IDA | RESEARCH NOTES

WELCH AWARD 2018

Fall 2019

IDA is the Institute for Defense Analyses, a trusted and decisive contributor to the national security and science policy debates. A nonprofit corporation, IDA operates three Federally Funded Research and Development Centers (FFRDCs) that provide objective analyses of national security issues and related national challenges, particularly those requiring extraordinary scientific and technical expertise.

The summaries in this edition of *IDA Research Notes*, as well as the original publications on which the summaries are based, were written by researchers in the following IDA research groups. The directors of these groups would be glad to respond to questions on topics related to their work.

Cost Analysis and Research Division (CARD)
Dr. David E. Hunter, Director, 703.575.4686, dhunter@ida.org

Information Technology and Systems Division (ITSD)
Dr. Margaret E. Myers, Director, 703.578.2782, mmyers@ida.org

Intelligence Analyses Division (IAD)
RAdm. Richard B. Porterfield, Director, U.S. Navy (retired). 703.578.2812, rporterf@ida.org

Joint Advanced Warfighting Division (JAWD)
Dr. Daniel Y. Chiu, Director, 703.845.2439, dchiu@ida.org

Operational Evaluation Division (OED)
Dr. Robert R. Soule, Director, 703.845.2482, rsoule@ida.org

Science and Technology Division (STD)
Dr. Leonard J. Buckley, 703.578.2800, lbuckley@ida.org

Science and Technology Policy Insitute (STPI)
Dr. Mark J. Lewis, Director, 202.419.5491, mjlewis@ida.org

The Larry D. Welch Award is named in honor of former IDA president and U.S. Air Force (USAF) Chief of Staff, General Larry D. Welch, USAF (retired). The annual award recognizes IDA researchers who exemplify General Welch's high standards of analytic excellence through their external publication in peer-reviewed journals or other professional publications, including books and monographs.

The articles in this issue of *IDA Research Notes* are derived from the winner and finalists in the 2018 Larry D. Welch Award competition. The Welch Award Selection Committee named four additional nominated publications as being worthy of note given their success in the open literature and the quality of research they reflect.

Names in bold type have current or former affiliations with IDA. The original publications that were nominated are cited, along with a link where available.[1]

WINNER

This year the best example of high-quality, relevant research published in the open literature is "Deterrence Is Not a Credible Strategy for Cyberspace," by Information Technology and Systems Division (ITSD) researcher **Michael P. Fischerkeller** and co-author Richard J. Harknett. Their paper was published in *Orbis*, May 18, 2017.

FINALISTS

"An Abridged History of Federal Involvement in Space Weather Forecasting," published in *Space Weather*, October 2017, by former Science and Technology Policy Institute (STPI) researchers **Becaja M. Caldwell** and **Eoin D. McCarron** and STPI researcher **Seth Jonas**, builds on STPI analyses for the Office of Science and Technology Policy.

Based on IDA research for the Military Compensation and Retirement Modernization Commission, Cost Analysis and Research Division (CARD) researchers **Sarah K. Burns** and **Philip M. Lurie** and former CARD researcher **John E. Whitley** wrote "Analysis of an Alternative Military Healthcare Benefit Design," published in *Defence and Peace Economics*, July 2017.

Intelligence Analyses Division (IAD) researchers **Stephanie M. Burchard** and **Dorina A. Bekoe** coauthored "The Contradictions of Pre-election Violence: The Effects of Violence on Voter Turnout in Sub-Saharan Africa," published in *Africa Studies Review*, September 2017. This article is an extension of research related to election violence in Africa conducted in IDA's Africa program.

"Power Approximations for Generalized Linear Models Using the Signal-to-Noise Transformation Method," published in *Quality Engineering*, October

[1] IDA assumes no responsibility for the persistence of URLs for external and third-party internet websites referred to in this publication. Further, IDA does not guarantee the accuracy or appropriateness of these websites' content now or in the future.

2017, by Operational Evaluation Division (OED) researchers **Thomas H. Johnson** and **Colin E. Anderson**, former OED Assistant Director **Laura J. Freeman**, and IDA consultant **James R. Simpson**, is based on IDA research conducted for the Director, Operational Test and Evaluation.

Science and Technology Division (STD) researcher **Shelley M. Cazares** based her article “The Threat Detection System That Cried Wolf,” published in *Defense Acquisition Research Journal*, January 2017, on multiple IDA projects for the Department of Defense and the Department of Homeland Security.

“Winning Indefinite Conflicts: Achieving Strategic Success Against Ideologically-Motivated Violent Non-State Actors,” published in *Small Wars Journal*, March 2017, by Joint Advanced Warfighting Division (JAWD) researcher **Mark E. Vinson**, is informed by IDA research for the Joint Staff.

NOTEWORTHY

“Five Actions to Improve Military Hospital Performance,” published in *IBM Center for the Business of Government*, 2017 Improving Performance Series, by former CARD researcher **John E. Whitley**, is based on multiple IDA analyses for the Department of Defense and the Military Compensation and Retirement Modernization Commission.

“Getting ‘Cyber’ Right for the Department of Defense,” published in *War on the Rocks*, November 2017, by ITSD researchers **Gregory V. Cox** and **Priscilla E. Guthrie**, is based on knowledge gained through multiple IDA projects for the Department of Defense.

“Operational Graphics for Cyberspace,” published in *Joint Forces Quarterly,* second quarter 2017, by OED researchers **Erick D. McCroskey** and **Charles A. Mock**, is based on IDA analyses for the Director, Operational Test and Evaluation.

“Orbital Debris Momentum Transfer in Satellite Shields Following Hypervelocity Impact, and Its Application to Environmental Validation,” published in *Science Direct, Procedia Engineering* 204, 14th Hypervelocity Impact Symposium, April 2017, by OED researcher **Joel E. Williamsen** and IDA consultant **Steven W. Evans**, is based on IDA analyses conducted for NASA.

Deterrence Is Not a Credible Strategy for Cyberspace (and What Is)

Michael P. Fischerkeller and Richard J. Harknett

Much of U.S. defense policy over the past 20 years has been grounded in a deterrence framework. When the cyberspace operational domain emerged, it was promptly and similarly considered a domain of restraint and reaction, with insufficient attention paid to its unique characteristics and the strategic context. This article makes two central arguments. First, within cyberspace, the protection or advancement of national interests cannot rest on deterrence as the central strategy but can be realized through a strategic approach that captures and takes advantage of unique characteristics of the domain and the current strategic context—persistent engagement. Second, if the United States is to shape the development of international cyberspace norms that will bring stability and security, it can do so primarily through strategic cyber campaigns that begin to shape directly and indirectly the parameters of responsible behavior.

[1] The winning publication originally appeared in *Orbis* 61, no. 3 (Summer 2017): 381–393, https://doi.org/10.1016/j.orbis.2017.05.003.

Challenge of a New Domain

In a 2010 essay, William J. Lynn III, then U.S. Deputy Defense Secretary, outlined a new strategy for a new operating domain—cyberspace (Lynn 2010). In describing the strategy, consideration reasonably turned to a strategic framework to suggest norms of behavior for operating within cyberspace. Consistent with much of U.S. defense policy over the past 20 years, those norms were grounded in a deterrence framework. The operational norms associated with the air, land, and maritime domains are fundamentally derived from the centuries-old concept of Westphalian sovereignty, a structural feature rooted in segmentation (bounded territories) and derived from respect for the principle of non-intervention and territorial integrity that marked the end of the Thirty Years' War in 1648. Although specifics regarding these norms have evolved, the basic principle is still widely accepted by state actors in the international system and is codified in the United Nations Charter article 2(4), which states, "All members shall refrain in their international relations from the use of force against the territorial integrity or political independence of any state." Consistent with this language, the United States and its allies adopted and advocated for the principle of relative operational restraint associated with deterrence strategies (i.e., a "doctrine of restraint" came to anchor U.S. cyberspace strategy and inform perspectives on the substance of norms). Unfortunately, this perspective was adopted without comprehensive consideration of whether a strategy of deterrence was appropriate given cyberspace's unique characteristics and the current strategic context. It was not—as many actors realized their national interests could be advanced through strategic cyber campaigns comprised of continuous operations with strategic effects short of use of force or armed attack equivalence. While many of these actors might be considered "unlike-minded,"[2] the number and effectiveness of their aggressive cyber campaigns suggest that a sizeable number of effective actors are leveraging the U.S. default to restraint.

Uniqueness of Cyberspace

The cyberspace operational domain is defined as "a global domain within the information environment consisting of the interdependent networks of information technology infrastructures and resident data, including the Internet, telecommunications networks, computer systems, and embedded processors and controllers" (Joint Chiefs of Staff 2018, GL-4). Thus, it is argued that cyberspace is uniquely a human-constructed domain, and thus malleable. Moreover, the scale and scope of this constantly shifting space is distinctive—state and non-state actors' abilities to modify other operational domains cannot occur at the pace and on the scale being witnessed in cyberspace. Strategy must recognize that there is a qualitative difference between the capacity to modify terrain and to create it whole cloth.

[2] See White House (2018, 21) for the strategy's specification of working with "like-minded" states to develop norms.

The uniqueness of cyberspace is also reflected in the low cost of entry, which allows a number of actors who can affect relative national power to operate in cyberspace that is orders of magnitude higher than the small number of states that operate with consequence in the land, air, maritime, and space operational domains. Moreover, no internationally agreed upon concept of cyberspace sovereignty prevails. This suggests a corollary—international relations (and nature) abhor vacuums; consequently, cyber security strategy should assume that states and other significant actors are continually seeking to exert their influence in cyberspace through strategic cyber campaigns or operations.

Agreement to compete robustly short of armed conflict may be the grand strategic consequence of cyberspace.

Whereas segmentation is the core structural feature of the air, land, and maritime domains, interconnectedness is the oft-cited, but rarely embraced, core structural feature of cyberspace. If one accepts interconnectedness as such, then fundamental international relations concepts for understanding or explaining actor behaviors and making strategic choices, such as sovereignty and territoriality, come into question because the core condition that follows from interconnectedness is constant contact, a term used by the United States Cyber Command (USCYBERCOM) to describe the cyberspace operating environment (USCYBERCOM 2018, 4).[3] This condition, when coupled with the nature and substance of cyberspace—a vulnerable yet resilient technological system that is a global warehouse of and gateway to troves of sensitive strategic information—encourages persistent opportunism to access and leverage those sensitive data while simultaneously requiring states to continuously seek to secure those data and data flows from others.[4] The combination of interconnectedness and constant contact with cyberspace's ever-changing character, both in "terrain" and in the capacity to maneuver across that terrain, further encourages operational persistence in order to secure and leverage critical data and data flows. When these factors are considered together, in operational reality, operational persistence/engagement (not operational restraint) becomes the appropriate strategic choice (if not imperative) for states seeking to secure and advance their interests in, through, and from cyberspace.[5] The past decade of voluminous and exploitative adversarial behavior in cyberspace suggests adversaries recognized and adapted to this imperative early in cyberspace's maturation. The consequence for the United States has been the gradual degradation of U.S. sources of national power by adversarial strategic cyber

[3] See also Fischerkeller and Harknett (2017).

[4] For a discussion of the nature, character, and substance of cyberspace and its implications for cyberspace strategy, see Fischerkeller (2018).

[5] This was the critical and concluding argument of the 2018 Welch Award–winning publication (Fischerkeller and Harknett 2017). The remainder of this article highlights extensions and applications of that argument as represented in the authors' publication (Fischerkeller and Harknett 2018).

campaigns targeting those same sources of power. This situation has not gone unnoticed by U.S. policy makers.

A strategic approach to securing national interests and pursuing norms codification in cyberspace that is based primarily on operational restraint, then, fails to take into account that the unique characteristics of cyberspace argue for a strategic approach of operational persistence. Analyses of behaviors in, through, and from cyberspace over the past decade reveal that state and non-state actors have increasingly understood and aggressively leveraged the value of cyberspace and strategic cyber campaigns short of armed conflict to support their interests. It is likely that these actors have also come to recognize that because norms emerge first through behaviors, then mature and are codified through international discourse, when the time comes for international discourse regarding codification, those who operationally dominate the domain will be in the strongest position to argue for norms supporting their positions.

Current Strategic Context

National Security Strategy of the United States of America, issued in December 2017, and its complement, *National Defense Strategy of the United States of America*, stand in marked contrast to their predecessors in their declarations that adversaries are executing strategic campaigns short of armed attack to secure and advance national interests. Indeed, both documents assert that the central challenge to U.S. security and prosperity is the re-emergence of a long-term, *strategic competition* with revisionist and rogue regimes and actors that have become skilled at operating *below the threshold of armed conflict* (White House 2017, 3, 31; Department of Defense 2018, 2). Cyberspace and its derivative cyber operations, in particular, have been identified as offering state and non-state adversaries the ability to wage strategic campaigns against American political, economic, and security interests without physically crossing U.S. borders (White House 2017, 12). This view is presented most comprehensively in *Achieve and Maintain Cyberspace Superiority: Command Vision for U.S. Cyber Command*. Adversaries are described as continuously operating against the United States below the threshold of armed conflict—demonstrating the resolve, technical capability, and persistence to undertake strategic cyberspace campaigns to weaken U.S. democratic institutions and gain economic, diplomatic, and military advantages (USCYBERCOM 2018, 3).[6]

Strategic Approach of Persistent Engagement

Taking into consideration the unique characteristics of cyberspace and the current strategic context, USCYBERCOM recently described a strategic approach that is better aligned than deterrence with these realities. The approach

[6] Concern has been expressed regarding "the *persistence* [emphasis added] exhibited by adversary attempts to penetrate critical infrastructure and the systems that control these services" Rogers (2017, 2).

prescribes that the United States increase resiliency; defend forward as close as possible to the origin of adversary activity; and contest cyberspace actors to generate continuous tactical, operational, and strategic advantage.[7] USCYBERCOM argues that this strategic approach of *persistent engagement*—described operationally as the combination of seamless resiliency, forward defending, contesting, and countering—will compel many U.S adversaries to shift resources to defense and reduce attacks. Moreover, *persistent engagement* is expected to allow greater freedom of maneuver to impose tactical friction and strategic costs on U.S. adversaries pursuing activities that are more dangerous before they impair U.S. national power. This effort seeks to render the majority of adversary cyber and cyber-enabled activity inconsequential.

We have recently argued that through the adoption of this strategic approach, the United States would become an active participant in an ongoing *agreed competition* below the threshold of armed attack among major actors in cyberspace, all of whom are seeking to protect and/or gain strategic advantage short of armed attack through the same (Fischerkeller and Harknett 2018). The term *agreed competition* is a derivative of *agreed battle*, a term strategist Herman Kahn described as a concept rooted in factors relating to particular levels of escalation.[8] The concept emphasizes that in an escalation situation in which both sides are accepting limitations, there is in effect an agreement, whether or not it is explicit or even well understood. "Thus the term does not have any connotation of a completely shared understanding, an intention of containing indefinitely with the limitation, or even a conscious quid pro quo arrangement" (Kahn 2017, 3). From a norms-development perspective, what is important to note in Kahn's rendering is that agreement rests on *interactions* between adversaries, which, despite being complex and nuanced, can come to be understood and shared between actors. He notes that states can come to recognize "what the 'agreed battle' is and is not, what the legitimate and illegitimate moves are, and what are 'within the rules' and what are escalatory moves" (Kahn 2017, xiii).[9]

And so, to come full circle, in contrast to a strategy of deterrence, which emphasizes cyberspace operational restraint and norms establishment with like-minded significant actors, a strategic approach of *persistent engagement* emphasizes competitive interaction within an *agreed competition* and norms

[7] USCYBERCOM argues that superiority through persistence seizes and maintains the initiative in cyberspace by continuously engaging and contesting adversaries and causing them uncertainty wherever they maneuver. It describes how USCYBERCOM would operate (maneuvering seamlessly between defense and offense across the interconnected battlespace_; where they would operate (globally, as close as possible to adversaries and their operations_; when they would operate (continuously, shaping the battlespace); and why they operate (to create operational advantage for the United States while denying the same to U.S. adversaries) (USCYBERCOM 2017, 5).

[8] Kahn attributes the term *agreed battle* to Max Singer.

[9] For a comprehensive discussion of interaction and escalation dynamics that would emerge from a strategic approach of persistent engagement, see Fischerkeller and Harknett (2018).

construction (through interaction) with all actors. Security and stability will emerge through interaction because more clarity will emerge on the demarcations between illegitimate and legitimate cyber operations and between operations outside and within the "rules" of *agreed competition*.

Conclusion

Several years ago, U.S. adversaries waded cautiously but strategically into the strategic competitive space between war and peace, perhaps most fulsomely in cyberspace. In response, the United States adopted a strategy of deterrence, one that was misaligned with both cyberspace's unique structural and operational characteristics and the strategic context. Consequently, adversaries are now pursuing aggressive strategic campaigns short of armed conflict in, through, and from cyberspace to gain strategic advantage in military, economic, and diplomatic arenas. As evidenced in recent U.S. strategic guidance, however, the United States has now recognized that it must operate persistently in this competitive space if it hopes to re-gain the upper hand on adversaries who have been reaping the benefits of their early strategic adaptation to cyberspace at the expense of U.S. national interests. A strategic approach of persistent engagement in cyberspace supports this newly adopted orientation while simultaneously, through continuous competitive interaction, supporting the development of norms of responsible behavior. Agreement to compete robustly short of armed conflict may be the grand strategic consequence of cyberspace.

References

Department of Defense. 2018. *Summary of the 2018 National Defense Strategy of the United States of America: Sharpening the American Military's Competitive Edge*. https://www.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf.

Fischerkeller, M. P. 2018. *Offense-Defense Theory, Cyberspace, and the Irrelevance of Advantage*. Alexandria, VA: Institute for Defense Analyses. IDA Document NS D-8939.

Fischerkeller, M. P., and R. J. Harknett. 2017. "Deterrence Is Not a Credible Strategy for Cyberspace." *Orbis* 61, no. 3: 381–393. https://doi.org/10.1016/j.orbis.2017.05.003.

———. 2018. *Persistent Engagement, Agreed Competition, Cyberspace Interaction Dynamics and Escalation*. Alexandria, VA: Institute for Defense Analyses. IDA Document NS D-9076. https://www.idalink.org/D-9076.

Joint Chiefs of Staff. 2018. *Cyberspace Operations*. Joint Publication 3-12. http://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_12.pdf.

Kahn, H. (with a new introduction by Thomas C. Schelling). 2017. *On Escalation: Metaphors and Scenarios*. London: Routledge.

Lynn III, W. J. 2010. "Defending a New Domain: The Pentagon's Cyberstrategy." *Foreign Affairs* 89, no. 5. https://www.foreignaffairs.com/articles/united-states/2010-09-01/defending-new-domain.

Rogers, M. S. 2017. *Statement of Admiral Michael S. Rogers, Commander United States Cyber Command, Before the Senate Committee on Armed Services, 9 May 2017*. https://www.armed-services.senate.gov/imo/media/doc/Rogers_05-09-17.pdf.

United States Cyber Command (USCYBERCOM). 2018. *Achieve and Maintain Cyberspace Superiority: Command Vision for US Cyber Command*. https://assets.documentcloud.org/documents/4419681/Command-Vision-for-USCYBERCOM-23-Mar-18.pdf.

White House. 2017. *National Security Strategy of the United States of America*. https://www.whitehouse.gov/wp-content/uploads/2017/12/NSS-Final-12-18-2017-0905.pdf.

——— 2018. *National Cyber Strategy of the United States of America*. https://www.whitehouse.gov/wp-content/uploads/2018/09/National-Cyber-Strategy.pdf.

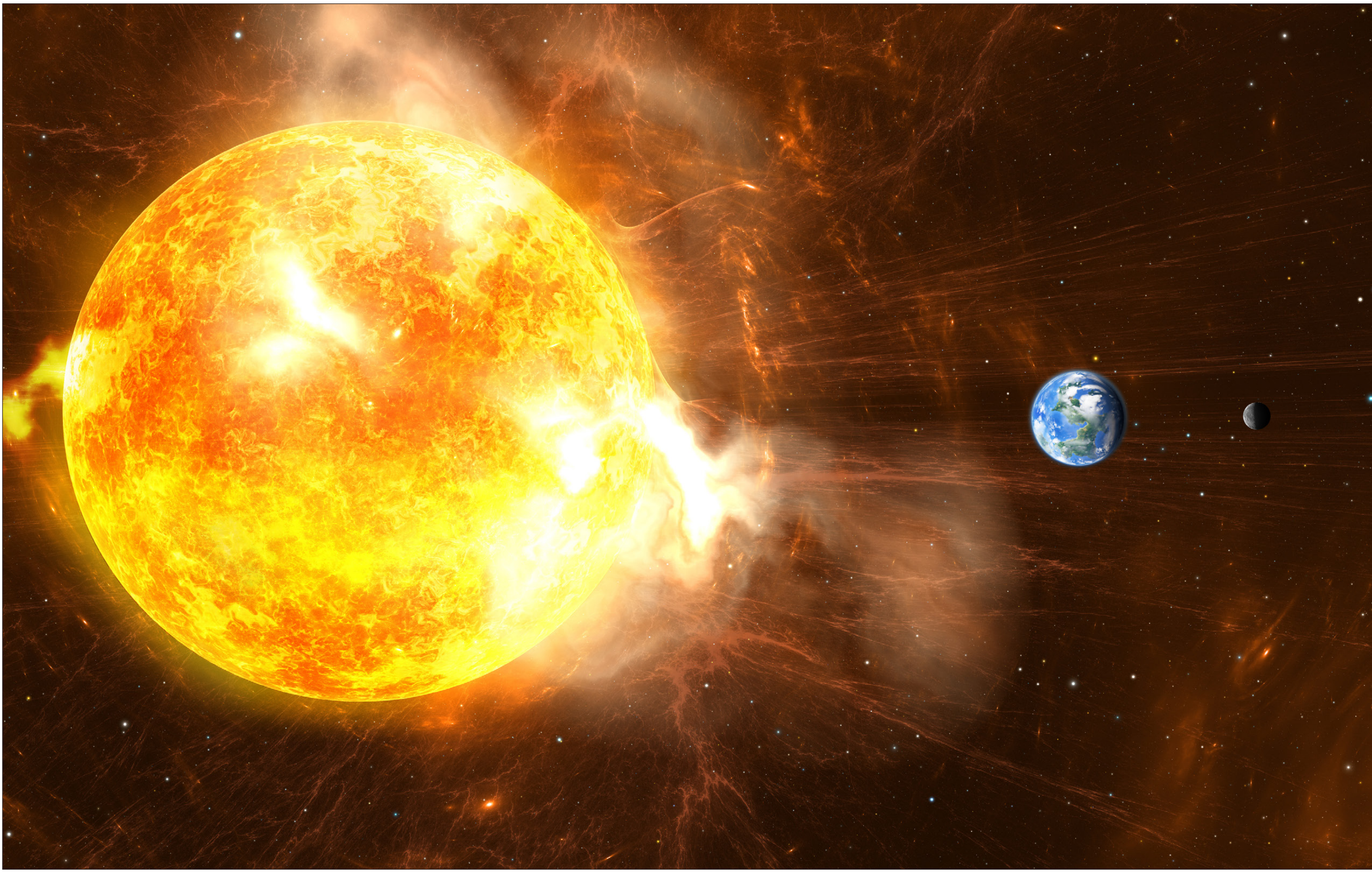
Michael Fischerkeller (with IDA President David S. C. Chu) is a Research Staff Member in the Information Technology and Systems Division of IDA's Systems and Analyses Center. He holds a doctorate in political science from the Ohio State University.

Richard Harknett (right), professor and department head of political science at the University of Cincinnati, holds a doctorate in political science from the Johns Hopkins University.

Brief History of Federal Involvement in Space Weather Forecasting[1]

Becaja Caldwell, Eoin McCarron, and Seth Jonas

Space weather has the potential to adversely affect systems and technologies critical to public health and safety and poses a significant risk to national security. As a result, the federal government has taken efforts to prepare for and mitigate the effects of space weather events. Over the past decade, there has been an increase in federal activities to improve the nation's resilience to the hazards of space weather and to prepare for future space weather events. However, federal involvement in space weather policy and forecasting dates back much further. This paper provides an overview of the history of federal involvement in space weather forecasting from the early space weather-related research and forecasting conducted between World Wars I and II, through the Cold War and the Space Race, to the present.



[1] The original article, "An Abridged History of Federal Involvement in Space Weather Forecasting," was published in *Space Weather*, October 2017, https://doi.org/10.1002/2017SW001626.

Understanding the Ionosphere

The early decades of the 1900s were marked by efforts to understand the ionosphere and its effects on new technologies and to develop and expand ionospheric forecasting capabilities. These efforts were spurred by the discovery and expansion of high-frequency radio wave propagation for long-distance communications and the growth of commercial aviation.

In 1902, scientists Arthur Kennelly and Oliver Heaviside independently published articles suggesting that an ionized upper region in the atmosphere existed that reflected radio waves (Kennelly 1902, 473; Heaviside 1902, 215; Kirby et al. 1934, 16). This layer would become known as the Kennelly-Heaviside layer, but it took over two more decades before direct experimentation confirmed its existence. In 1925, physicist Edward Appleton and his student Miles Barnett proved the existence of the ionosphere through a series of experiments (Appleton and Barnett 1925, 333). This evidence, combined with independent observations made by American physicists Gregory Breit and Merle Tuve, showed for the first time that radio waves could be reflected reliably from the ionized portion of the atmosphere (Breit and Tuve 1925, 357; Appleton and Barnett 1925, 333). The discovery of the ionosphere and the growing use of long-range wireless communication prompted researchers to explore further the connection between the Sun and interruptions in radio transmissions.

In the fall of 1935, a series of severe solar, ionospheric, and magnetic events coincided with radio fade-outs and transmission disruptions. John Dellinger, who would later become chief of the U.S. Department of Commerce National Bureau of Standards (NBS) Radio Section, mapped the reported occurrences of radio fade-outs and their duration to the associated phenomena (Dellinger 1935, 351). He later noted that the increased ionization was caused “by electromagnetic waves from a solar eruption” (Dellinger 1937, 51). Through his observations and analysis, Dellinger determined the existence of a direct correlation between radio fade-outs and disruptions and disturbances on the Sun.

During the early twentieth century, radio communications and aviation advanced as complementary technologies. Aviators came to depend on radio for wireless communications, radar, and navigation, so it is unsurprising the NBS received its first request for information on disruptions to radio communications from the nascent commercial aviation industry (Snyder and Bragaw 1987, 231). A scientist at the NBS prepared a report addressing the issue and providing a recommendation to adjust the frequency to avoid disruptions (Gilliland 1934, 231). This response illustrated the NBS’s expanding role of providing practical support to radio users and, in 1939, the Radio Section initiated a formal service for forecasting radio transmission information and maximum useable frequencies.

These predictions represent the first publicly available, federally developed space weather products and are one of several examples of the growing number of products and services that the NBS began to provide to users (Gilliland et al. 1939, 227). The activities undertaken to understand the ionosphere and its impact on radio wave propagation in the early 1900s, together with the development of services to communicate the effects of ionospheric disruptions, provided a foundation for subsequent federal space weather research in both the civilian and defense sectors.

World War II and Postwar Years

Given the wide usage of radio in wireless communications and navigation, predicting potential ionospheric disruptions and ensuring continuity of these services became critical to the war effort. In 1942, the United States established the Interservice Radio Propagation Laboratory (IRPL) at the Radio Section of the NBS. The functions of the IRPL were to "centralize data on radio propagation and related effects, from all available sources; keep continuous world-wide records of ionosphere characteristics and related solar, geophysical, and cosmic data; and prepare the resulting information and furnish it to the Allied Military Services" (Gladden 1959, 16). By 1942, the IRPL provided Allied armed forces around the globe with predictions of useful radio frequencies for transmission (Cochrane 1966, 405).

In 1943, the IRPL published the "IRPL Radio Propagation Handbook," which described the behavior of the ionosphere and the theory behind lowest and maximum useful frequencies. In addition, the Handbook described the products, forecasts, and warnings provided by IRPL (Cochrane 1966, 404). After the war ended, the federal government recognized the need to continue centralized radio propagation and radio standards and services due to the expanding use of telecommunications among both civilians and the military. In 1946, the Central Radio Propagation Laboratory replaced the IRPL and assumed responsibility for radio propagation research and prediction at the NBS (Gladden 1959, 25). The laboratory continued to expand on the research of its predecessor, specifically by improving the understanding of solar disturbances and how to predict them.

On the military side, the U.S. Air Force (USAF), established in 1947, assumed the responsibility of weather reporting and forecasting for both the USAF and the Army (Nolan and Murphy 2000, 3). The USAF carried out this responsibility through its Air Weather Service. The Air Force Cambridge Research Laboratories (a predecessor to today's Air Force Research Laboratory) was established to continue building on advancements made during World War II including conducting research to improve ionospheric data and forecasting in support of military customers and operations. The USAF established the Sacramento Peak Observatory in New Mexico, operated by the Air Force Cambridge Research Laboratories, which enabled better imaging and analysis of the Sun. The research conducted there would later inform solar observations and forecasting efforts conducted during the Cold War.

The postwar years also saw a growing emphasis on international cooperation, particularly in areas of science. The International Geophysical Year is an important example of such cooperation, and the United States made significant contributions toward its success, particularly for the Third International Geophysical Year, in which the United States provided financial contributions and logistical support (Bullis 1973). Though international cooperation on scientific issues was gaining momentum, it was also strained by growing political tensions between the United States and the Soviet Union.

Cold War and Space Race

Continued federal investment in efforts to improve the understanding of and prediction capabilities for space weather events can make the United States not only better prepared for such hazards, but also better able to mitigate their effects.

The outbreak of the Cold War and the subsequent Space Race spurred U.S. federal investment and interagency cooperation in efforts to explore the space environment. In October 1957, the Soviet Union launched Sputnik. The United States followed with the launch of Explorer I in January 1958. Explorer I led to the development and refinement of research instrumentation, while data gathered from subsequent launches, primarily from Explorer III, led to the discovery of the Van Allen belts, solar radiation, and the magnetosphere (Newell 2010). The discovery of the Van Allen belts and the knowledge of solar radiation and its potential impacts proved critical to the newly established National Aeronautics and Space Administration (NASA) and its manned spaceflight program.

The United States succeeded in its mission to send the first human to the Moon in July 1969 with its Apollo program. Supporting this mission with critical space weather information was the Space Disturbances Laboratory (SDL). The SDL was responsible for operating the Space Disturbances Forecasting Center, which provided space weather forecasting, warnings, and alerts for manned space missions. The SDL also had several research programs aimed at better understanding the space environment and its impacts. These programs included studies of solar energetic particles, the magnetosphere, disturbed ionosphere, and solar physics (Olson 1969, 241).

U.S.-Soviet tensions nearly resulted in armed conflict when space weather storms in May 1967 significantly affected military and civilian operations. As a result, the USAF Air Weather Service was tasked with establishing an operational space weather capability (Knipp et al. 2016, 614; Townsend et al. 1982). The Space Environmental Support System (SESS) was conceived to support operations in space and to develop an operational network of solar optical telescopes to monitor the Sun (Townsend et al. 1982). The SESS and its capabilities continued to evolve as the Department of Defense increasingly came to rely on the USAF for information on the space environment and space assets.

Following successes of both manned and unmanned space exploration missions, the federal government continued to expand its space weather observation and forecasting capabilities. For example, NASA funded a series of ground-based solar observatories, collectively referred to as the Solar Particle Alert Network (SPAN), to support space weather monitoring during the Apollo missions (Reid 1971, 367). The SPAN, established in 1965, consisted of seven solar observatories located around the world that gave 24-hour coverage of solar activity at optical and radio frequencies, providing a near-continuous stream of solar data to better understand and forecast space weather (Robbins and Reid 1969, 502; (Hill et al. 2013, 392).

Federal departments and agencies also sought to develop new and enhanced scientific instruments and better satellite payloads in support of improved solar observational capabilities. For example the Geostationary Operational Environmental Satellite (GOES) program, a joint effort launched in 1975 between NASA and National Oceanic and Atmospheric Administration (NOAA), carries crucial space environment instruments that have been critical to obtaining valuable measurements of solar protons, electrons, and X-rays. The data collected by GOES continue to serve as a resource for space weather forecasters. The technological advancements made during and immediately after the Cold War served as a foundation for future innovation and exploration in the space environment.

Recent History

In the decades since the Cold War to the present, the federal government has funded, led, or supported myriad efforts and activities to improve space weather observation and forecasting capabilities. It has also sought to improve interagency, and international, collaboration and to enhance federal space weather forecasting services to address the hazard of space weather.

In 2014, the National Science and Technology Council established the Space Weather Operations, Research, and Mitigations (SWORM) Task Force, bringing together agencies focused on science and technology with those focused on homeland and national security to produce the National Space Weather Strategy and the National Space Weather Action Plan (Jonas and McCarron 2016, 54). The 2015 National Space Weather Strategy and the National Space Weather Action Plan documents articulate how the federal government will work to enhance national preparedness for space weather events and identifies high-level goals and nearly 100 specific activities in support of these broader goals. These goals include establishing benchmarks for space weather events; improving assessment, modeling, and prediction of impacts on critical infrastructure; and improving space weather services through advances in understanding and forecasting.

Pursuant to Executive Order 13744, “Coordinating Efforts to Prepare the Nation for Space Weather Events,” the SWORM Task Force became a permanent subcommittee of the National Science and Technology Council, serving as the interagency coordination body for space weather across the federal government. The Executive Order builds on the significant progress represented by the National Space Weather Strategy and National Space Weather Action Plan and further establishes the commitment of the federal government to prepare for space weather events.

Conclusion

The evolving threat of space weather to the interconnected electric power grid, satellites in orbit, public health and safety systems, and other critical infrastructure continues to drive federal action and cooperation in space weather forecasting. Continued federal investment in efforts to improve the understanding of and prediction capabilities for space weather events can make the United States not only better prepared for such hazards, but also better able to mitigate their effects. Increasing reliance on technology for the provision of essential services, economic vitality, and social well-being will likely be primary drivers for continued federal involvement in forecasting space weather events.

References

Appleton, E., and M. Barnett. 1925. "Local Reflection of Wireless Waves from the Upper Atmosphere." *Nature,* 115, no. 2888: 333–334. https://www.nature.com/nature/journal/v115/n2888/pdf/115333a0.pdf.

Breit, G., and M. A. Tuve. 1925. "A Radio Method of Estimating the Height of the Conducting Layer." *Nature* 116, no. 2914: 357. https://doi.org/10.1038/116357a0.

Bullis, H. 1973. *The Political Legacy of the International Geophysical Year*. Washington, DC: U.S. Government Printing Office. https://babel.hathitrust.org/cgi/pt?id=uc1.a0000088468;view=1up;seq=1.

Cochrane, R. 1966. *Measures for Progress: A History of the National Bureau of Standards.* Washington, DC: U.S. Department of Commerce.

Dellinger, J. H. 1935. "A New Cosmic Phenomenon." *Science* 82, no. 2128: 351. https://doi.org/10.1126/science.82.2128.351.

Dellinger, J. H. 1937. "Sudden Ionospheric Disturbance." *Terrestrial Magnetism and Atmospheric Electricity* 42, no. 1: 49–53. https://doi.org/10.1029/TE042i001p00049.

Gilliland, T. R. 1934. "Application of Ionospheric Measurements to a Practical Radio Communication Problem." In *Achievement in Radio: Seventy Years of Radio Science, Technology, Standards, and Measurement at the National Bureau of Standards,* edited by W. F. Snyder and C. L. Bragaw, 230–231, Washington DC: National Bureau of Standards.

Gilliland, T. R., S. S. Kirby, and N. Smith 1939. "Characteristics of the Ionosphere at Washington, D.C." P*roceedings of the Institute of Radio Engineers* 27, no. 3: 226–227. https://doi.org/10.1109/JRPROC.1939.228215.

Gladden, S. C. 1959. *A History of Vertical-Incidence Ionosphere Sounding at the National Bureau of Standards*. Washington, DC: U.S. Department of Commerce. Technical Note No. 28. https://www.gpo.gov/fdsys/pkg/GOVPUB-C13-e4f91b98a62d26a7a67171a1d92d7f3d/pdf/GOVPUB-C13-e4f91b98a62d26a7a67171a1d92d7f3d.pdf.

Heaviside, O. 1902. "Telegraph Theory." In *Encyclopedia Britannica,* 10th edition. Vol. 33, 215.

Hill, F., M. J. Thompson, and M. Roth. 2013. "Workshop Report: A New Synoptic Solar Observing Network." *Space Weather* 11: 392–393. https://doi.org/10.1002/swe.20068.

Jonas, S., and E. McCarron. 2016. "White House Releases National Space Weather Strategy and Action Plan." *Space Weather* 14: 54–55. https://doi.org/10.1002/2015SW001357.

Kennelly, A. E. 1902. "On the Elevation of the Electrically-Conducting Strata of the Earth's Atmosphere." *Electrical World and Engineer* 39: 473.

Kirby, S. S., L. V. Berkner, and D. M. Stuart. 1934. "Studies of the Ionosphere and Their Application to Radio Transmission." U.S. Department of Commerce, Bureau of Standards, Research Paper RP632. *Bureau of Standards Journal of Research* 12 (January): 15–51. https://nvlpubs.nist.gov/nistpubs/jres/12/jresv12n1p15_A2b.pdf.

Knipp, D. J., A. C. Ramsay, E. D. Beard, A. L. Boright, W. B. Cade, I. M. Hewins, R. H. McFadden, W. F. Denig, L. M. Kilcommons, M. A. Shea, and D. F. Smart. 2016. "The May 1967 Great Storm and Radio Disruption Event: Extreme Space Weather and Extraordinary Responses." *Space Weather* 14: 614–633. https://doi.org/10.1002/2016SW001423.

Newell, H. 2010. *Beyond the Atmosphere: Early Years of Space* Science. Dover Publications, Inc. Unabridged republication of original published in 1980 as NASA SP-4211.

Nolan, L. E., and J. M. Murphy. 2000. *Air Force Weather: A Brief History*. Offutt AFB, Air Force Weather Agency.

Olson, R. H. 1969. "Solar-Terrestrial Research and Services in the ESSA Research Laboratories Environmental Science Services Administration, Boulder, Colo., U.S.A." *Solar Physics* 8, no. 1: 240–247. https://doi.org/10.1007/BF00150672.

Reid, J. H. 1971. "Solar Activity as Observed by the NASA Solar Particle Alert Network 1976–1969." *Publications of the Astronomical Society of the Pacific* 83, no. 493: 365–369. http://iopscience.iop.org/article/10.1086/129140/pdf.

Robbins, D. E., and J. H. Reid. 1969. "Solar Physics at the NASA Manned Spacecraft Center." *Solar Physics 10*, no. 2: 502–510. https://doi.org/10.1007/BF00145537.

Snyder, W., and C. Bragaw. 1987. *Achievement in Radio: Seventy Years of Radio Science, Technology, Standards, and Measurement at the National Bureau of Standards*. Washington, DC: National Bureau of Standards.

Townsend, R. E., R. W. Cannata, R. D. Prochaska, G. E. Rattray, and J. C. Holsbrook. 1982. *Source Book of the Solar-Geophysical Environment*. Air Force Global Weather Central, Offutt Air Force Base.

Becaja (Bebe) Caldwell (with IDA President David S. C. Chu), a former Research Associate in IDA's Science and Technology Policy Institute (STPI), holds a master's degree in international relations from King's College London. *Eoin McCarron*, a former Science Policy Fellow in STPI, is studying law at University of California, Los Angeles. *Seth Jonas* (no photo available), a Research Staff Member in STPI, holds a doctorate in physics from Johns Hopkins University.

Analysis of an Alternative Military Health Benefit Design[1]

Sarah K. Burns, Philip M. Lurie, John E. Whitley

Congress established the Military Compensation and Retirement Modernization Commission in 2013 to systematically review military compensation and recommend ways to address rising costs and other trends. The commission's recommendation for reforming the TRICARE program was sweeping, and differed greatly from earlier proposals that focused on increasing beneficiary cost shares. Specifically, the commission proposed overhauling the current benefit delivery model and replacing it with a premium-based insurance model offering a menu of private health plans the Department of Defense (DoD) sponsored. An estimate of the budgetary impact of its proposed reforms indicate that movement towards the premium-based model would produce an annual budgetary cost savings in the $2 billion to $4 billion range, with a best savings estimate of $3.2 billion.

[1] Based on "Analysis of an Alternative Military Healthcare Benefit Design," *Defence and Peace Economics*, July 2017, https://doi.org/10.1080/10242694.2017.1349302.

Introduction

Military health care reform is a topic that has received much attention over the last decade, with particular attention to the subject of fiscal sustainability. The costs of the Military Health System (MHS) have grown rapidly during this period, peaking at $53 billion, or roughly 10 percent of the DoD's total outlays, in fiscal year (FY) 2012.[2] Fiscal sustainability is not the only topic driving calls for reform, however. Another topic that has been gaining attention in the reform debate is that of beneficiary satisfaction and access. More specifically, military beneficiaries have consistently reported frustration over their inability to access care in a timely and convenient matter and their limited choice in providers due to the narrow TRICARE network (Military Compensation and Retirement Modernization Commission 2015).

Movement toward a premium-based model would constitute a fundamental shift in DoD health care.

To address these concerns, the Military Compensation and Retirement Modernization Commission proposed a comprehensive reform plan that would have overhauled the current system and replaced it with a premium-based insurance model consistent with an employer-sponsored benefit program that offers a menu of private health plans. Under the commission's proposed policy change, care provisions for active duty service members and Medicare-eligible military retirees covered by TRICARE for Life would remain unchanged. The populations affected by the change would primarily include active duty family members and retirees not yet eligible for Medicare and TRICARE for Life. These beneficiary groups would now select a private health plan and assume financial responsibility for a portion of the premium cost. A Basic Allowance for Health Care would be introduced for all active duty family members to help cover premium shares, co-pays, deductibles, and other out-of-pocket expenses.

Our analysis developed the estimated cost and potential savings from providing a DoD health benefit under such a model. The cost to DoD of purchasing care under such a system would depend on the premium costs of the health plans available within the new program and the enrollment behavior of the eligible population. A cost estimate that would reflect these considerations requires data on a population currently covered under such a system. To meet this requirement, we worked with the Office of Personnel Management to obtain data on the civilian population enrolled in the Federal Employees Health Benefits Program (FEHBP).[3] FEHBP is the largest employer-sponsored health benefit program in the United States, and its enrollees constitute an analytically desirable comparison group for the DoD beneficiary population given the program's size and extensive geographic span.

[2] The FY 2012 Unified Medical Budget was $53 billion. See Defense Health Agency, Support Division (2016).

[3] The Office of Personnel Management provided support for the commission's analysis without endorsing it.

Methodology and Results

Using data on the FEHBP population's demographics, plan choices, and plan costs combined with data on the DoD population, we modeled which FEHBP plans military beneficiaries would select and what premium rates would be set for each plan.

Plan Choice

To develop our cost estimate, we applied federal civilian plan choices to the military beneficiary population, using data on current FEHBP enrollees. A simple approach would be to obtain the distribution of plan enrollment for this population and allocate the DoD population across each plan accordingly (e.g., if 44 percent of FEHBP contract holders are enrolled in the BlueCross BlueShield Standard plan, we would assume 44 percent of DoD beneficiaries will select this plan). However, this would fail to account for important differences in the demographic, socioeconomic, and geographic composition of the FEHBP and DoD populations. The age distributions for the two beneficiary populations illustrates this point. A glance at Table 1 reveals that the DoD population is significantly younger than the FEHBP population. Nearly 50 percent of the DoD population is under age 35, while less than 10 percent of FEHBP population falls into this category. Conversely, for the categories that would be eligible for the proposed policy change, less than 1 percent of the DoD population are over age 65, compared to nearly 36 percent of the FEHBP population.

Table 1. Enrollee population age comparison, FY 2013

	FEHBP Contract Holders			DoD Sponsors		
Age	**Count**	**Percentage**	**Cumulative Percentage**	**Count**	**Percentage**	**Cumulative Percentage**
<23	3,938	0%	0%	413,703	14%	14%
23–34	358,678	9%	9%	894,572	31%	46%
35–44	475,730	12%	21%	431,988	15%	61%
45–54	750,288	19%	39%	518,715	18%	79%
55–64	1,003,588	25%	64%	595,488	21%	100%
65–74	694,849	17%	81%	4,819	0%	100%
75+	753,857	19%	100%	3,734	0%	100%
Total	**4,040,928**			**2,863,019**		

Note: The FEHBP age distribution is based on the age of all contract holders enrolled in the system (active employees and annuitants). The DoD age distribution is based on all active duty and non-Medicare-eligible retiree sponsors.

To properly account for such differences in the composition of the two populations, a cohort-based approach was implemented. This allowed the DoD population to be allocated across plans based on within-group enrollment distributions. The cohort grouping was based on observable demographic and socioeconomic factors known to influence health plan choice. While many demographics are thought to have some bearing on plan choice, age (which can be viewed as a proxy for health and expected expenditures) and income are widely recognized as the most important (Scanlon et al. 1997). Geographic considerations are also important, given that many plans are available only in select market areas. The cohort grouping for this analysis was therefore based on age, income, and state of residence.

Premium Adjustments Choice

The cohort methodology allows us to control for some of the compositional differences between the FEHBP and DoD beneficiary populations when modeling the predicted enrollment behavior of DoD beneficiaries. However, plan choice is not the only parameter affected by the demographic composition of beneficiary populations. Premium amounts must also be considered.

Under a premium-based model, participating health plans assume the financial risk for the beneficiary population they cover. Insurance underwriters therefore determine plan premiums based upon a careful assessment of each population's specific risk pool. For instance, even when controlling for age, a significant difference in health may still exist between the average 17- to 24-year-old male in the FEHBP population compared to the average 17- to 24-year-old male in the DoD population. To account for these factors fully, insurers calculate risk scores based on claims data for subsets of beneficiaries (such as 17- to 24-year-old males) within a population. These risk scores, together with the populations' composition, determine the premium amounts. Our analysis developed a methodology to adjust each plan's premium to reflect the characteristics of the DoD population projected to enroll in the plan. It involved adjustments for population risk score, population composition factor, and retirees' use of the Department of Veterans Affairs (VA) and other (civilian) health insurance.

Results

The importance of these adjustments was found to be significant—especially the PCF adjustment. This is illustrated by Table 2, which shows the total estimated premium costs as each adjustment is applied.

Table 2. Unadjusted and adjusted premium cost estimates (millions)

Estimate	Population risk score	Population composition factor	VA & other health care	Total cost to DoD
Unadjusted	—	—	—	$22,152
Partially adjusted	x	—	—	$21,770
	x	x	—	$18,907
Final	x	x	x	$18,046

The combination of adjustments combined reduced our estimated cost of delivering care under the commission's proposed reform by just over $4.1 billion, resulting in a final estimate of $18 billion.[4]

Discussion of Results

Determining whether our final baseline estimate represents a cost decrease or increase requires an estimate of what DoD currently spends providing a health benefit to this population. The DoD premium equivalent cost, or the cost of covering the same population under the current program, was estimated to be $21.2 billion, suggesting a baseline annual savings of $3.2 billion.[5] Sensitivity analyses showed variations in those savings ranged generally from between $2 billion and $4 billion, although some sensitivity analyses found wider ranges. For instance, if we assume that the Medicare-eligible population in FEHBP costs less than we predicted, the resulting premium reduction factor would be low and our savings estimate would fall to $822 million. In another excursion, we estimated savings would be just under $7.5 billion if all beneficiaries were placed in a lower cost plan, using Government Employees Health Association (GEHA) as an example.[6]

The GEHA example provided an interesting illustration of the magnitude of savings that could be gained from switching from the current TRICARE model to a private insurance model. Under the commission's proposal, where beneficiaries were free to select their health plan, we estimated DoD would see a budgetary savings of roughly $3.2 billion dollars. The quality of the benefit was not held constant under this reform proposal; however, beneficiary choice and access were greatly increased. If DoD were to attempt a quality-neutral type reform—replace the TRICARE plan with a private plan like GEHA that approximately equals TRICARE in non-price quality attributes—savings could more than double. To test whether the GEHA plan was similar to TRICARE in terms of non-price quality attributes, we explored several comparison metrics, including network size, patient satisfaction, access standards, and covered services. Our analysis concluded that the GEHA plan generally had more providers than the TRICARE network, slightly higher beneficiary satisfaction, and similar access standards and covered services.[7]

[4] The weighted premiums used to construct these cost estimates are contained in Appendix A of Burns et al. (2015).

[5] The DoD premium equivalent cost was a concept created to ensure a fair comparison. We attempted to identify all costs associated with delivering care to the population of interest that would have been covered by premiums under a premium-based model. We included certain budgeted costs associated with overhead, management, and capital but excluded costs associated with readiness (for example readiness and training). See Burns et al. (2015) for an explanation of the development of the DoD premium equivalent cost.

[6] GEHA Standard seemed a natural candidate for the comparison analysis, given it was the plan with the third-highest predicted DoD enrollment (after BlueCross BlueShield Basic and Standard) but had a relatively low premium cost.

[7] The full network comparison analysis can be found in Burns et al. (2015).

References

Burns, S. K., P. M. Lurie, and S. A. Horowitz. 2015. *Analyses of Military Healthcare Benefit Design and Delivery: Study in Support of the Military Compensation and Retirement Modernization Commission*. IDA Paper P-5213. Alexandria, VA: Institute for Defense Analyses. https://idalink.org/P-5213.

Defense Health Agency, Support Division. 2016. *Evaluation of the TRICARE Program: Access, Cost, and Quality, Fiscal Year 2016 Report to Congress.* Office of the Assistant Secretary of Defense (Health Affairs).

Military Compensation and Retirement Modernization Commission. 2015. *Report of the Military Compensation and Retirement Modernization Commission: Final Report*. http://www.dtic.mil/dtic/tr/fulltext/u2/a625626.pdf.

Scanlon, D. P., M. Chernew, and J. R. Lave. 1997. "Consumer Health Plan Choice: Current Knowledge and Future Discussions." *Annual Review of Public Health* 18, no. 1: 507–528. https://doi.org/10.1146/annurev.publhealth.18.1.507.

Phil Lurie (left), a Research Staff Member in the Cost Analysis and Research Division (CARD) of IDA's Systems and Analyses Center, holds a doctorate in statistics from Harvard University. *John Whitley* (center), a former Adjunct Research Staff Member in CARD, holds a doctorate in economics from the University of Chicago. *Sarah Burns* (right), a Research Staff Member in CARD, holds a doctorate in economics from the University of Kentucky.

Effects of Violence on Voter Turnout in Sub-Saharan Africa[1]

Dorina A. Bekoe and Stephanie M. Burchard

Because of its coercive nature, many researchers have assumed that election-related violence has a depressive effect on voter turnout. Out of fear for physical safety or the desire to keep out of harm's way, potential voters might remain home and abstain from the polls in the face of violent threats. The empirical record, however, does not substantiate this assumption. After examining violence and voter turnout in nearly 300 elections held in sub-Saharan Africa from 1990 to 2014, we find no significant aggregate effect of pre-election violence on voter turnout. A closer look at the nature of election violence and its intended targets explains this finding. Violence entrepreneurs strategically employ violence for a multitude of sometimes conflicting reasons. For some audiences, coercion is used to mobilize support, and for others, it is used to prevent electoral participation. And sometimes violence is used to displace potential voters and change the partisan competition of constituencies.

[1] Originally published as "The Contradictions of Pre-election Violence: The Effects of Violence on Voter Turnout in Sub-Saharan Africa," *Africa Studies Review*, September 2017.

Introduction

Over the last ten years, the phenomenon of electoral violence has gained considerable attention from policy makers, practitioners, and academics. This field of study has now produced many works investigating the underlying rationale, dynamics, and consequences of electoral violence (e.g., Höglund 2009; Bekoe 2012; Hafner-Burton et al. 2014; Burchard 2015). Recent research indicates that, at least in the case of incumbents, violence is frequently used as a strategy when a politician is uncertain about the likelihood of victory or fears the loss of a political position, particularly in an environment of weak institutions and few consequences of violence (Hafner-Burton et al. 2014). The dominance of pre-election violence, in particular, indicates that the purpose of the violence is to influence the election through intimidation, harassment, assassination, or other large-scale acts of aggression. In certain cases, pre-election violence has resulted in a politician's withdrawal from the contest (e.g., Morgan Tsvangirai in Zimbabwe in 2009) or a boycotting of the election by the opposition party (e.g., in Burundi in 2010)—mostly to the benefit of the party most responsible for the violence. Beyond these national-level effects, however, the influence of electoral violence—specifically, the effect of pre-election violence on voter turnout—has been unclear.

The working assumption by the policy and academic communities is that voter turnout is negatively affected by pre-election violence. Indeed, the possibility of lower voter turnout in the face of pre-election violence is one of the driving factors behind the electoral security framework developed by the United States Agency for International Development (USAID). USAID's *Electoral Security Framework* asserts that voter turnout is suppressed when insurgents delay or discredit an election; when candidates attempt to "capture an election"; when political parties boycott the polls; or as a direct consequence of electoral violence (USAID 2010, 6). Similarly, the United Nations Development Program (UNDP) guide *Elections and Conflict Prevention* states that voter turnout may be decreased by the use of violence by political parties or armed groups in order to ensure a particular outcome (UNDP 2009, 5). Scholars also assume that voter turnout is generally negatively affected by electoral violence: Höglund (2009, 412) states that "voter turnout may be influenced if large sections of the population refrain from casting their vote due to fear of violence." Individual case studies of Nigeria's 2007 election also start from an assumption that violence affects voter turnout (e.g., Bratton 2008; Collier and Vicente 2011). Thus, from both a policy and an academic perspective, it is accepted as fact that violence leads to fewer people showing up at the polls.

Despite this inclination to view pre-election violence as a suppressant of voter turnout, it has not been clear how—or even if—this takes place. Politicians and political parties that employ electoral violence are often interested in affecting the results of an election, not in suppressing voting per se. In Zimbabwe's 2008 election, violence was used to punish opposition supporters, as well as to persuade people to vote for the ruling party (Human Rights Watch 2008). In Ethiopia's 2010 election, many were intimidated into voting for the government

(Human Rights Watch 2010). In Kenya, violence was also used to turn out voters. A closer look at the data is needed to determine the motivations and effects of pre-election violence.

Voter Turnout and Pre-Election Violence

Under a democratic system in which political participation is voluntary, voter turnout is the sum effect of citizen involvement in the formal exercise that selects a country's political leadership. According to the International Institute for Democracy and Electoral Assistance (International IDEA 2017), average voter turnout in Africa is 65 percent. High voter turnout generally reflects an energized constituency that sees value in the effort required to cast a ballot, while low voter turnout may reflect a paucity of electoral options or low interest in the outcome of the election. Low voter turnout may also indicate that voters lack confidence in the electoral process or in the legitimacy of the existing regime. In either case, voters may refrain from voting if they believe their vote will have little effect on the outcome (Karp and Banducci 2008; Birch 2010).

Some (e.g., USAID 2013) argue that low voter turnout signals trouble in a young or fragile democracy and that electoral violence is a direct cause. Unfortunately, however, in this context the meaning of voter turnout is particularly difficult to interpret. Countries transitioning to democracy from authoritarian regimes may not have the necessary safeguards in place to ensure a free or fair vote, and in some cases, turnout can be coerced and artificially inflated.

Politicians and political parties that employ electoral violence are often interested in affecting the results of an election, not in suppressing voter turnout.

Data Analysis

Our primary motivation was to examine how pre-election violence affects voter turnout. We began our analysis with the assumption that instigators of violence use it to deter participation due to the simple fact that voting becomes more cumbersome when the threat of violence looms. Following this logic, we hypothesized that pre-election violence should deter participation and therefore decrease turnout, all else being equal.

We tested our hypothesis using multiple methods and different levels of data, building upon the African Election Violence Database assembled by Straus and Taylor (2012). For the years 1990–2008, Straus and Taylor categorized the level of violence during the six months prior to an election and the three months after an election for each election in sub-Saharan Africa. The categories were 0 for cases in which no violence occurred; 1 for cases in which voter intimidation and harassment occurred; 2 for cases in which violent repression, including political assassinations and fatalities, occurred; and 3 for elections in which large-scale violence took place with at least twenty reported fatalities. Using the same scheme, we updated the data set to cover elections that were held through 2014. For the purposes of our analysis, we collapsed the four categories into a dummy variable; however, in order to address the concern that the severity of electoral violence could also have an impact on voter turnout, we conducted all analyses using both our binary treatment and Straus and Taylor's original scheme, which treats electoral violence as an ordinal-level variable.

Our data set contained a total of 287 observations of elections in 47 countries. We conducted separate analyses of legislative voter turnout (including both singular and concurrent elections, for a total of 191 elections) and executive voter turnout (again, including both singular and concurrent elections, for a total of 166 elections). In our sample, average voter turnout in Africa for presidential and legislative elections was nearly the same: 66 percent and 63 percent, respectively.

Table 1 reports average voter turnout by election type (executive or legislative) and incidence of electoral violence. These data come from the pooled data set that includes all elections in all countries with available data. The differences in average turnout are not statistically significant. Complicating our data analysis was the fact that some countries in our sample have historically had violent elections (Kenya and Zimbabwe) and others have never had them (Botswana, São Tomé, and Príncipe). In these extreme cases, the key independent variable shows no variation, so absence or presence of violence cannot explain variation in voter turnout over time. Our solution was to perform an isolated analysis of countries that do demonstrate variance in the absence or presence of electoral violence over time. This removed approximately 40 percent (19) of the countries in our sample and left us with data from 28 countries to examine.

Table 1. Voter Turnout and Violence, Pooled Sample

Election violence	Executive turnout	Legislative turnout
Violence before election	67.4% (*n* = 101)	62.4% (*n* = 102)
No violence before election	63.7% (*n* = 65)	63.8% (*n* = 89)
t-test	*t* = –1.45, *p* = 0.15	*t* = 0.53, *p* = 0.59

Table 2 reports voter turnout by type of election and whether violence took place before the election or not for our isolated sample. While turnout was on average lower in legislative elections where violence occurred, the difference is not statistically significant. Based on this descriptive analysis, thus far there appears to be no significant difference in voter turnout between violent elections and nonviolent elections.

Table 2. Voter Turnout and violence, isolated sample

Election violence	Executive turnout	Legislative turnout
Violence before election	65.6% (*n* = 66)	59.0% (*n* = 58)
No violence before election	62.3% (*n* = 46)	59.3% (*n* = 58)
t-test	*t* = –1.19, *p* = 0.23	*t* = 0.11, *p* = 0.90

In addition to performing descriptive analysis, we tested our hypothesis using generalized least squares (GLS) regression analysis on our isolated sample. Due to the structure of our data set— elections nested within countries and variation in number of elections per country resulted in unbalanced short-panel data— we addressed dependency within panels/countries (Gelman and Hill 2006). By including random effects in our model, we accounted for unspecified country-level effects that could potentially bias our estimates.

We ran several regression analyses with voter turnout as our dependent variable and election violence as our key independent variable. To identify the relevant control variables, we relied specifically on the literature on voter turnout and African voters. Much of the broader literature on voter turnout focuses on how institutional, political, and socioeconomic factors affect voter turnout (Blais 2006; Geys 2006). Proportional electoral institutions are generally found to increase voter turnout, whereas plurality/majoritarian electoral institutions tend to decrease it (Banducci and Karp 2009). We determined type of electoral system using a categorical variable, where the values 1–4 correspond to plurality, majoritarian, mixed, and proportional representation electoral rules, respectively.

We ran separate random-effects GLS regressions for executive and legislative turnout with controls for electoral system, type of election, political climate, and socioeconomic status. In all model specifications, the coefficient for violence was negative but insignificant. In none of the models did it come close to reaching significance. In both executive and legislative elections the "youth" bulge was significant and negative, meaning that countries with younger populations overall have lower than average voter turnout rates compared to countries with older populations. In executive elections, incumbent participation was significant and positive (for one of the models), meaning that when an incumbent executive runs for re-election, voter turnout increases. This may reflect intense mobilization efforts that incumbent presidents undertake, in part due to their access to state resources. For legislative elections, this finding was inconsistent across our two measures of political environment.

Based on our cross-national analysis, election violence does not appear to affect voter turnout in the aggregate.

Conclusion

Electoral violence has many motivations. In Kenya, for example, violence has been used to suppress, motivate, or punish voters. Moreover, different actors have fomented the violence. In early elections, the Kenyan government was the main perpetrator, but violence was also used by opponents in later years and at the subnational level in 2013. In addition, the impact of electoral violence on voter turnout can vary because voters react to violence in different ways: they may flee the country or stay home but not vote or they may adjust their vote. Voter response can depend on how widespread the violence is, how much risk the voters are willing to bear, and how they view the election. The rate of violence preceding the 2013 Kenyan election was higher than that preceding the 2002 and

2007 elections, yet voter turnout was higher. However, the 2013 elections were also publicized as an opportunity for the country to move beyond the violence of 2007; they were managed by a more respected electoral commission and commissioner, framed by a relatively well-received new constitution, conducted under the aegis of a well-respected and newly reformed judiciary, and monitored by a national and international institutions.

Does pre-election violence, then, suppress voter turnout, as we hypothesized? Our overall conclusion is that over time and across countries in Africa, electoral violence does not result in lower voter turnout. Indeed, it has no perceptible overall effect. Pre-election violence and its intended effects are specific to each situation—resulting in either suppressing voters or pushing them to turn out at the polls—congruent with the goals of the perpetrators and electoral environment. Pre-election violence, it seems, can achieve many objectives, depending on the political and social context. This finding suggests the need for a more nuanced analysis—one that looks more closely at the rhetoric surrounding specific elections, the motivations behind electoral violence, and the coercive powers of the perpetrators of violence.

References

Banducci, S., and J. A. Karp. 2009. "Electoral Systems, Efficacy, and Voter Turnout." In *The Comparative Study of Electoral Systems*, edited by Hans-Dieter Klingemann, 109–34. Oxford: Oxford University Press.

Bekoe, D. A., ed. 2012. *Voting in Fear: Electoral Violence in Sub-Saharan Africa*. Washington, D.C.: United States Institute of Peace Press.

Birch, S. 2010. "Perceptions of Electoral Fairness and Voter Turnout." *Comparative Political Studies* 43 (12): 1601–22. https://doi.org/10.1177/0010414010374021.

Blais, Ae. 2006. "What Affects Voter Turnout?" *Annual Review of Political Science* 9: 111–25. https://doi.org/10.1146/annurev.polisci.9.070204.105121.

Bratton, M. 2008. "Vote Buying and Violence in Nigerian Election Campaigns." *Electoral Studies* 27: 621–32. https://doi.org/10.1016/j.electstud.2008.04.013.

Burchard, S. 2015. "The Resilient Voter? An Exploration of the Effects of Post-Election Violence in Kenya's Internally Displaced Person Camps." *Journal of Refugee Studies* 28, no. 3: 331–49. https://doi.org/10.1093/jrs/feu038.

Collier, P., and P. C. Vicente. 2013. "Votes and Violence: Evidence from a Field Experiment in Nigeria." *Economic Journal* 124, no. 574: 356–87. https://doi.org/10.1111/ecoj.12109.

Gelman, A., and J. Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge, U.K.: Cambridge University Press.

Geys, B. 2006. "Explaining Voter Turnout: A Review of Aggregate-Level Research." *Electoral Studies* 25, no. 4: 637–63. https://doi.org/10.1016/j.electstud.2005.09.002.

Hafner-Burton, E. M., S. D. Hyde, and R. S. Jablonski. 2014. "When Do Governments Resort to Election Violence?" *British Journal of Political Science* 44, no. 1: 149–79. https://doi.org/10.1017/S0007123412000671.

Höglund, K. 2009. “Electoral Violence in Conflict-Ridden Societies: Concepts, Causes, and Consequences.” *Terrorism and Political Violence* 21, no. 3: 412–27. https://doi.org/10.1080/09546550902950290.

International IDEA. 2017. Voter Turnout Database. https://www.idea.int/data-tools/data/voter-turnout.

Human Rights Watch. 2008. “*Bullets for Each of You”: State-Sponsored Violence since Zimbabwe’s March 29 Elections*. https://www.hrw.org/reports/2008/zimbabwe0608/zimbabwe0608webwcover.pdf.

———. 2010. “‘One Hundred Ways of Putting Pressure’: Violations of Freedom of Expression and Association in Ethiopia.” https://www.hrw.org/report/2010/03/24/one-hundred-ways-putting-pressure/violations-freedom-expression-and-association.

Karp, J. A., and S. A. Banducci. 2008. “Political Efficacy and Participation in Twenty-Seven Democracies: How Electoral Systems Shape Political Behavior.” *British Journal of Political Science* 38, no. 2: 311–334. https://doi.org/10.1017/S0007123408000161.

Straus, S., and C. Taylor. 2012. “Democratization and Electoral Violence in Sub-Saharan Africa, 1990–2008.” In *Voting in Fear: Electoral Violence in Sub-Saharan Africa*, edited by Dorina A. Bekoe, 15–38. Washington, D.C.: United States Institute of Peace Press. https://doi.org/10.1017/asr.2017.50.

United Nations Development Program (UNDP). 2009. *Elections and Conflict Prevention: A Guide to Analysis, Planning and Programming*. http://www.undp.org/content/undp/en/home/librarypage/democratic-governance/electoral_systemsandprocesses/elections-and-conflict-prevention-guide.html.

United States Agency for International Development (USAID). 2010. *Electoral Security Framework: Technical Guidance Handbook for Democracy and Governance Officers*. https://www.usaid.gov/sites/default/files/documents/1866/1-Electoral-Security-Framework.pdf.

———. 2013. *Best Practices in Electoral Security: A Guide for Democracy, Human Rights and Governance Programming*. https://www.usaid.gov/sites/default/files/documents/2496/Electoral_Security_Best_Practices_USAID.pdf.

Stephanie Burchard (left), a Research Staff Member in the Intelligence Analyses Division (IAD) of IDA’s Systems and Analyses Center, holds a doctorate in political science from Rice University. *Dorina Bekoe*, also a Research Staff Member in IAD, holds a doctorate in public policy from Harvard University.

Power Approximations for Generalized Linear Models Using the Signal-to-Noise Transformation Method[1]

Thomas H. Johnson, Laura J. Freeman, James D. Simpson, and Colin E. Anderson

Statistical power is a useful measure for assessing the adequacy of an operational test. It is the probability of correctly concluding that a factor in the experiment significantly impacts the response variable. For normally distributed response variables, power calculations are widely available in experimental design software. However, many defense testing applications use non-normal response variables. Generalized linear models provide many useful analysis methods for non-normal responses. While statistical software routinely includes generalized linear models in model-fitting packages, power calculations for generalized linear models are not widely available in experimental design modules. This paper proposes a signal-to-noise transformation method (SNRx) that enables generalized linear model power approximations using normal linear model power equations, making them generally available to all practitioners.

[1] Based on "Power Approximations for Generalized Linear Models Using the Signal-to-Noise Transformation Method," *Quality Engineering*, October 2017, https://www.tandfonline.com/doi/full/10.1080/08982112.2017.1361537.

Introduction

Experimental designs are used to help with planning, executing, and analyzing an experiment. In the planning phase test objectives are determined. These objectives guide the development of the factors, levels, and response variables (Freeman et al. 2013). Recent Department of Defense policy has emphasized the importance of using principles of design of experiments in all operational testing (Johnson et al. 2012, 61).

Our goal is to provide a simple method to obtain power for a generalized linear model by transforming the effect size in the power calculation for a classical linear model.

Equally important in the planning phase is the assessment of the experimental design. An assortment of measures is available to assess the goodness of an experiment prior to data collection. Hahn, Meeker, and Feder (1976) call these *measures of precision*. These include standard error of predicted mean responses, standard error of coefficients, correlations metrics, and optimality criteria values. Measures of precision are affected by many aspects of the plan for the experiment, including the choice of factors and levels, the assumed model form, the combination of factor settings from run to run, and the total number of runs.

Power—the focus of this paper—is an important measure of precision. Power is the probability of correctly concluding that an effect has an impact on the response variable. In general, the power of an effect increases with sample size, making it a useful measure for determining the scope of an operational test. Here, we focus on a second-order model for designs with multilevel categorical factors. Effects considered include the main effects and two-factor interactions (Montgomery 2008, 4).

Experimental design software that calculates power for classical linear models is widely available. However, power calculations should reflect the knowledge that the result will not be normally distributed, when it is known before running the experiment. Techniques for calculating power for experimental designs with generalized linear models are not widely available in commercial software; such calculations usually require Monte Carlo simulation studies. Accounting for the knowledge of the planned analysis is important when planning the test because different distributions can require dramatically different sample sizes to achieve high-effect power.

Our goal is to provide a simple method to obtain power for a generalized linear model by *transforming* the effect size in the power calculation for a classical linear model. Existing software (e.g., JMP, Minitab, and Design Expert) that accommodates classical linear model power calculations allows the user to adjust the signal-to-noise ratio or alter the model coefficients under the alternative hypothesis. SNRx provides a means of setting the signal-to-noise ratio or the coefficients so that the calculation represents the generalized linear model power calculation. The target audience of SNRx is the analyst who has statistical design experience and is comfortable working with popular statistical software, but who is not inclined to calculate power for generalized linear models using custom code and Monte Carlo simulation.

Model Formulation

A generalized linear model generalizes the classical linear model and is defined in terms of its three components (McCullagh and Nelder 1989, 27):

Random component. Response variables $Y_1,..., Y_n$ share the same distribution from the exponential distribution family, where the vth response of the experiment has an expected value equal to the mean, μ_v.

Systematic component. The unknown coefficients systematically specify the linear predictor η_v such that $\eta_v = Z_v \psi + X_v \lambda$, where Z_v and X_v represent the vth row of the test and nuisance matrix.

Link between the random and systematic components. The link function $g(\cdot)$ relates the mean and linear predictor in the expression $(\mu_v) = \eta_v$.

Generalized linear models may also include as special cases linear regression, logistic regression, and log-linear models for count data.

Model Inference

We are interested in a hypothesis test for the significance of a multilevel categorical factor or interaction between multilevel categorical factors. Specifically, we want to be able to test whether the coefficients belonging to a main effect or two-factor interaction effect are equal to zero. Thus, the hypothesis test for an individual effect is

$$H_0: \psi = 0,$$

$$H_1: \psi \neq 0.$$

The classical and generalized linear models use similar techniques for evaluating these hypothesis tests. A classical linear model uses analysis of variance (ANOVA), which is based on an F statistic. The analogue of an ANOVA for generalized linear models is an analysis of deviance, which is based on a likelihood ratio statistic.

Some classical linear model software allows the user to specify the details of a planned experiment, and the software outputs the power associated with this hypothesis test. The user can input the design matrix, choose the model form, set the anticipated coefficients (i.e., set ψ under H_1), and obtain power.

The SNRx method is useful in situations where the practitioner only has access to classical linear model software, but is interested in calculating power for a specific generalized linear model. In this situation, the SNRx method sets ψ under H_1 so that the ANOVA hypothesis test well represents an analysis of deviance for the specific generalized linear model.

SNRx Method

The approach assumes that for each run in the experiment $(v = 1, 2,..., N)$ the linear predictors η_v in a generalized linear model can be modeled as the response variable Y_v in a classical linear model. That is, $Y_v = \eta_v = Z_v\,\psi + X_v\lambda + \epsilon_v$, where $\epsilon_v \sim N(0, \sigma^2)$, and the error term ϵ_v is independent and identically distributed. The variance σ^2 is the transformed noise, meaning it represents the variance of the linear predictor for the generalized linear model.

Another assumption in this approach is that σ^2 is constant and is evaluated at the overall mean across the design space $\bar{\mu}$. For example, an analyst may anticipate a 70 percent average probability of success across the design space that can be fit with a logistic regression model. The overall mean $\bar{\mu}$ impacts σ^2 and, in turn, affects power.

A tenet of generalized linear models is that the variance of Y depends on the mean μ and the dispersion parameter ϕ. Since we are assuming a nonzero effect size for ψ under the alternative hypothesis, an implication is that μ is not constant; thus, neither is σ^2. For this reason, only small effect sizes should be considered.

Another assumption is that the hypothesis test is constructed without considering nuisance effects. That is, for the hypothesis test $\psi = 0$, the nuisance coefficients take the form $\lambda = (\lambda_{int}|0)T$. Without this assumption, significant values of λ could further invalidate the assumption that σ^2 is constant because λ impacts μ, which, in turn, affects the variance of Y.

We define the signal-to-noise ratio as $\kappa = \delta/\sigma$. For SNRx, we must transform δ and σ to the linear predictor space. Since Y is a random variable with $E(Y) = \mu$, we can use $g(Y)$ as an estimator of $g(\mu)$. Using the delta method from Casella and Berger (2002), we can approximate that

$$\mathrm{E}\big(g(Y)\big) \approx g(\mu),$$

$$\mathrm{Var}\big(g(Y)\big) \approx [g'(\mu)]^2\mathrm{Var}(Y).$$

We also know that $\mathrm{Var}(Y) = a(\phi)\mathrm{Var}(\mu)$ for generalized linear models. Substituting this into the above equation, taking the square root, and evaluating $g'(\mu)$ and $\mathrm{Var}(\mu)$ at $\bar{\mu}$, we obtain the following estimate of the noise:

$$\sigma = \sqrt{\mathrm{Var}(g(Y))} = g'(\bar{\mu})\sqrt{a(\phi)\mathrm{Var}(\bar{\mu})}.$$

Now that the noise is transformed, we turn our attention to the signal. If the upper and lower bounds of the signal of interest are $\bar{\mu} + \delta/2$ and $\bar{\mu} - \delta/2$, we can convert this quantity to a value in the linear predictor space as $g(\bar{\mu} + \delta/2)$ and $g(\bar{\mu} - \delta/2)$, respectively, where $g(\cdot)$ is the link function for the generalized linear model of interest.

The signal-to-noise ratio is described as the ratio of the signal and noise within the linear predictor space, as shown in the equation below.

$$\kappa = \frac{g(\overline{\mu} + \delta/2) - g(\overline{\mu} - \delta/2)}{g'(\overline{\mu})\sqrt{a(\phi)V(\overline{\mu})}}$$

Mission Success Example

In logistic regression, the response variable is binary (1 or 0). For this example, let 1 and 0 represent a mission success and failure, respectively. In an experiment with N groups or strata, Y_v represents the number of successes in the vth group out of m_v attempts, where $v = 1, 2, \ldots, n$. Then, a logistic regression model assumes that $Y_v \sim \text{binom}(m_v, \pi_v)$.

A few pieces of information are needed to set up the power calculation. The first is the assumed mean response across the design space $\overline{\mu}$. For logistic regression, the mean response is bounded between zero and one and represents the average probability of success across the design space. For this example, we assume a nominal 70 percent probability of success, or $\overline{\mu} = 0.7$.

The second element is the effect size δ. Recall that δ is the change in the mean response that is symmetric about $\overline{\mu}$. In this example, we assume $\delta = 0.3$ so that the change of interest ranges from 55 to 85 percent probability of success.

The next step is to calculate the signal-to-noise ratio κ. The signal-to-noise ratio can be directly inputted into some software, such as Design Expert, and the corresponding effect power is outputted. In other software, such as JMP, the coefficients anticipated under the alternative hypothesis must be manually inputted using the approach outlined below. Using the assumed values for this example, we get

$$\kappa = \frac{g(\overline{\mu} + \delta/2) - g(\overline{\mu} - \delta/2)}{g'(\overline{\mu})\sqrt{a(\phi)V(\overline{\mu})}}$$

To obtain the approximate coefficients, we first construct the marginal mean effect so that its range is equal to κ and then convert it to coefficients. The coefficients for a three-level main effect are

$$\psi = [.70/2 - .70/2)]T.$$

In this example, assume the experiment includes three factors and the sample size is 96. That is, the operational test includes 96 missions. The experimental design is a full factorial that is replicated four times so the model matrix M is size 96 × 18. The first column of M corresponds to the intercept, columns 2 through 7 correspond to the main effects, and columns 8 through 18 correspond to the two factor interactions. The coefficient vector β is size 18 × 1. The power calculation requires that we split the model matrix into the test matrix Z and the nuisance matrix X.

For the test on the main effect, the test matrix Z is size 96 × 2, and the previously calculated test coefficient vector ψ is size 2 × 1. The nuisance matrix is 96 × 16. We calculate the hat matrix W, and use W, Z, and ψ in the equation for the noncentrality parameter, which is given as

$$\gamma_F = (Z\psi)^T (\mathrm{I} - W)(Z\psi),$$

and we find that $\gamma_F = 7.91$. By setting the significance $\alpha = 0.05$, we then calculate the critical F value that is equal to $f_{crit} = 3.11$. Finally, we calculate power, which is equal to 0.69. Clearly, 96 missions does not provide enough power to determine if the main effect significantly affects mission success. Additional missions are required to provide a robust evaluation. Further details about this calculation can be found in the full-length version of this paper.

Conclusion

This work provides a practical approach for sizing operational tests. Compared to current approaches, our hope is that this methodology will be more accessible to the test and evaluation community. Properly scoped tests should lead to more rigorous evaluations, which, in turn, should lead to well-informed acquisition decisions.

References

Casella, G., and R. L. Berger. 2002. *Statistical Inference*. 2nd Edition. Pacific Grove, CA: Duxbury.

Freeman, L. J., A. G. Ryan, J. L. Kensler, R. M. Dickinson, and G. G. Vining. 2013. "A Tutorial on the Planning of Experiments." *Quality Engineering* 25: 315–332.

Hahn, G. J., W. Q. Meeker Jr., and P. I. Feder. 1976. "The Evaluation and Comparison of Experimental Designs for Fitting Regression Relationships." *Journal of Quality Technology* 8: 140–157. https://doi.org/10.1080/00224065.1976.11980735.

Johnson, R. T., G. T. Hutto, J. R. Simpson, and D. C. Montgomery. 2012. "Designed Experiments for the Defense Community." *Quality Engineering* 24: 60–79. https://doi.org/10.1080/08982112.2012.627288.

McCullagh, P., and J. A. Nelder. 1972. *Generalized Linear Models*. 2nd edition. Chapman and Hall, Inc.

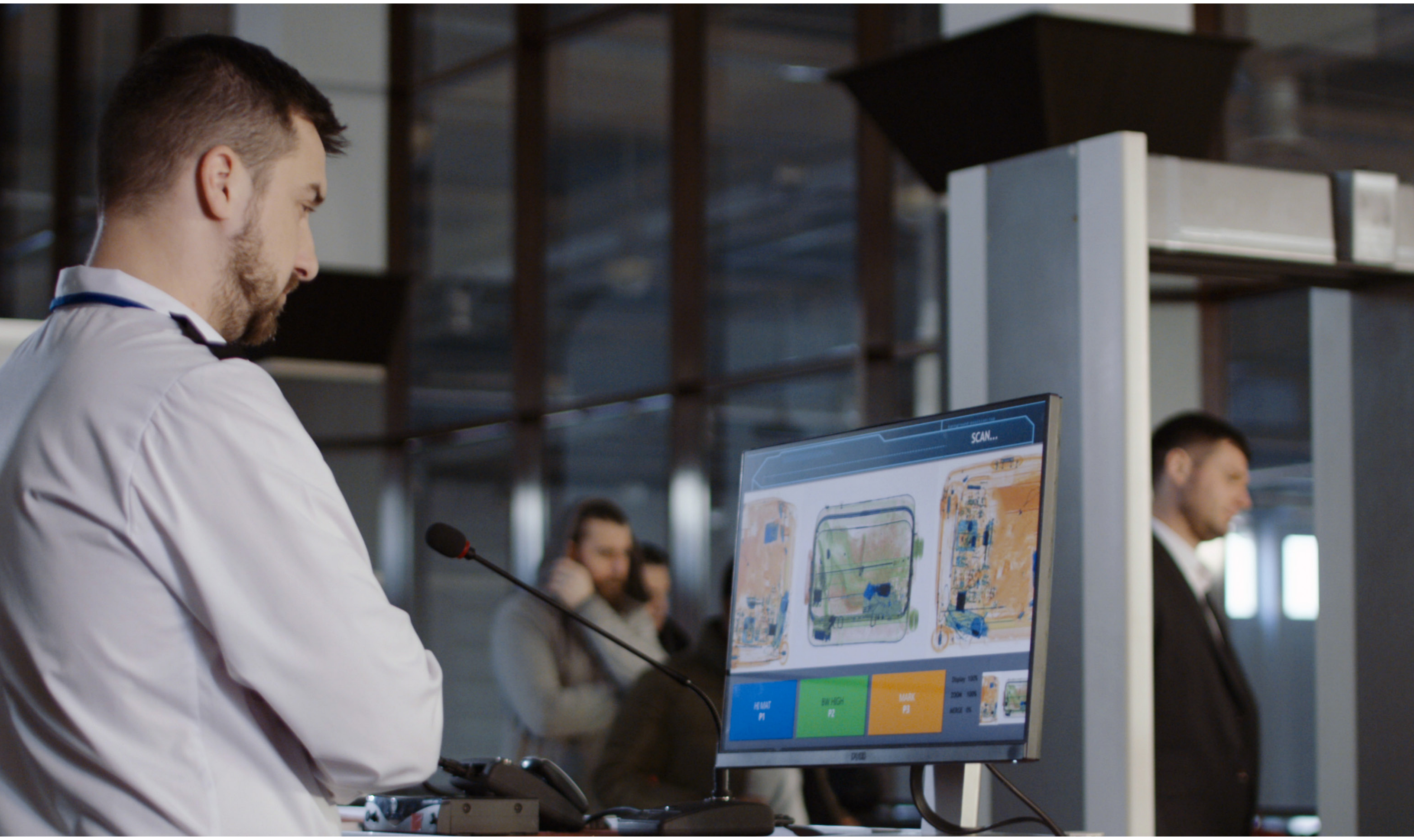
Montgomery, D. C. 2008. *Design and Analysis of Experiments*. 7th edition. John Wiley & Sons.

Colin Anderson (left), a Research Staff Member in the Operational Evaluation Division (OED) of IDA's Systems and Analyses Center, holds a doctorate in physics from Yale University. *Laura Freeman*, former Assistant Director in OED, holds a doctorate in statistics from Virginia Tech. *Tom Johnson* (right), a Research Staff Member in OED, holds a doctorate in aerospace engineering from Old Dominion University. *Jim Simpson* (no photo available), a consultant in OED, holds a doctorate in industrial engineering from Arizona State University.

The Threat Detection System That Cried Wolf: Reconciling Developers with Operators[1]

Shelley M. Cazares

Both the Department of Defense (DoD) and the Department of Homeland Security (DHS) use threat detection systems, such as airplane cargo screeners and counter–improvised-explosive-device (IED) systems. These systems may perform well during testing but "cry wolf" in the field (i.e., generate false alarms when true threats are not present). As a result, operators can lose faith in the systems—ignoring them or even turning them off and taking the chance that a true threat will not occur. This paper reviews statistical concepts to reconcile the performance metrics that summarize a developer's view of a system during testing with the metrics that describe an operator's view of the system during real-world missions. Program managers can still make use of systems that cry wolf by arranging them into a tiered system that performs better than each individual system alone.



[1] The original article of the same title was published in *Defense Acquisition Research Journal*, January 2017, https://doi.org/10.22594/dau.16-749.24.01. The original article illustrates how a PM can make use of a system that frequently cries wolf by incorporating it into a tiered system that, overall, exhibits better performance than each individual system does alone.

Introduction

DoD and DHS operate counter-mine systems, counter-IED systems, airplane cargo screening systems, and other threat detection systems, all of which share a common purpose: to detect potential threats among clutter.

Threat detection systems are often assessed based on their Probability of Detection (P_d) and Probability of False Alarm (P_{fa}) (Urkowitz 1967). P_d describes the fraction of true threats for which the system correctly declares an alarm. Conversely, P_{fa} describes the fraction of true clutter (true nonthreats) for which the system *in*correctly declares an alarm—a false alarm. A perfect system will exhibit a P_d of 1 and a P_{fa} of 0. P_d and P_{fa} are defined in Table 1.

While the Probability of Detection and the Probability of False Alarm summarize how much of the truth causes an alarm, Positive Predictive Value and Negative Predictive Value summarize how many alarms turn out to be true.

Table 1. Definitions of Common Metrics Used to Assess the Performance of Threat Detection Systems

Metric	Definition	Perspective
Probability of Detection (P_d)	The fraction of all items containing a true threat for which the system correctly declared an alarm	Developer
Probability of False Alarm (P_{fa})	The fraction of all items not containing a true threat for which the system incorrectly declared an alarm	Developer
Positive Predictive Value (*PPV*)	The fraction of all items causing an alarm that did end up containing a true threat	Operator
Negative Predictive Value (*NPV*)	The fraction of all items not causing an alarm that did not end up containing a true threat	Operator
Prevalence (*Prev*)	The fraction of items that contained a true threat (regardless of whether the system declared an alarm)	Not applicable

Threat detection systems with good P_d and P_{fa} performance metrics are not always well received by system operators, because some systems may "cry wolf," generating false alarms when true threats are not present. As a result, operators may lose faith in the systems, delaying their response to alarms (Getty et al. 1995) or ignoring them altogether (Bliss et al. 1995), potentially leading to disastrous consequences. This issue has arisen in military, national security, and civilian scenarios (Cushman 1987; Stuart 1987; Oldham 2006).

This issue often stems from an inappropriate choice of metrics—P_d and P_{fa}—used to assess the system's performance during testing. While P_d and P_{fa}

encapsulate the *developer's* perspective of the system's performance, these metrics do not encapsulate the *operator's* perspective. The operator's view can be better summarized with other metrics, namely Positive Predictive Value (*PPV*) and Negative Predictive Value (*NPV*) (Altman and Bland 1994). *PPV* describes the fraction of all alarms that correctly turn out to be true threats—a measure of how often the system does not cry wolf. Similarly, *NPV* describes the fraction of all *lack* of alarms that correctly turn out to be true clutter. From the operator's perspective, a perfect system will have *PPV* and *NPV* values equal to 1. *PPV* and *NPV* are also defined in Table 1.

Interestingly enough, the same threat detection system that satisfies the developer's desire to detect as much truth as possible can also disappoint the operator by crying wolf too often (Scheaffer and McClave 1995). A system can exhibit excellent P_d and P_{fa} values, while also exhibiting a poor *PPV* value. Unfortunately, low *PPV* values naturally occur when the Prevalence (*Prev*) of true threat among true clutter is extremely low (Parasuraman 1997; Scheaffer and McClave 1995), as is often the case in defense and homeland security scenarios. As summarized in Table 1, *Prev* is a measure of how widespread or common the true threat is. A *Prev* of 1 indicates a true threat is always present, while a *Prev* of 0 indicates a true threat is never present. As we shall see, a low *Prev* can lead to a discrepancy in how developers and operators view the performance of threat detection systems in DoD and DHS.

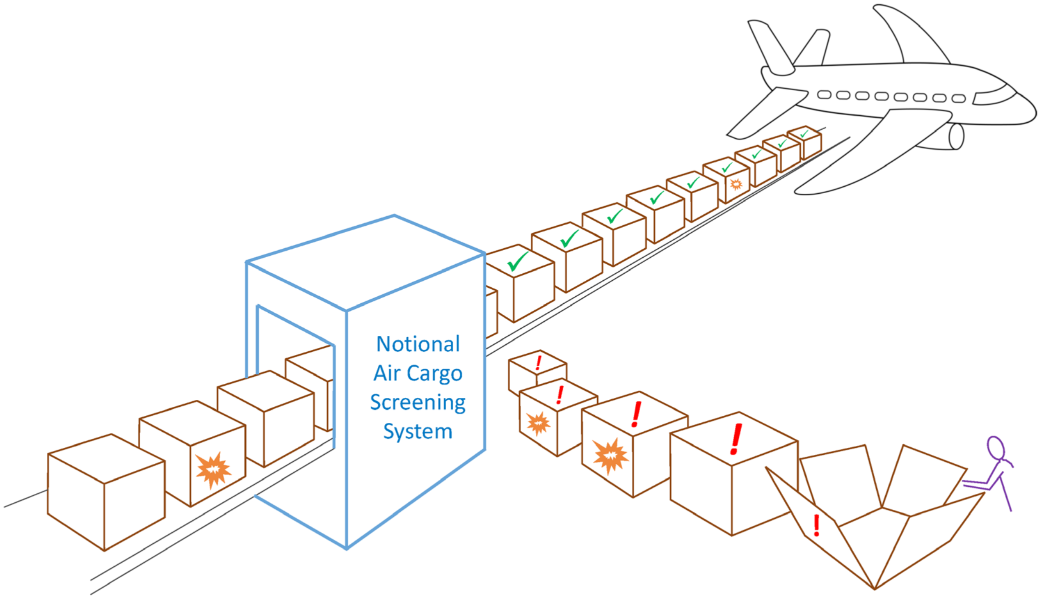
In the following sections, I reconcile the performance metrics used to quantify the developer's versus operator's views of threat detection systems. Although these concepts are already well known within the statistics and human factors communities, they are not often immediately understood in DoD and DHS science and technology acquisition communities. This review is intended for program managers (PMs) of threat detection systems in DoD and DHS.

Testing a Threat Detection System

Consider the notional air cargo screening system in Figure 1. The purpose of this notional system is to detect explosive threats packed inside items that are about to be loaded into the cargo hold of an airplane. To determine how well this system meets capability requirements, its performance must be quantified. A large number of items are input into the system, and each item's ground truth (whether the item contained a true threat) is compared to the system's output (whether the system declared an alarm). The items represent those that the system would likely encounter in an operational setting. At the end of the test, the following items are counted:

- True Positive (*TP*), an item containing a true threat for which the system correctly declared an alarm;
- False Positive (*FP*), an item *not* containing a true threat for which the system *in*correctly declared an alarm (a Type I error);

- False Negative (*FN*), an item containing true threat for which the system *in*correctly did *not* declare an alarm (a Type II error); and
- True Negative (*TN*), an item *not* containing a true threat for which the system correctly did *not* declare an alarm.



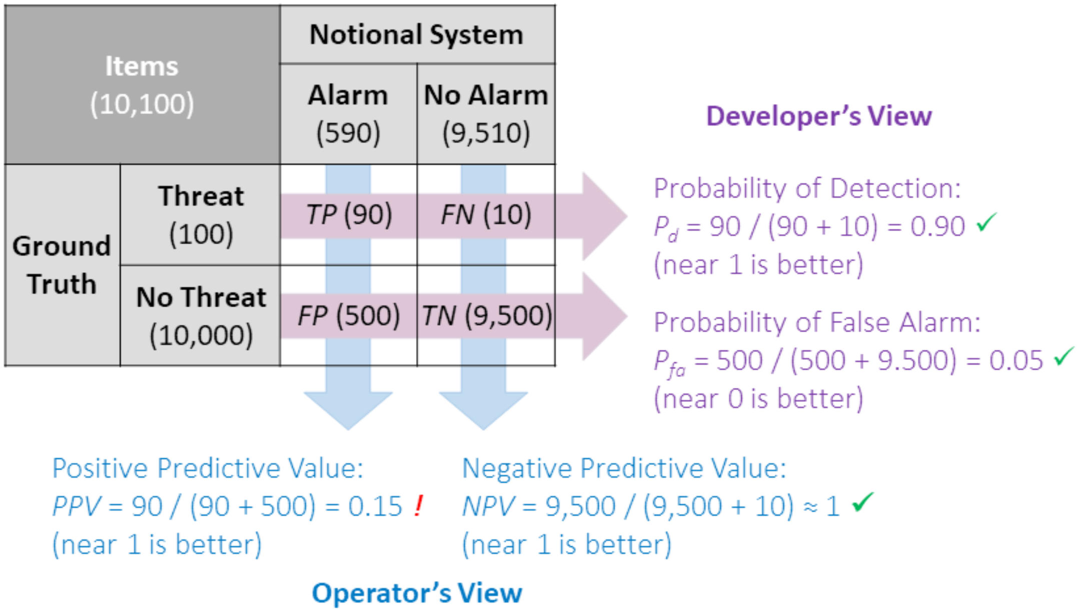
Note: A set of predefined, discrete items (small brown boxes) are presented to the system one at a time. Some items contain a true threat (orange star) among clutter, while other items contain clutter only (no orange star). For each item, the system declares either one or zero alarms. All items for which the system declares an alarm (red exclamation point) are further examined manually by trained personnel (purple figure). In contrast, all items for which the system does not declare an alarm (green checkmark) are left unexamined and loaded directly onto the airplane.

Figure 1. Notional Air Cargo Screening System

As shown in Figure 2, a total of 10,100 items passed through the notional air cargo screening system. One hundred items contained a true threat, while 10,000 items did not. The system declared an alarm for 590 items and did not declare an alarm for 9,510 items. Comparing the items' ground truth to the system's alarms (or lack thereof), there were 90 *TP*s, 10 *FN*s, 500 *FP*s, and 9,500 *TN*s.

Developer's View: P_d and P_{fa}

A PM must consider how much of the truth the threat detection system is able to identify. This can be done by considering two questions: Of those items that contain a true threat, for what fraction does the system correctly declare an alarm? And of those items that do *not* contain a true threat, for what fraction does the system *in*correctly declare an alarm? These questions often guide developers during the research and development phase of a threat detection system.



Items (10,100)		Notional System	
		Alarm (590)	No Alarm (9,510)
Ground Truth	Threat (100)	*TP* (90)	*FN* (10)
	No Threat (10,000)	*FP* (500)	*TN* (9,500)

Developer's View

Probability of Detection:
P_d = 90 / (90 + 10) = 0.90 ✓
(near 1 is better)

Probability of False Alarm:
P_{fa} = 500 / (500 + 9.500) = 0.05 ✓
(near 0 is better)

Positive Predictive Value:
PPV = 90 / (90 + 500) = 0.15 !
(near 1 is better)

Negative Predictive Value:
NPV = 9,500 / (9,500 + 10) ≈ 1 ✓
(near 1 is better)

Operator's View

Note: This 2 × 2 matrix tabulates the number of *TP*, *FN*, *FP*, and *TN* items processed by the system. P_d and P_{fa} summarize the developers' view of the system's performance, while *PPV* and *NPV* summarize the operators' view. In this notional example, the low *PPV* of 0.15 indicates a poor operator experience (the system often cries wolf, since only 15 percent of alarms turn out to be true threats) even though the good P_d and P_{fa} are well received by developers.

Figure 2. 2 × 2 Confusion Matrix of a Notional Air Cargo Screening System

P_d and P_{fa} can be easily calculated from the confusion matrix to answer these questions. From a developer's perspective, the notional air cargo screening system exhibits good performance:[2]

$$P_d = \frac{TP}{TP + FN} = \frac{90}{90 + 10} = 0.90 \text{ (compared to 1 for a perfect system)} \qquad (1)$$

$$P_{fa} = \frac{FP}{FP + TN} = \frac{500}{500 + 9{,}500} = 0.05 \text{ (compared to 0 for a perfect system).} \qquad (2)$$

Equation 1 shows that, of all items that contained a true threat ($TP + FN = 90 + 10 = 100$), a large subset ($TP = 90$) correctly caused an alarm. These counts resulted in $P_d = 0.90$, close to the value of 1 that would be exhibited by a perfect

[2] PMs must determine what constitutes a "good" performance. For some systems operating in some scenarios, $P_d = 0.90$ is considered good, since only 10 FNs out of 100 true threats is considered an acceptable risk. In other cases, $P_d = 0.90$ is not acceptable. Appropriately setting a system's capability requirements calls for a frank assessment of the likelihood and consequences of *FN*s versus *FP*s and is beyond the scope of this paper.

system.[3] Based on this P_d value, the PM can conclude that 90 percent of items that contained a true threat correctly caused an alarm, which may (or may not) be considered acceptable within the capability requirements for the system. Furthermore, Equation 2 shows that, of all items that did *not* contain a true threat ($FP + TN = 500 + 9{,}500 = 10{,}000$), only a small subset ($FP = 500$) caused a false alarm. These counts led to $P_{fa} = 0.05$, close to the 0 value that would be exhibited by a perfect system. In other words, only 5 percent of items that did *not* contain a true threat caused a false alarm.

Operator's View: *PPV* and *NPV*

The PM must also anticipate the operator's view of the threat detection system. One way to do this is to answer the following questions: Of those items that caused an alarm, what fraction turned out to contain a true threat (i.e., what fraction of alarms turned out *not* to be false)? And of those items that did *not* cause an alarm, what fraction turned out *not* to contain a true threat? On the surface, these questions seem similar to those posed previously for P_d and P_{fa}. Upon closer examination, however, they are quite different. While P_d and P_{fa} summarize how much of the truth causes an alarm, *PPV* and *NPV* summarize how many alarms turn out to be true.

PPV and *NPV* can also be easily calculated from the 2 × 2 confusion matrix. From an operator's perspective, our notional air cargo screening system exhibits a conflicting performance:

$$NPV = \frac{TN}{TN + FN} = \frac{9{,}500}{9{,}500 + 10} \approx 1 \text{ (compared to 1 for a perfect system)} \tag{3}$$

$$PPV = \frac{TP}{TP + FP} = \frac{90}{90 + 500} = 0.15 \text{ (compared to 1 for a perfect system)} \tag{4}$$

Equation 3 shows that, of all items that did *not* cause an alarm ($TN + FN = 9{,}500 + 10 = 9{,}510$), a large subset ($TN = 9{,}500$) correctly turned out to *not* contain a true threat. These counts resulted in $NPV \approx 1$, approximately equal to the 1 value that would be exhibited by a perfect system.[4] In the absence of an alarm, the operator could rest assured that a threat was highly unlikely. However, Equation 4 shows that, of all items that did indeed cause an alarm ($TP + FP = 90 + 500 = 590$), only a small subset ($TP = 90$) turned out to contain a true threat (i.e., were not false alarms). These counts unfortunately led to $PPV = 0.15$, much lower than the 1 value that would be exhibited by a perfect system. When an alarm was declared, the operator could not trust that a threat was present, since the system cried wolf so often.

[3] For P_d and P_{fa} values from equations (1) and (2), statistical tests can determine whether the system's value is significantly different from the perfect value and if it is different from the capability requirement (Fleiss et al. 2013).

[4] For *NPV* and *PPV* values from equations (3) and (4), statistical tests can determine whether the system's value is significantly different from the perfect value and if it is different from the capability requirement (Fleiss et al. 2013.

Reconciling Developers with Operators: P_d and P_{fa} versus *PPV* and *NPV*

The discrepancy between *PPV* and *NPV* versus P_d and P_{fa} reflects the discrepancy between operators' and developers' views of the threat detection system. Developers are often primarily interested in how much of the truth correctly cause alarms—concepts quantified by P_d and P_{fa}. In contrast, operators are often primarily concerned with how many alarms turn out to be true—concepts quantified by *PPV* and *NPV*. As shown in Figure 2, the very same system that exhibits excellent values for P_d, P_{fa}, and *NPV* can also exhibit poor values for *PPV*.

Poor *PPV* values can be expected for DoD and DHS threat detection systems. Such performance is often merely a reflection of the low *Prev* of true threats among true clutter that commonly occurs in defense and homeland security scenarios.[5] *Prev* describes the fraction of all items that contain a true threat, including those that did and did not cause an alarm. In the case of our notional air cargo screening system, *Prev* is very low:

$$Prev = \frac{TP + FN}{TP + FN + FP + TN} = \frac{90 + 10}{90 + 10 + 500 + 9{,}500} = 0.01. \tag{5}$$

Equation 5 shows that, of all items ($TP + FN + FP + TN = 90 + 10 + 500 + 9{,}500 = 10{,}100$), only a small subset ($TP + FN = 90 + 10 = 100$) contained a true threat, leading to $Prev = 0.01$. When true threats are rare, most alarms turn out to be false, even for an otherwise strong threat detection system, leading to a low value for *PPV*. In fact, to achieve a high value of *PPV* when *Prev* is extremely low, a threat detection system must exhibit so few *FP*s (false alarms) as to make P_{fa} approximately zero.

Recognizing this phenomenon, PMs should not necessarily dismiss a threat detection system simply because it exhibits a poor *PPV*, provided that it also exhibits an excellent P_d and P_{fa}. Instead, PMs can estimate *Prev* to help determine how to guide such a system through development. *Prev* does not depend on the threat detection system and can, in fact, be calculated in the absence of the system. Knowledge of ground truth (i.e., which items contain a true threat) is all that is needed to calculate *Prev* (Scheaffer and McClave 1995).

Of course, ground truth is not known *a priori* in an operational setting. However, it may be possible for PMs to use historical data or intelligence tips to roughly estimate whether *Prev* is likely to be particularly low in operation. A *Prev* that is estimated to be particularly low can cue the PM to anticipate discrepancies in P_d and P_{fa} versus *PPV*, forecasting the inevitable discrepancy between the developers' versus operators' views early in the system's

[5] Conversely, when *Prev* is *high*, threat detection systems often exhibit poor values for *NPV*, even while exhibiting excellent values for P_d, P_{fa}, and *PPV*. Such cases are not discussed here, since fewer scenarios in DoD and DHS involve a *high* prevalence of threat among clutter.

development, while there are still time and opportunity to make adjustments. At that point, the PM can identify concepts of operations in which the system can still provide value to the operator for his or her mission. A tiered system may provide one such opportunity.

References

Altman, D. G., and J. M. Bland. 1994. "Diagnostic Tests 2: Predictive Values." *British Medical Journal* 309, no. 6947: 102. https://doi.org/10.1136/bmj.309.6947.102.

Bliss, J. P., R. D. Gilson, and J. E. Deaton. 1995. "Human Probability Matching Behavior in Response to Alarms of Varying Reliability." *Ergonomics* 38, no. 11: 2300–2312. https://doi.org/10.1080/00140139508925269.

Cushman, J. H. 1987. "Making Arms Fighting Men Can Use." *New York Times*. June 21. http://www.nytimes.com/1987/06/21/business/making-arms-fighting-men-can-use.html.

Fleiss, J. L., B. Levin, and M. C. Paik. 2013. *Statistical Methods for Rates and Proportions* 3rd ed. Hoboken, NJ: John Wiley.

Getty, D. J., J. A. Swets, R. M. Pickett, and D. Gonthier. 1995. "System Operator Response to Warnings of Danger: A Laboratory Investigation of the Effects of the Predictive Value of a Warning on Human Response Time." *Journal of Experimental Psychology: Applied* 1, no. 1: 19–33. https://doi.org/10.1037/1076-898X.1.1.19.

Oldham, J. 2006. "Outages Highlight Internal FAA Rift." *Los Angeles Times*. October 3. http://articles.latimes.com/2006/oct/03/local/me-faa3.

Parasuraman, R. 1997. "Humans and Automation: Use, Misuse, Disuse, Abuse." *Human Factors* 39, no. 2: 230–253. https://doi.org/10.1518/001872097778543886.

Scheaffer, R. L., and J. T. McClave. 1995. "Conditional Probability and Independence: Narrowing the Table." In *Probability and Statistics for Engineers*, 85–92. 4th ed. Belmont, CA: Duxbury Press.

Stuart, R. 1987. "U.S. Cites Amtrak for Not Conducting Drug Tests."' *The New York Times*. January 8. http://www.nytimes.com/1987/01/08/us/us-cites-amtrak-for-not-conducting-drug-tests.html.

Urkowitz, H. 1967. "Energy Detection of Unknown Deterministic Signals." *Proceedings of the IEEE* 55, no. 4:523–531. MV{.

Shelley Cazares, a Research Staff Member in the Science and Technology Division of IDA's Systems and Analyses Center, holds a doctorate in engineering science from the University of Oxford.

Winning Indefinite Conflicts: Achieving Strategic Success against Ideologically Motivated Violent Non-State Actors[1]

Mark E. Vinson

In Afghanistan, Iraq, Syria, and many other countries around the globe, violent non-state actors, motivated by religious, political, and ethnic ideas, have been remarkably resilient, perseverant, and influential. With broad, ambiguous strategic objectives, and an indefinite, changing path to strategic success, the United States has struggled to define, much less achieve, strategic success. Traditional military victory in such conflicts is not sufficient against an ideology-based movement. Rather, the military must support a holistic strategy that defeats the ideology with a better idea. This article, based on a cooperative study by the U.S. and Israeli militaries, examines this challenge and offers ideas and recommendations on how the United States might address it. Strategic success against such actors requires a long-term, comprehensive, and indirect approach with the United States serving as a patron to encourage and support regional partners, who in turn directly enable local partners to holistically address their local populations' basic needs in terms of security, legitimate governance, and sustainable services. The better idea that will enable strategic success in countering or defeating an ideology-driven violent non-state actor (VNSA) must be formed and legitimized by tangible actions and measured by concrete results at the local level.

[1] The original article, based on results of a cooperative examination by the U.S. and Israeli militaries, was published in *Small Wars Journa*l, March 2017, http://smallwarsjournal.com/jrnl/art/winning-indefinite-conflicts-achieving-strategic-success-against-ideologically-motivated-vi.

Elusive Success

If, as President Obama asserted in July 6, 2015, ideologies are defeated not by guns, but by better ideas (White House 2015), then how should the U.S. military be used to help achieve strategic success in the growing number of protracted, irregular conflicts with ideologically motivated VNSAs? In Afghanistan, Iraq, Syria, Yemen, Somalia, and many more countries around the globe, VNSAs, motivated by religious, political, ethnic, and other status-quo–challenging ideas, have been remarkably resilient, perseverant, and influential. By surviving and rapidly recovering from punishing attacks by the United States and its partners—while continuing to carry out violent agendas against local, regional, and even global adversaries—these VNSAs can credibly claim that they are succeeding strategically. With broad, ambiguous long-term strategic objectives, and an open-ended, evolving path to strategic success, the United States has generally conducted limited military operations intended to disrupt and degrade such VNSAs, followed by the hopeful but indefinite objective of ultimately defeating them. In view of the VNSAs' resilience, persistence, and ideological basis for conflict, the path to strategic success for the United States has remained elusive.

The U.S. military inevitably enters conflicts with a lack of local knowledge, language abilities, and cultural experience. Planners struggle to accurately understand and frame the operational problem, leading to flawed campaign design and planning.

Accelerating Treadmill

U.S. intelligence capabilities are ill suited for irregular conflicts with VNSAs, which tend to take place in complex, uncertain foreign operational environments. These environments are dynamic ecosystems containing a multitude of actors, each with unique tribal, religious, national, and ethnic identities that produce complex relationships based on myriad factors, all of which combine to make it impossible to predict system-wide effects of an action against any part of the system. In such unfamiliar environments, threat actors are conducting protracted, ideological conflicts, blending into populations, urban areas, and complex terrain. The U.S. military inevitably enters conflicts with a lack of local knowledge, language abilities, and cultural experience. Planners struggle to accurately understand and frame the operational problem, leading to flawed campaign design and planning.

The U.S. military generally lacks the essential support, both among the local population in a conflict zone and at home, to sustain its direct involvement in a protracted conflict with VNSAs. Local populations will naturally distrust the motives and long-term commitment of external forces, especially extra-regional forces with no tie to the local land or its people. As a foreign force in such conflicts, the U.S. military will naturally struggle to gain and maintain the local legitimacy required for successful direct involvement in a protracted campaign. Likewise, the sustained support of the U.S. public for direct involvement in such conflicts is unlikely unless political leaders can communicate a clear and compelling argument for U.S. interests. The protracted nature of conflicts with VNSAs, the huge cost of military operations, and the public's reluctance to accept casualties, make the substantial and long-term commitment of ground combat forces problematic for the United States.

In the face of complex and uncertain conflicts, U.S. leaders are challenged to describe specific long-term strategic objectives that align with those of U.S. partners. As a result, leaders initially provide broad, ambiguous objectives that may be insufficient to enable national or coalition unity of effort. Without specific strategic objectives, it is unclear whether U.S. operations are making progress toward strategic success.

Besides ambiguous strategic objectives, military operations against VNSAs generally suffer from a lack of effective strategic and operational orchestration. As a result, a series of tactically or operationally successful operations may not be integrated with interagency or other partners' lines of effort, and they may not contribute to strategic success. Without clear strategic objectives that find common ground with partners' various and competing objectives, U.S. operational planning will be unable to establish the integrating framework necessary to unify effort among all contributing actors.

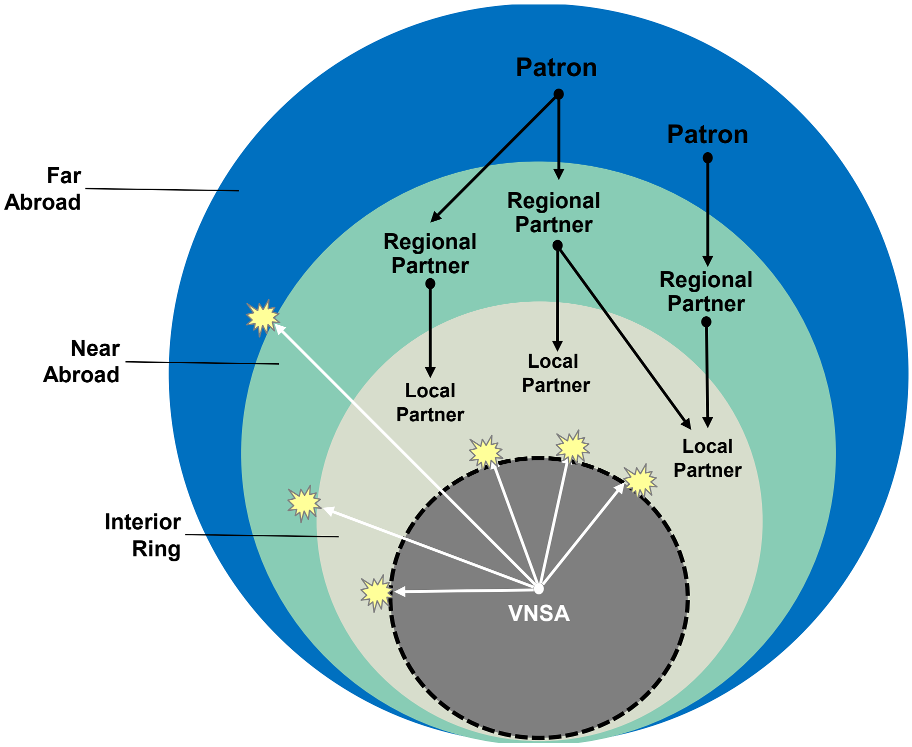
In 2005, while the United States was struggling to design and execute successful campaigns against VNSAs in Iraq and Afghanistan, Douglas Feith, then Under Secretary of Defense for Policy, warned that if the nation's efforts were limited to "protecting the homeland and attacking and disrupting terrorist networks, you're on a treadmill that is likely to get faster and faster…." (Schmitt and Shanker 2005). Twelve years later, the United States is arguably still on the accelerating treadmill, asking what strategic success looks like against such adversaries, what its role should be, and how its military should be used.

Comprehensive and Indirect Approach

Fundamentally, political leaders should not conflate military success with strategic success, particularly in complex conflicts with VNSAs. Although military and police operations play a critical security and stability role in a comprehensive approach, their contributions cannot be strategically decisive. Addressing the foundations of a conflict with VNSAs requires a tailored, integrated, strategic approach that comprehensively applies all elements of national and coalition partner power.

To enable a sustainable, long-term campaign that gains and maintains public support, the U.S. military must employ an indirect approach. This approach requires a sustainable patron–regional partner–local partner relationship that will enable a long-term campaign to succeed (Figure 1). To enable such a partnership, trust and cooperation based on an alignment of strategic objectives regarding the VNSA adversary must be sustained. The key ideas behind the indirect approach result from two complementary concepts: a top-down go-local concept and a bottom-up grassroots concept. As an external patron, the

United States goes local by encouraging and supporting regional partner states with a direct stake in the conflict and historical ties to the vulnerable territory and its local populations, who, in turn, encourage and enable local actors to be committed partners that holistically address their populations' needs. This means that vetted local partners—who are intrinsically committed to and inherently knowledgeable of the local population's needs—must be identified and enabled with sustainable support during a protracted conflict. In turn, the empowered local actors use a bottom-up grassroots approach to establish local security, legitimate governance, economic opportunity, and sustainable services, tailored to their constituent populations.



Notes: ISIS is framing a strategy to expand across three geographic rings referred to as interior, near abroad, and far abroad. Patrons in the far abroad ring enable local partners' success through regional partners in the near abroad ring. Regional partners enable sustained support to local partners in the interior ring. Local partners contain the VNSA, establish local security, oversee legitimate governance, and spur economic development.

Figure 1. Indirect Approach Model to Enable Regional and Local Partners

The primary conditions for strategic and operational success are security, legitimacy, and sustainability. Trained by regional partners, and equipped, supported, and coordinated by external patrons, local police and militia forces establish and maintain security. Likewise, local leaders are best suited to establish legitimate governance of local population groups. Local leaders have the obvious and essential advantage of intrinsically understanding the governance and other basic needs—security, economic, social, services—of their constituents. If legitimacy is a result of their success in addressing the population's needs, then local leaders have the best opportunity to gain and maintain the population's legitimacy and support. Local leaders are directly enabled by regional partners, who leverage their historical relationships with the local populations to gain trust and legitimate influence, while external patrons with international legitimacy and influence indirectly support them through their regional partners. Finally and critically, a campaign is sustainable when each actor (local, regional, and external), in consideration of its interests and likely long-term levels of public and political support, commits time, manpower, and resources to achieve its objectives.

Comprehensive Containment and a Better Idea

Harleen Gambhir summarized ISIS's strategy, writing that "ISIS intends to expand its Caliphate and eventually incite a global apocalyptic war. In order to do so, ISIS is framing a strategy to remain and expand across three geographic rings: the Interior Ring, the Near Abroad, and the Far Abroad" (Gambhir 2015, 9). How would a comprehensive and indirect approach be applied to contain such a threat?

Containment operations include complementary military and civilian lines of effort to build and manage a coalition, to halt VNSA territorial expansion, to prevent VNSA recruits from entering a regional partner's territory, to support local governance and economic opportunity, and to deny VNSA access to weapons, funds, and resources. A containment operation is a defensive approach unlikely to be decisive on its own, but it could provide a stable basis for follow-on offensive operations. Thus, containment should be considered as an intermediate objective in a broader campaign designed to ultimately succeed operationally against a VNSA. Such an operation might be employed early in a campaign to prevent expansion and to stabilize and protect vulnerable regional and local partners.

While territorially containing a threat is essential, the idea must be extended beyond the physical to comprehensively contain the influence of VNSAs that embody and promote violent ideologies. As James Dorsey observed, "[c]ontainment addresses the immediate problem but ignores factors that fuel radicalization far from the warring state's borders and make jihadism attractive to the disaffected across the globe" (Dorsey 2015). Addressing the spread of violent ideas and associated violent acts requires a different approach. This challenge returns us to President Obama's statement, which begs the practical question of how can a better idea be applied to defeat a violent, ideologically-motivated VNSA, or more specifically, to attain the key conditions of security, legitimacy, and sustainability?

Better ideas are more than information operations or persuasive philosophies; better ideas require a fusion of compelling messages and congruent actions. To counter or defeat an ideology-driven VNSA, better ideas must be formed and legitimized by tangible actions and measured by concrete results. These ideas and actions must address the fundamental issues that produced and supported the VNSA, and they must be tailored to achieve the key conditions of security, legitimacy, and sustainability for each relevant local population. Only then, will the United States and its regional and local partners demonstrate the idea's credibility, the integrity of which can then be used to influence other relevant populations and to proliferate the idea. As the idea is successfully implemented, using the indirect approach described earlier, it could then be spread incrementally via a cellular approach that first establishes an outer defensive containment ring of local security forces that consolidates their gains by establishing legitimate governance and sustainable services. As the containment ring succeeds, the idea and supporting actions could be extended to contract the VNSA territory and counter the credibility of its ideology, ultimately to achieve the necessary security, legitimacy, and sustainability conditions.

Implications for the U.S. Military

While the U.S. military needs to be able to fight and win major wars, it also needs the ready capabilities and capacity to sustain and eventually achieve strategic success in long-term campaigns against VNSAs.

To improve its ability to achieve strategic success in such conflicts, the military first needs improved intelligence capabilities to better understand local and regional populations, to assess root-cause issues, and to enable effective campaign design, planning, execution, and assessment. The U.S. military should consider developing more tailorable command and control capabilities to better enable a unified planning and execution effort with U.S. Government agencies, and across a broad coalition of patron states, regional partners, and myriad local partners.

To sustain its support to partners, the U.S. military requires sufficient regionally focused personnel with language and cultural training to rotate forces and sustain trusting relationships for the duration of a long-term campaign. Given the specialized nature of U.S. enabling operations, the military needs special operations forces and other high-demand forces that can directly engage with partners. They must have the language skills and cultural knowledge to adequately understand the situation, and to gain and maintain influence. While special operations forces are best suited for these roles, many of the traditional intelligence, communications, joint fires, and logistics support functions reside in the conventional forces. Likewise, in view of persistent, region-wide conflicts, the military requires the capability to rapidly and effectively organize, train, and deploy conventional forces to expand its special operations forces' capabilities and capacity without breaking the conventional force.

References

Dorsey, J. M. 2015. "Fighting Islamic State: Getting Down to Root Causes." *International Policy Digest*. October2. https://intpolicydigest.org/2015/10/02/fighting-islamic-state-getting-down-to-root-causes/.

Gambhir, H. 2015. *ISIS'S Global Strategy*: A Wargame. Washington, DC: Institute for the Study of War. Middle East Security Report 28, July.

Schmitt, E., and T. Shanker. 2005. "U.S. Officials Retool Slogan for Terror War." *New York Times*. July 26. https://www.nytimes.com/2005/07/26/politics/us-officials-retool-slogan-for-terror-war.html.

White House. 2015. "Remarks by the President on Progress in the Fight against ISIL." July 6. https://obamawhitehouse.archives.gov/the-press-office/2015/07/06/remarks-president-progress-fight-against-isil.

يعرض نفسه للخطر
MORTAL DANGER - MILITARY ZONE
ANY PERSON WHO PASSES OF DAMAGES THE FENCE
ENDANGERS HIS LIFE

Mark Vinson, an Adjunct Research Staff Member in the Joint Advanced Warfighting Division of IDA's Systems and Analyses Center, holds a master of science degree in operations research from the Georgia Institute of Technology. He also is certified in military operational arts and science by the U.S. Army Command and General Staff College and in national security and strategic studies by the U.S. Army War College.

Past Welch Award Winners

2017

"Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review"

Review of Educational Research

J. A. Kulik and J. D. Fletcher

2016

"Mining Measured Information from Text"

Proceedings of the 38th International ACM CIGR Conference on Research and Develoment in Information Retrieval

A. Maiya, D. Visser, and A. Wan

2015

"Visible Signatures of Hypersonic Reentry"

Journal of Spacecraft and Rockets

J. Teichman and L. Hirsch

2014

"Mixed Models Analysis of Radar Residuals Data"

IEEE Access

C. Gaither, D. Loper, C. Jackson, and J. Pozderac

2013

"Comparison of Predicted and Measured Multipath Impulse Responses"

IEEE Transactions on Aerospace and Electronic Systems

K. Haspert and M. Tuley

Past Issues

2018 | Multidisciplinary Research for Securing the Homeland

- Countering Terrorism One Technology at a Time
- Does Imposing Consequences Deter Attempted Illegal Entry into the United States
- Improving Shared Understanding of National Security and Emergency Preparedness Communications
- and more...

2016 | Acquisition, Part 2

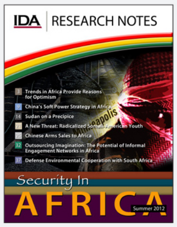
- Cost Growth, Acquisition Policy, and Budget Climate
- Improving Predictive Value of Indicators of Poor Performance
- Root Cause Analysis of VTUAV Fire Scout's Nunn-McCurdy Breach
- and more...

2014 | Technological Innovation for National Security

- Countering Terrorism One Technology at a Time
- Does Imposing Consequences Deter Attempted Illegal Entry into the United States
- Improving Shared Understanding of National Security and Emergency Preparedness Communications
- and more...

2013 | Acquisition, Part 1

- Defining Acquisition Trade Space Through "DERIVE"
- Supporting Acquisition Decisions in Air Mobility
- Assessing Reliability with Limited Flight Testing
- and more...



2012 | Security in Africa

- Trends in Africa Provide Reasons for Optimism
- China's Soft Power Strategy in Africa
- Sudan on a Precipice
- and more...

IDA | RESEARCH NOTES

© Institute for Defense Analyses
4850 Mark Center Drive • Alexandria, VA 22311-1882
ida.org

@IDA_org