For example, in the figure below, the majority of players voted to put the program in the middle-tier prototyping pathway, with two dissenters voting not to acquire the program at all. Players then placed three more poker chips in the bottom of the board to indicate their assessment of the program's risk in performance, schedule, and cost. In other words, we asked players to project where they thought the program might fail looking forward. If players opted to place the program in an MTA pathway, they then completed a worksheet that allowed them to select core elements of an acquisition strategy, including decisions regarding testing regime, contract type, and target cost and schedule. This step was designed to allow players to buy down some of the risks they identified in their initial analysis.

After players had selected which programs to acquire using an MTA pathway and developed an acquisition strategy for each, the game staff then projected the performance of the MTA programs two years into the future to provide feedback on program performance, schedule, and cost. To accomplish this projection, the team used an adjudication process that combined an assessment of the program risk based on program characteristics and the acquisition strategy developed by players to generate a probability distribution of potential outcomes. The researchers then injected an element of chance to reflect aspects of program success that lay outside the control of acquisition officials—including technological progress, contractor performance, and outside economic events—to determine a specific outcome for the program from the distribution. Players were updated on the status of the program and asked if they wanted to maintain the program in the MTA pathway or transition to another option. If the players elected to retain the program in MTA, they could update their choice for an acquisition strategy, triggering a repeat of the adjudication process, which provided an update on program performance at four years. Players were then asked to decide whether they believed the program would successfully meet MTA guidance to field a residual capability in five years or whether the program needed to be transitioned to an alternative pathway.

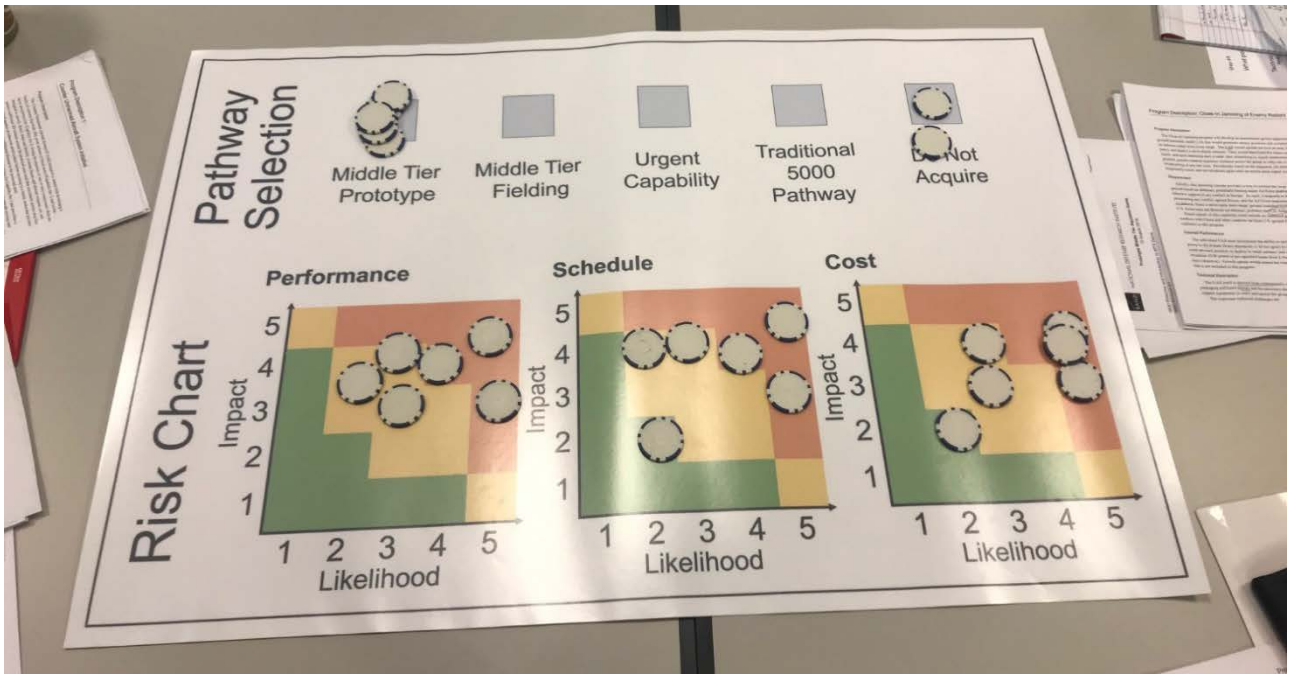

Figure 1. Pathway Selection and Risk Game Board

Nine RAND experts gathered for one day to play our prototype game. We drew the players from among RAND staff with diverse acquisition experience, including former project managers, former service PEOs, and leaders of RAND's acquisition research programs. Over the course of the game, we asked the players to consider seven program descriptions, five of which they opted to acquire using the MTA pathway.

While this game was only a prototype, it suggested several insights into MTA policy design and implementation that should be explored further in a more refined game:

- **The decision to acquire a program through MTA may depend on factors beyond cost and schedule.** The appropriateness for a specific program to use a particular acquisition pathway has often been determined based on cost (ACAT level) or schedule (JUON). The interim MTA policy focused on schedule (deployable capability in less than five years). Large programs projected to take more than five years were sometimes slotted into MTA based on other factors, particularly perceived operational demand, signaling value of the program, and technological maturity. This outcome suggests that current statutory guidance may not consider important factors in program appropriateness for MTA.
- **The risks of transitioning programs between pathways is not well understood.** Participants often opted to use the MTA prototyping program when they believed the program would likely need to be transitioned to an alternative pathway in order to take advantage of the greater flexibility and lower oversight requirements. However, participants noted potential risks, including staff and contractor changes, as well as requirements and budget-related issues, posed a concern with this strategy. The risks of transitioning between pathways was not addressed in any of the MTA policies promulgated at the time the game was played.
- **The perceived flexibility of MTA may translate into novel acquisition strategies.** In several cases, the participants broke large complex programs into more manageable elements that could be executed within the MTA. This included designing incremental acquisition strategies or splitting programs into components or major subsystems to take advantage of more than one pathway to buy down program risk, though it was recognized that this later approach introduces a form of integration risk. The interim policy provided little

guidance on how MTA acquisition strategies should be constructed. It might also be useful to capture and document novel strategies for use by future MTA programs with similar characteristics.

- **MTA may need governance to align policy implementation across the DoD.** Guidance on MTA was interpreted differently by different players. This included data collection and reporting, decision-making and internal milestones, oversight and approvals, and the firmness of MTA constraints (i.e., schedule). Policy alignment should include establishing the primary purpose of the pathway in contrast to other pathways, as well as establishing consistent roles, responsibilities, and authorities for key stakeholders (i.e., DAE/SAE/MDA, PEO, PM, functional staff [contracting, testing, engineering], etc.).

MTA has been in place for nearly two years, so individuals deeply familiar with existing policies and practices may argue that these insights are borne out by experience. However, we emphasize observations were developed with just a day of exercises. This suggests that a gaming platform like this could serve as a sandbox to experiment with new acquisition policies.

Limitations

These findings illustrate the types of evidence that a game can provide. However, it is also important to note that the scope and scale of this prototype effort imposed several limitations on game findings. First, we were not able to examine service-specific implementation of MTA as part of this effort. Because of known difference in service culture and MTA implementation approach, our players' decisions may be different from those of real decision-makers. Second, due to limited time, the description of each program was limited to between two to three pages of text. It may be that inclusion of additional information may have introduced other key factors into the discussion or even shifted player decisions regarding pathway. A third limitation was that requirements to generate unclassified materials on operationally relevant but not currently available capabilities excluded some categories, like COTS programs, from consideration. Such limitations could be mitigated in future iterations by making adjustments to the game's processes and materials.

Next Steps

We offer several next steps for USD(A&S) to consider.

First, look to expand beyond traditional evidence-based approaches to acquisition policy. While empirical data are critical, they have limitations. Adding complementary methods, such as games, will build a more rigorous evidence-based policy than depending on data-driven approaches alone. Thus, we suggest USD(A&S) fund pilot efforts using new tools in order to better understand their utility to evaluating the health of the DAS.

Second, instrument the adaptive acquisition system to enable continuous, incremental improvements to acquisition policy. The value of an empirical, data-driven approach appears indisputable. But this analysis reveals that historical data scarcity and the existence of many explanatory variables both limit the evidence-based approach. In particular, our work suggests that data-based evidence is likely to be informative when considering incremental policy changes but may be insufficient for informing larger changes in policy.

Third, experiment with developing and using policy games to rapidly prototype more revolutionary changes to acquisition policies for which an empirical approach may not apply. Evidence from our prototype effort suggests that problems involving policy implementation, developing acquisition strategies, and understanding information requirements may benefit from gaming. Games developed to inform policy-makers may also have additional utility as an educational tool in DoD schoolhouses.

References

Bartels, E. M. (2020). *Building better games for national security policy analysis* (RGSD-437). Santa Monica, CA: RAND Corporation.

Drezner, J. A., Schmid, J., Grana, J., McKernan, M., & Ashby, M. (2020). *Benchmarking data use and analytics in large, complex private sector organizations: Implications for Department of Defense acquisition*. Santa Monica, CA: RAND Corporation.

Halliday, J., & Adams, J. (2006). *Identifying performance based logistics best practices*.

Klitgaard, R., & Light, P. C. (Eds.). (2005). *High-performance government: Structure, leadership, incentives* (MG-256-PRGS). Santa Monica, CA: RAND Corporation.

Levine, R. A., Schelling, T., & Jones, W. M. (1991). *Crisis games 27 years later: Plus c 'est déjà vu* (P-7719). Santa Monica, CA: RAND Corporation.

Parson, E. (1996). What can you learn from a game? In R. L. Keeney, R. J. Zeckhauser, & J. K. Sebenius (Eds.), *Wise choices: Decisions, games, and negotiations*. Boston, MA: Harvard Business School Press.

Perla, P. P. (2011). *The art of war gaming: A guide for professionals and hobbyists* (2nd ed.). Naval Institute Press.

Perla, P. P., & McGrady, E. D. (2011). Why wargaming works. *Naval War College Review*, *64*(3).

Section 809 Panel. (2019). *Report of the advisory panel on streamlining and codifying acquisition regulations: Volume 3 of 3.* Retrieved from https://discover.dtic.mil/wp-content/uploads/809-Panel-2019/Volume3/Sec809Panel_Vol3-Report_Jan2019_part-1_0509.pdf

Work, R., & Silva, P. (2015). Revitalizing wargaming is necessary to be prepared for future wars. *War on the Rocks.* Retrieved from https://warontherocks.com/2015/12/revitalizing-wargaming-is-necessary-to-be-prepared-for-future-wars/

Acquisition Research Program
Graduate School of Defense Management
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net