Calhoun: The NPS Institutional Archive

DSpace Repository

Theses and Dissertations | 1. Thesis and Dissertation Collection, all items

2019-12

MEASURING THE EFFECTIVENESS OF SURVEILLANCE TECHNOLOGY AT THE U.S. SOUTHERN BORDER

Hudspeth, Robert A.

Monterey, CA; Naval Postgraduate School

http://hdl.handle.net/10945/64187

Downloaded from NPS Archive: Calhoun

http://www.nps.edu/library

Calhoun is a project of the Dudley Knox Library at NPS, furthering the precepts and goals of open government and government transparency. All information contained herein has been approved for release by the NPS Public Affairs Officer.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

MEASURING THE EFFECTIVENESS OF SURVEILLANCE TECHNOLOGY AT THE U.S. SOUTHERN BORDER

by

Robert A. Hudspeth

December 2019

Thesis Advisor:	Erik J. Dahl
Second Reader:	Scott E. Jasper

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** December 2019	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE MEASURING THE EFFECTIVENESS OF SURVEILLANCE TECHNOLOGY AT THE U.S. SOUTHERN BORDER	**5. FUNDING NUMBERS**
6. AUTHOR(S) Robert A. Hudspeth	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

The United States' investment in southern border security has consistently been a topic of discussion regarding technological improvements and measurements of effectiveness. There have been multiple failed programs designed to combine infrastructure, personnel, and technology, ranging from the America's Shield Initiative (ASI) to the Secure Borders Initiative Network (SBInet). These efforts have resulted in billions of dollars of wasted funding. The latest initiative, named the Southwest Border Technology Plan, claims to use lessons learned from previous failures and focuses on integrating systems tailored to individual sectors of the border. A related issue is the use of apprehension rates and other passive metrics as the measures of effectiveness for the security of the southern border, continuing the historical inconsistency of inaccurate reporting methods. The Department of Homeland Security (DHS) has recognized the problem of inadequate measurement and is developing new methods with the assistance of improved data captured with biometric systems; however, the issue of inaccurate reporting remains. An alternate and more active option to consider for measuring security effectiveness is red teaming. This thesis explores the following questions: what technologies are currently utilized for border security and how can their effectiveness be measured? And, can red teaming be used to improve on existing measures of effectiveness?

14. SUBJECT TERMS technology, border, security, effectiveness, red teaming	**15. NUMBER OF PAGES** 89
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

MEASURING THE EFFECTIVENESS OF SURVEILLANCE TECHNOLOGY AT THE U.S. SOUTHERN BORDER

Robert A. Hudspeth
Captain, United States Air Force
BS, Park University, 2011
MPA, Valdosta State University, 2012

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF ARTS IN SECURITY STUDIES
(HOMELAND SECURITY AND DEFENSE)**

from the

**NAVAL POSTGRADUATE SCHOOL
December 2019**

Approved by: Erik J. Dahl
Advisor

Scott E. Jasper
Second Reader

Afshon P. Ostovar
Associate Chair for Research
Department of National Security Affairs

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The United States' investment in southern border security has consistently been a topic of discussion regarding technological improvements and measurements of effectiveness. There have been multiple failed programs designed to combine infrastructure, personnel, and technology, ranging from the America's Shield Initiative (ASI) to the Secure Borders Initiative Network (SBInet). These efforts have resulted in billions of dollars of wasted funding. The latest initiative, named the Southwest Border Technology Plan, claims to use lessons learned from previous failures and focuses on integrating systems tailored to individual sectors of the border. A related issue is the use of apprehension rates and other passive metrics as the measures of effectiveness for the security of the southern border, continuing the historical inconsistency of inaccurate reporting methods. The Department of Homeland Security (DHS) has recognized the problem of inadequate measurement and is developing new methods with the assistance of improved data captured with biometric systems; however, the issue of inaccurate reporting remains. An alternate and more active option to consider for measuring security effectiveness is red teaming. This thesis explores the following questions: what technologies are currently utilized for border security and how can their effectiveness be measured? And, can red teaming be used to improve on existing measures of effectiveness?

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AAR	after action report
APSS	Agent Portable Surveillance System
ASI	America's Shield Initiative
ATP	Arizona Border Surveillance Technology Plan
BCI	Border Conditions Index
CAS	Civil Aviation Security
CBP	Customs and Border Protection
CIA	Central Intelligence Agency
CRS	Congressional Research Service
CRT	Computational Red Teaming
CTX	Computer Tomography X-ray
DHS	Department of Homeland Security
DOJ	Department of Justice
FAA	Federal Aviation Administration
FBI	Federal Bureau of Investigation
GAO	Government Accountability Office
IBM	International Business Machines
IFT	Integrated Fixed Tower
INS	Immigration and Naturalization Service
I-UGS	Imaging Unattended Ground Sensor
JCATS	Joint Conflict and Tactical Simulation Software
MANA	Map Aware Non-Uniform Automata
MMP	Mexican Migration Project
MSC	Mobile Surveillance Capability
MVSS	Mobile Video Surveillance System
NATO	North Atlantic Treaty Organization
OI	Office of Inspection
OIG	Office of Inspector General
OPCON	Operational Control
RAND	Reasonable and Non-Discriminatory

RVSS	Remote Video Surveillance System
SBInet	Secure Border Initiative
SBTP	Southwest Border Technology Plan
SUAS	small unmanned aerial vehicles
TID	Thermal Imaging Device
TSA	Transportation Security Administration
UFMCS	University of Foreign Military and Cultural Studies
UGS	Unattended Ground Sensor
USBP	United States Border Patrol

ACKNOWLEDGMENTS

First, I would like to thank my family for supporting me through the stressors of school and for providing a wonderful home for me to come back to every day. While Monterey has been a wonderful place to visit during our short time here, it is not an easy task moving everything we own and the children in and out of yet another school.

To my colleagues, thank you for your camaraderie on campus as we kept each other in line and on the path to success. We had a great time exploring many of the hikes Monterey has to offer as well as plenty of special events.

Finally, I would like to thank my thesis advisory team. Professor Dahl made selecting my primary advisor especially simple by demonstrating a positive, professional attitude every time I met with him and always being available in his office during duty hours. I selected my second reader, Professor Jasper, for similar reasons of professionalism and reliable office hours that allowed for face-to-face communication whenever I needed assistance.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. MAJOR RESEARCH QUESTION

The United States' investment in southern border security has consistently been a topic of discussion regarding technological improvements and measurements of effectiveness. There have been multiple failed programs designed to combine infrastructure, personnel, and technology ranging from the America's Shield Initiative (ASI) to the Secure Borders Initiative Network (SBInet). These efforts have resulted in billions of dollars of wasted funding. The latest initiative, named the Southwest Border Technology Plan, claims to use lessons learned from previous failures and focuses on integrating systems tailored to individual sectors of the border. A related issue is the use of apprehension rates and other passive metrics as the measures of effectiveness for the security of the southern border, continuing the historical inconsistency of inaccurate reporting methods. The Department of Homeland Security (DHS) has recognized the problem of inadequate measurement and is developing new methods with the assistance of improved data captured with biometric systems; however, the issue of inaccurate reporting remains. An alternate and more active option to consider for measuring security effectiveness is red teaming. This thesis explores the following questions: what technologies are currently utilized for border security and how can their effectiveness be measured? And, can red teaming be used to improve on existing measures of effectiveness?

B. SIGNIFICANCE OF THE RESEARCH QUESTION

Border security is a major priority for the current presidential administration and has been for decades. On January 25, 2017, President Trump signed Executive Order 13767 and charged the Department of Homeland Security with gaining total operational control (OPCON) of the United States' borders. Section 2 of the executive order defines operational control as "The prevention of all unlawful entries into the United States, including entries by terrorists, other unlawful aliens, instruments of terrorism, narcotics,

and other contraband."[1] However, this new order is not the first time DHS has been charged with acquiring OPCON of the United States' southern border. OPCON was used as a goal by DHS from 2004 to 2010 during an effort to measure the qualitative effect of enforcement at the southern border in order to "determine the proper mixture of personnel, technology, and infrastructure to deny or deter illegal entry into the United States."[2] OPCON was later dropped as a goal because of the lack of empirical evidence of success, as reported by the Government Accountability Office (GAO) and other organizations.[3]

Over the last two decades DHS has tracked several performance metrics for border security only to drop them soon after. From 2001–2004, optimum deterrence was used as a measure of effectiveness, meaning that success would be demonstrated when increased security measures no longer resulted in increased apprehensions. In 2005, OPCON became the new measure and showed success as each mile along the border was controlled by Border Patrol agents with the ability to detect, identify, and respond to illegal activity. After operational control was dropped, the apprehension rate became the measure in 2011. Though the apprehension rate was only intended to be the interim measure until 2013, while DHS developed a comprehensive measure called the border conditions index.[4] The border conditions index was intended to stand as the sole measure of border security by tracking estimated flows at entry ports, the quality of life in regions around the border, public safety, and wait times for legal flows at ports of entry.[5] However, the measure did not meet the demands required for a single comprehensive measure and its development was discontinued in 2013. Therefore, the apprehension rate still stands as the primary measurement today.

---

[1] Exec. Order, No. 13767, 3 C.F.R. 8793 (2017), https://www.federalregister.gov/documents/2017/01/30/2017-02095/border-security-and-immigration-enforcement-improvements.

[2] Department of Homeland Security, *Efforts by DHS to Estimate Southwest Border Security between Ports of Entry* (Washington, DC: Department of Homeland Security, 2017), 18, https://www.dhs.gov/sites/default/files/publications/17_0914_estimates-of-border-security.pdf.

[3] Department of Homeland Security, 18.

[4] Carla N Argueta, *Border Security Metrics Between Ports of Entry*, CRS Report No. R44386 (Washington, DC: Congressional Research Service, 2016), 3, https://fas.org/sgp/crs/homesec/R44386.pdf.

[5] Argueta, 3.

The inability to accurately measure the effectiveness of technology improvements at the border significantly affects the ability of DHS to acquire additional infrastructure and technology.[6] The data currently offered is incomplete, often unreliable, and easily subject to misinterpretation.[7] The acquisition, placement, and measuring of technological effectiveness is a necessity for proving that a given project is worth continuing. According to a 2007 RAND testimony, new border security technologies are extremely expensive and it is imperative that the level of performance gained from them justify the cost of developing and deploying them.[8] While the apprehension rate used by Border Patrol is simple to track, according to a study by the Bipartisan Policy Center tracking the rate of apprehensions does not indicate whether or not Border Patrol is meeting the goal of deterring or preventing illegal entry.[9]

A new metric used by Border Patrol in an attempt to demonstrate technological improvements is "technology assists." Technological assists consist of any apprehension that occurs where assistance was provided to an agent by a variety of assets, such as a ground sensor or an Unmanned Aerial Vehicle. A major issue is the fact that the data is dependent on reporting by individual agents who have received very little guidance by management on how to categorize inputs or understanding of the purpose behind the collection of data. An investigation performed by the GAO in 2017 proved the data set to be inaccurate by sampling the output and finding sufficient mistakes to render the compilation of data useless.[10]

---

[6] Office of Inspector General, *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry* (Washington, DC: Office of Inspector General DHS, 2017), 13, https://www.oig.dhs.gov/sites/default/files/assets/2017/OIG-17-39-Feb17.pdf.

[7] Office of Inspector General, 13.

[8] Brian A Jackson, *Developing Robust Border Security Technologies to Protect Against Diverse and Adaptive Threats* (Santa Monica, CA: RAND, 2007), 3, https://www.rand.org/content/dam/rand/pubs/testimonies/2007/RAND_CT294.pdf.

[9] Bryan Roberts, *Measuring the Metrics: Grading the Government on Immigration Enforcement* (Washington, DC: Bipartisan Policy Center, 2015), 15, https://bipartisanpolicy.org/wp-content/uploads/2015/02/BPC_Immigration_MeasuringEnforcement.pdf.

[10] Rebecca Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, GAO-18-119 (Washington, DC: Government Accountability Office, 2017), 30–32, https://www.gao.gov/assets/690/688666.pdf.

C. LITERATURE REVIEW

This literature review provides an overview of published research and expert opinions on the methods the Department of Homeland Security can use to measure the effectiveness of technology at the United States southern border. Congress directed DHS to provide detailed reporting on southwest border security through the Consolidated Appropriations Act of 2017. The act requires DHS to provide metrics capable of measuring the effectiveness of between ports of entry security as well as the data and methodology supporting the measure.[11] The Office of Immigration Statistics in DHS focused on three main categories to measure security effectiveness between the ports of entry: 1) the apprehension and interdiction rate, which uses apprehension survey data, interdiction effectiveness, total interdiction rates, and partial apprehension rates; 2) the deterrence rate, which uses the recidivism rate, deterrence survey data, and measures of illegal inflows; and 3) border crossing costs, which uses survey data on smuggler fees.[12] However, according to the most recent Border Security Metrics Report from DHS, research on these methods is "still a work in progress and DHS is not able to validate the modeling assumptions or quantify the uncertainty within the new estimation procedures."[13]

1. Measuring Border Security Effectiveness

The common issue among scholars and other professionals is that they do not agree on what the best method is to measure border security effectiveness. After a review of multiple GAO reports, Congressional Research Service (CRS) reports, RAND studies, and other documents, it is apparent that most of the literature focuses on indirect metric data collection over long periods of time to measure effectiveness instead of a more direct method. For example, RAND suggestions are: capture-recapture methods that tag apprehended individuals and determine immigration flow by their recaptures over time; sampling border segments by placing assets in areas with low, medium, and high flow and

---

[11] Department of Homeland Security, *Efforts by DHS*, 1.

[12] Department of Homeland Security, 3.

[13] Department of Homeland Security, *Department of Homeland Security Border Security Metrics Report* (Washington, DC: Department of Homeland Security, 2018), 7, https://www.dhs.gov/sites/default/files/publications/BSMR_OIS_2016.pdf.

calculating apprehensions; community surveys with questions that promote honesty; and lastly, synthetic modeling based on migrant risk and costs of "coyote" services.[14] The Bipartisan Policy Center recommends using methodology built from data that is already collected by Border Patrol, the first being known-flow data which combines the apprehension rate, individuals who give up while attempting to cross, and those known to get away. The second method is analysis of the recidivism rate, which is the percentage of border crossers caught more than once during the same fiscal year. The third and final method is migrant surveys focused on asking how many times individuals have been apprehended and how many attempts it took to cross.[15] As noted in the previous section DHS currently uses apprehension rates as a measure of border security effectiveness. As far back as its legacy organization the Immigration and Naturalization Service, the department has recognized the challenges involved with relying on the apprehension rate for measuring effectiveness, but as stated earlier, they are still working on refining the estimation techniques.[16]

DHS did adopt a few recommendations from outside agencies in its 2018 Border Security Metrics report, which is a step in the right direction from previous accounts of its reluctance to accept outside recommendations. But it still falls short in meeting the basic criteria established by the Bipartisan Policy Center report for good performance measurements.[17] A few requirements for this measurement criteria established by individuals from public administration and policy analysis are as follows: 1) measures should be meaningful, clear and readily understandable; 2) measures should be capable of being used by government agencies to inform decisions and resource allocation should be timely and actionable; and 3) measures should be stable over time.[18] Though border

---

[14] Andrew Morral, Henry Willis, and Peter Brownell, *Measuring Illegal Border Crossing Between Ports of Entry* (Santa Monica, CA: RAND Corporation, 2011), 11, https://www.rand.org/pubs/occasional_papers/OP328.html.

[15] Roberts, *Measuring the Metrics*, xi.

[16] Department of Homeland Security, *Department of Homeland Security Border Security Metrics Report*, 7.

[17] Roberts, *Measuring the Metrics*, vii.

[18] Roberts, 17.

security data has been public since 1950, there has never been a consistent measurement other than the apprehension rate.[19] Because the apprehension rate can appear to indicate successful performance whether the rate is rising or falling, it is not a valuable tool, but it is essentially the only outcome measure that the public and Congress have to measure border security effectiveness.[20]

Multiple GAO and CRS reports dating back to 2003 have concluded that DHS has instituted some programs that align with recommendations. But as the Inspector General stated in the 2017 OIG report, CBP still faces challenges with measuring effectiveness of its programs and operations in regards to securing the southwest border.

2. Red Teaming

Scholars and other experts have proposed very few direct methods for measuring border security effectiveness that would meet the criteria discussed above. Carla Argueta provides a few recommendations in the final section of her 2016 CRS report such as population surveys, regression models, and stratified samplings by placing surveillance resources to test certain areas, but she does not offer any detail on how they would be conducted. In the final sentence, however, she does mention red teaming as a method for detection in a sampled area and having migrants or agents attempt border crossings in order to establish interdiction probabilities.[21]

Red teaming encompasses a multitude of structured tests used to determine the intentions and capabilities of a competitor or institution using alternative analysis, penetration tests, and simulations to make a better-informed decision. However, other than the final sentence in the CRS report, there has been no mention about red teaming as a direct testing method for surveillance technology effectiveness. In a 2007 RAND testimony, Brian Jackson makes a recommendation to use red teaming as a means of testing

---

[19] Roberts, 10.

[20] Edward Alden, *Measuring the Effectiveness of Border Enforcement* (Washington, DC: Council on Foreign Relation, 2013), 3, https://cfrd8-files.cfr.org/sites/default/files/pdf/2013/03/Alden_Border_Security_Testimony_03-14-13%20-%20Final.pdf.

[21] Argueta, *Border Security Metrics Between Ports of Entry*, 18.

technology prior to broad implementation with a team of individuals capable of discovering new ways to penetrate them.[22] While this is useful in ensuring that technologies will perform over time, the method does not demonstrate surveillance technology effectiveness in a given area before and after deployment of surveillance technology.

Not focusing on the potential gains offered by red teaming could be a missed opportunity for DHS. An important point from the 9–11 Commission report states that red teaming is "notably lacking within the homeland security and intelligence elements of the Federal government."[23] The Department of Defense has, however, integrated red teaming in decision making for acquisitions and testing for years. A 2003 report on the status of red teaming stated that red teaming in the DoD plays an important role in training, concept development, and experimentation both during the experimental phase and after implementation. The report also found red teaming useful in the testing of secure systems where an opportunity does not usually exist, such nuclear storage and transportation.[24]

Though the name "red teaming" is new, the concept has been used by NATO for years under the name of "alternative analysis," and by the U.S. Naval War College under the title of "war gaming" as far back as 1923.[25] After witnessing the benefits of red teaming, the 2003 Defense Science Board recommended a strong presence of red teaming within the DoD but also suggested the development of a guide and coursework to demonstrate best practices.[26] Such a guide was released by the United States in 2005 for the University of Foreign Military and Cultural Studies, and a similar guide was produced by the United Kingdom Ministry of Defence. However, the guides only act as an academic introduction for individuals new to red teaming and do not provide technical guidance for the use of red teaming. The guides provide assistance with understanding the history of red

---

[22] Jackson, *Developing Robust Border Security Technologies to Protect Against Diverse and Adaptive Threats*, 2007, 7.

[23] 9/11 Commission, *The 9/11 Commission Report* (Washington, DC: Government Printing Office, 2004), 352, http://govinfo.library.unt.edu/911/report/911Report.pdf.

[24] Department of Defense, *The Role and Status of DoD Red Teaming Activities* (Washington, DC: Department of Defense, 2003), 3, https://fas.org/irp/agency/dod/dsb/redteam.pdf.

[25] Department of Defense, 31.

[26] Department of Defense, 1.

teaming, the successful conditions of red teaming, how to apply red teaming techniques, and highlighting red teaming's potential benefits, but they remain at a very abstract level and do not provide specific examples of red teaming used in history. The Ministry of Defence guide states that red teaming is prevalent in reducing risk and problem solving for commercial enterprises, such as IBM, and other governmental agencies such as the Central Intelligence Agency.[27] However, neither guide specifies how red teaming is used in these fields.

Recent literature on red teaming offers several more specific examples of how red teaming has been used in U.S. national security. Two examples are the Federal Aviation Administration (FAA) covert red team that tested and found airport security deficiencies prior to 9/11 and the simulations conducted by the Navy SEALs prior to the 2011 raid on Osama bin Laden in Pakistan.[28] The recommendations provided from these examples are applicable to both government and private sector requirements, and while they provide useful outcomes, the advice from the reports is not always taken. For example, after Micah Zenko describes a decade's worth of successful red teaming penetrations of airports across the globe in his book *Red Team*, he writes about the 400 pages' worth of vulnerability reporting that was dismissed by leadership within the FAA prior to 9/11.[29]

After reviewing the descriptions of red teaming in the private sector and DoD, it is apparent that the focus is on cyber security, with very little focus on physical security. The U.S. Army takes the lead on instructing red teaming in the DoD, but the Air Force Red Team program under the Air Force Directorate of Electronics and Special Programs offers the most relevancy to the testing of surveillance security systems at the border. Though there is no example provided, the 2003 DoD report states that the Air Force team incorporates a red vs blue interaction that evaluates and makes improvements to the systems of the defender, also known as the blue team. The outcomes of the tests are then

---

[27] Ministry of Defence, *Red Teaming Guide*, 2nd ed. (Wiltshire, UK: Ministry of Defence, 2012), ii, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/142533/20130301_red_teaming_ed2.pdf.

[28] Micah Zenko, *Red Team: How to Succeed by Thinking Like the Enemy* (New York: Basic Books, 2015).

[29] Zenko, 124.

used to guide decision making on technology development, provide warning in regards to the vulnerability of fielded technology, and shows success when the provided data has altered a development plan or acquisition to a better product.[30] The DoD recognizes red teaming's important roles in training; concept development and experimentation, both before and during concept development; testing the security of complex systems and networks; and exercising activities when there is typically no option to perform real tests, such as responding to nuclear weapon mishaps.[31]

D. POTENTIAL EXPLANATIONS AND HYPOTHESES

Dating back to the Immigration and Naturalization Service, outside agency reports[32] and DHS internal studies have recognized that the reliance on alien apprehension rates as the measurement of effectiveness between ports of entry security is ineffective, and efforts to find new measurements continue.[33] As stated by the Office of Inspector General for DHS, "CBP does not measure the effectiveness of its programs and operations well; therefore, it continues to invest in programs and act without the benefit of the feedback needed to help ensure it uses resources wisely and improves border security."[34]

This thesis tests the hypothesis that the method of red teaming is a direct measure of surveillance technology effectiveness and offers a clear picture for both the policy makers and the agents charged with enforcing border security. A dedicated red team much like the one established in the FAA prior to 9/11 or dedicated simulation team can perform penetration testing of border sectors. The results from these teams provides capture rate data both before and after technology deployment as well as offers continued testing to improve implementation for years after deployment. Instead of relying on passive data

---

[30] Department of Defense, *The Role and Status of DoD Red Teaming Activities*, 11.

[31] Department of Defense, 3.

[32] Marc R Rosenblum, *Border Security: Immigration Enforcement Between Ports of Entry*, CRS Report No. R42138 (Washington, DC: Congressional Research Service, 2013), 27, https://securityassistance.org/sites/default/files/R42138.pdf.

[33] Department of Homeland Security, *Efforts by DHS*, 1.

[34] Office of Inspector General, *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry*, 2.

collected over a period of time, this thesis argues that DHS should consider a red teaming option for measuring the effectiveness of improved surveillance technology at the border.

E. RESEARCH DESIGN

This thesis has three main goals: (1) to take a brief historical look at the attempts DHS has made to measure effectiveness of border security initiatives and better understand the resulting billions of dollars in wasted investment; (2) to provide an overview of the technology currently used at the border and how its effectiveness is measured; and finally (3) to describe red teaming and develop a method similar to that used by the FAA or other simulations that DHS could use to test technology efficiency at the southern border.

In order to meet the first goal, the primary sources and materials used were reviews by the Office of Inspector General grading DHS and assessing the history of mismanagement with recommendations. The second goal is met with GAO and Congressional Research Service reports that assess the technology at the border and provide feedback on the need to improve data quality in order to better assess effectiveness. The Bipartisan Policy Center recommendations tied with internal reviews and testimonies of senior DHS leadership describing future actions to better measure effectiveness were used to meet this goal as well. The third goal was met with the U.S. military and Ministry of Defence red teaming guides, and the Department of Defense's *The Role and Status of DoD Red Teaming Activities* report to provide an introduction and outline to red teaming; while literature focused on red teaming such as Micah Zenko's *Red Team: How to Succeed by Thinking Like the Enemy* provided a baseline for physical penetration and computer simulation red teams.

F. THESIS OVERVIEW

This thesis includes five chapters. Following Chapter I's introduction and literature review, Chapter II provides a background on the history of performance measurements and failures in order to provide the reader with a better understanding of the significance for effective measurement. Chapter III describes the current surveillance technology at the border and the more recent measures of effectiveness proposed by scholars and outside agencies. Chapter IV provides an overview of what is red teaming, requirements for

successful red teaming, and examples of physical and simulated red team models used in similar environments. Chapter V uses the pre-9/11 FAA red team as a case study to provide an outline for red teaming as well as lessons learned for future red teams. Chapter VI concludes with a summary of the findings and provides a recommendation for DHS to use for future measurements of surveillance technology effectiveness.

THIS PAGE INTENTIONALLY LEFT BLANK

II. THE BACKGROUND AND THE PROBLEM

In order for DHS to maintain Congressional funding for surveillance technology, there must be measures of effectiveness that help justify the security expenditures.[35] This leads to the question, how can the effectiveness of surveillance technology at the United States southern border be measured? The importance of this question first results from the executive order signed by President Trump in 2017 charging DHS to gain operational control of the border;[36] and second from the Consolidated Appropriations Act of 2017, in which Congress directed DHS to provide detailed reporting on the status of the southwest border security.[37] A defining moment for Border Patrol recognizing the need for improved measurements to show the effectiveness of its efforts was after the failure of the SBInet. In 2010, the Chief of Border Patrol testified that even after $3.5 billion was spent on border security, less than 3 percent of the border was under control.[38] The ability to accurately identify measures of effectiveness has been recognized as critical to border control for decades, yet Border Patrol continues to use the number of apprehensions, a recognizably poor indicator, as a measurement for illegal migration flows and successful security implementation.[39]

To provide a complete understanding of the historical problems of measuring the effectiveness of security at the border, the first section of this chapter begins with a background of the immigration enforcement system and major changes that have occurred to make it as large as it is today. The remaining sections discuss the previous attempts at measuring effectiveness, the problems with relying on apprehension rates, and finally the importance of reliable measurements.

---

[35] Alden, *Measuring the Effectiveness of Border Enforcement*, 10.

[36] Exec. Order, No. 13767, 3 C.F.R. 8793.

[37] Department of Homeland Security, *Efforts by DHS*, 3.

[38] Robert D. Schroeder, *Holding the Line in the 21st Century* (Washington, DC: U.S. Customs and Border Protection), 30, accessed May 14, 2019, https://www.cbp.gov/sites/default/files/documents/Holding%20the%20Line_TRILOGY.pdf.

[39] Office of Inspector General, *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry*, 8.

A. BACKGROUND

The U.S. immigration enforcement system is a network of several agencies that include DHS, the Department of Justice (DOJ), and the Department of State. The primary purpose of the immigration enforcement system is to prevent unlawful entry into the U.S. by arresting, detaining, and removing individuals who pose a threat to national security or threaten border security. DHS handles the majority of immigration functions through the U.S. Customs and Border Protection, which includes the Office of Field Operations and the U.S. Border Patrol. While the Office of Field Operations handles the ports of entry, the Border Patrol prevents unauthorized entry between the ports of entry. This portion of the thesis focuses on the role of Border Patrol and the ways to measure the effectiveness of its primary mission, preventing the unlawful entry of immigrants.

In the early 1900s, there were no specific legal channels for immigration and very little enforcement. The primary function of immigration officials was to process new arrivals, record their information, and inspect them for disqualifications such as disease.[40] The only restrictions before 1920 were directed towards Chinese immigrants who were banned starting in the 1880s.[41] The Border Patrol was established in 1925 to enforce new laws that enacted quotas based on national origin. Large-scale immigration was not a concern until World War II when labor shortages in the agricultural industry required a program to import Mexican nationals. The issue with illegal immigration occurred after the war when the Mexican national labor quota was cut, but the need for workers still existed. To combat this rise in unlawful entries, President Eisenhower initiated policies to tackle the issue and brought Border Patrol apprehensions down through the 1950s and 1960s.[42] After the total cancelation of the labor program in 1965, the apprehension rate rose again, this time with the inclusion of illegal immigration from Central American countries such as El Salvador and Nicaragua due to civil conflict.[43]

---

[40] Roberts, *Measuring the Metrics*, 4.

[41] Roberts, 4.

[42] Roberts, 4.

[43] Roberts, 4.

In order to control the influx of illegal immigration, the Immigration Reform and Control Act of 1986 was passed. This act granted legal status to many laborers as well as unlawful residents who had been present since 1982 but was also intended to increase enforcement at U.S. borders.[44] When the act failed to prevent the inflow of illegal immigration, significant enforcement buildup occurred in the late 1990s and again in the late 2000s, doubling the number of Border Patrol agents, developing major fencing projects, and improving the technology.[45] This buildup has made immigration enforcement an extremely large federal effort; the expenditures are roughly 50 percent of all federal law enforcement agencies combined, and the work-force makes up 45 percent of all federal law enforcement officers.[46]

B. PROBLEM

After understanding the size and extreme expenses that border security consumes, it is essential to look at the historical problem the U.S. government has had with developing measures of border security effectiveness. Immigration data has been collected by the federal government in the form of workflow data since 1892, and the data has been publicly available since the 1950s.[47] The various workflow data results consisted of useful information such as the number of citizens inspected at ports, aliens denied entry, apprehensions between ports, and the number of deportations annually. The problem is that the Immigration and Naturalization Service (INS) ceased publication of this data without explanation in 2002.[48] The need for identifying measures of effectiveness was recognized well before that date by Sandia National Laboratories in 1993, which stated that measures are critical to controlling the border; and substantial recommendations were also made in a 1997 Government Accountability Office report.[49] Multiple attempts were made by INS

---

[44] Immigration Reform and Control Act of 1986, S. 1200, 99th Cong. (1986).

[45] Roberts, *Measuring the Metrics*, 5.

[46] Roberts, 6.

[47] Argueta, *Border Security Metrics Between Ports of Entry*, 10.

[48] Roberts, *Measuring the Metrics*, 10.

[49] Office of Inspector General, *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry*, 8.

between 1997 and 2000 to incorporate these methods, the first being recidivism, which is the percentage of border-crossers who get apprehended more than once in a year. Subsequent years brought the interdiction rate for illegal entry at ports and the between-port operational effectiveness rate, which accounts for the ratio of apprehensions versus total entries attempted. Even with these attempts to measure effectiveness, it was determined that they were highly variable and did not always relate to the core missions of INS.[50] These issues were the beginning of an era of start-and-stop performance measurements that would continue through the development of DHS and extend on today.

DHS's attempt to measure border security has led to multiple iterations of performance measurements. In the last 15 years, DHS has started and stopped five different methods and is currently refining and developing more.[51] When DHS took over immigration functions in 2003, it canceled the INS performance measures of optimum deterrence and replaced them with the number of border miles considered under OPCON. Optimum deterrence was calculated with apprehension rates, border-related crimes, recidivism, smuggling fees, and property values as a measure for individual corridors along the border. The level at which applying increased Border Patrol agents and assets did not lead to increased arrests was considered reaching optimum deterrence.[52] The OPCON replacement considered a mile to be within OPCON when Border Patrol "employed the proper mix of personnel, technology, and infrastructure to detect, respond, and interdict illegal entry at the immediate border."[53] Significant changes to performance measures under the Obama administration led DHS to drop the OPCON performance measurement. Border Patrol chiefs were unable to effectively use this method to assess different border areas, and a mile-by-mile border assessment did not seem practical.[54] The U.S.

---

[50] Roberts, *Measuring the Metrics*, 11.

[51] Argueta, *Border Security Metrics Between Ports of Entry*, 10.

[52] Argueta, 10.

[53] Department of Homeland Security, *Efforts by DHS*, 18.

[54] Argueta, *Border Security Metrics Between Ports of Entry*, 11.

Government Accountability Office stated the OPCON measure was lacking empirical measures and was not an effective tool.[55]

The apprehension rate was intended to be used from 2011 to 2013 as an interim fix while DHS worked on the Border Conditions Index (BCI), but it is still used today. The BCI was projected to stand as the sole measure of border security by tracking estimated flows at entry ports, the quality of life in regions around the border, public safety, and wait times for legal flows at ports of entry.[56] However, the measure did not meet the demands required for a single comprehensive measure, and its development was discontinued in 2013.

C. THE APPREHENSION RATE

There are many problems with relying on the apprehension rate as a performance metric for border security. A Congressional Research Service report in 2012 listed three reasons why the apprehension rate is a poor indicator that is highly affected by other trends. First, the data excludes successful unauthorized alien entries, unsuccessful unauthorized aliens, and would-be unauthorized aliens, which leads to an incomplete picture of migration enforcement and total unauthorized migration. Second, apprehension data accounts for events instead of people; when the same person is apprehended multiple times, it overestimates the actual number of illegal attempts. Finally, the apprehension rate does not account for the economic downturns or demographic changes that occur on the other side of the border.[57] RAND states that measures such as apprehension rates are indirect and unrelated to the mission of border control, which makes it an unreliable management tool. For the apprehension rate to be effective, the Border Patrol would need to know the total flow of immigrants, which is not yet possible.[58]

---

[55] Department of Homeland Security, *Efforts by DHS*, 18.

[56] Argueta, *Border Security Metrics Between Ports of Entry*, 3.

[57] Rosenblum, *Border Security: Immigration Enforcement Between Ports of Entry*, 22.

[58] Morral, Willis, and Brownell, *Measuring Illegal Border Crossing Between Ports of Entry*, vii.

While the apprehension rate can be useful in some aspects, it is not adequate for measuring performance, especially when it can be interpreted as both a positive and a negative. Like any law enforcement effort, increased arrests can mean either improved policing that catches more criminals, or it can be the outcome of more individuals violating laws.[59] Customs and Border Protection (CBP) has been known to interpret the numbers in both ways, with increased apprehension being a positive factor at checkpoints, while decreases in apprehensions represent improvements to deterrence from technology and staffing between ports of entry.[60] The Office of Inspector General (OIG) also believes that less apprehensions represent improvements in security measures between ports, but since Border Patrol does not calculate illegal circumvention of checkpoints, the OIG argued it was not a suitable measure of effectiveness.[61]

Even though the Border Patrol Chief stated in a 2017 testimony that the decline in apprehensions is an improvement, the OIG and GAO both concluded that using it as a performance measurement limited accountability and congressional oversight.[62] Though it is clear that a measure that demonstrates success whether it increases or decreases has little use as a tool for evaluation, as of now, it is essentially the only measurement that the public and Congress are provided to judge the accomplishments or failures of border enforcement.[63] A recent GAO report concluded that apprehension data does not lead to proper allocation decisions or program results, and until new measures are established, DHS and Congress can expect incomplete oversight and accountability.[64] The OIG concluded that CBP's inadequate or sometimes nonexistent performance measurements constitute an ongoing trend.[65]

---

59 Alden, *Measuring the Effectiveness of Border Enforcement*, 3.

60 Alden, 3.

61 Office of Inspector General, *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry*, 9.

62 Office of Inspector General, 9.

63 Alden, *Measuring the Effectiveness of Border Enforcement*, 3.

64 Argueta, *Border Security Metrics Between Ports of Entry*, 11.

65 Office of Inspector General, *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry*, 9.

D. IMPORTANCE OF MEASUREMENTS

The previous section demonstrated the consistent inability for DHS to measure performance. It is also essential to understand why these performance measurements are so important. In order to provide policymakers and the general public with information on how much expenditure is enough or if the technology is making a difference, there must be measurements. Regarding immigration and enforcement programs that, as mentioned earlier, consume 50 percent of all law enforcement expenditures, proper measurements quickly become a topic of controversy. When it comes to measuring the performance of border security, three sets of data are key: inputs, outputs, and outcomes.[66] An example of an input would be the cost of a new integrated fixed tower placed in a region. The output would be the documented assistance the tower provided to border patrol agents. Both the input and output data sets are easily measured through cost and documentation. For policymakers, a beneficial outcome of the investment in the tower would be to know the total number of illegal entries that were detected and apprehended in that area versus the total amount that was not, but this data is not feasibly attainable.

The information available to policymakers for decision making is typically the input measures such as the number of agents deployed, the amount of funding toward programs, or the number of miles covered with border fencing.[67] However, there are two reasons why the outcome measurement is needed to make decisions on how much spending is required or how effective a program is. First, outcome measures are capable of displaying the ability for DHS to meet its immigration enforcement goals of preventing illegal import and entry, or specifically OPCON.[68] Second, outcome measures are useful in decision making, and public debate over the security of the border, which many believe has not improved since 2005.[69] The 2017 OIG report states that the border is still porous, and it is

---

66 Roberts, *Measuring the Metrics*, 15.

67 Roberts, 15.

68 Paulina Orchard, *The 2014 Quadrennial Homeland Security Review* (Washington, DC: Department of Homeland Security, 2014), 77, https://www.dhs.gov/sites/default/files/publications/2014-qhsr-final-508.pdf.

69 Roberts, *Measuring the Metrics*, 15.

questionable whether or not the significant investments have made any improvements to border security.[70] Without a proper measurement for technology investment outcomes, it is unlikely that policymakers and the public will agree on the state of the border, making the decision of how much or what should be done more difficult.

Of the performance measures typically reported for government agencies, there are three that DHS should strive to meet with technology effectiveness. First, measures should be readily understandable, meaningful, and clear. Second, measures that inform decisions by government agencies for resource allocation should be actionable and timely. Third, the measures should be stable, consistent, reliable, and uniform over time.[71] As the previous section shows, it is apparent that these measures are not being met by DHS.

---

[70] Office of Inspector General, *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry*, 14.

[71] Theodore H Poister, *Measuring Performance in Public and Nonprofit Organizations* (San Francisco, CA: Jossey-Bass, 2003), 101, http://www.untag-smd.ac.id/files/Perpustakaan_Digital_2/NON%20PROFIT%20ORGANIZATION%20Measuring%20Performance%20in%20Public%20and%20Nonprofit%20Organizations.pdf.

III. BORDER SECURITY PROGRAMS—THEN AND NOW

DHS has experienced a series of failures in the past, partly due to the inability to build metrics into program management and planning, as well as "inadequately collecting reliable and complete data for program performance."[72] America's Shield Initiative (ASI) was initiated in 2004, and used sensors, cameras, and databases to create an Integrated Surveillance Intelligence System. After a review by the Government Accountability Office, ASI was deemed to not have effective program management elements such as defined roles and responsibilities and acquisition management.[73] Before the program was set into motion, it was reevaluated and rolled into the larger SBInet that was announced by the DHS in 2005.[74]

SBInet expanded the ASI integrated network idea to include radar/camera towers, ground sensors, unmanned aerial surveillance, and mobile surveillance all linked into a command post in contact with individual agents in the field.[75] GAO audited the program multiple times and found that CBP was once again not following good practices regarding the evaluation of overall effectiveness the technology had on border security.[76] After extreme cost overruns and no measurable benefit to border security, the Homeland Security Secretary canceled SBInet in January of 2011. By the time SBInet was canceled, it had cost taxpayers over $1 billion and covered only 53 miles of the 2000-mile southern border.[77]

Within a month of canceling SBInet, the Arizona Border Surveillance Technology Plan (ATP) was announced. While this plan took a different approach to acquisition by

---

[72] Office of Inspector General, *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry*, 10.

[73] Office of Inspector General, 11.

[74] Reed Abrahamson, "Fixing the Net: The Fall of SBInet, the Rise of Integrated Fixed Towers," *Georgetown Immigration Law Journal* 25, no. 3 (2011): 1.

[75] Abrahamson, 1.

[76] Office of Inspector General, *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry*, 11.

[77] Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 2.

purchasing previously tested technology instead of paying for the development and implementation costs that SBInet incurred, there were still issues. First, CBP did not develop documented analysis to justify increased border surveillance technology. Second, there were no predefined mission benefits to meet before implementing ATP. Finally, since there was no assessment for the effectiveness of the already placed SBInet systems, management was unable to make adequate decisions to improve the ATP further.[78] In 2014, the current Southwest Border Technology Plan (SBTP) was developed, which plans to incorporate the ATP developments and extend them beyond Arizona to the rest of the border.[79]

This chapter examines the current state of technology along the U.S. southern border, first by reviewing the Southwest Border Technology Plan and its major components. Next, it will describe a few recommended methods from outside organizations that DHS has adopted and refined as well as the primary methods that DHS is moving forward with as demonstrated in the 2018 DHS Metrics Report.

A. SOUTHWEST BORDER TECHNOLOGY PLAN

To better understand the current technology efforts at the southern border when discussing the performance measures, this section discusses the SBTP and its technological systems in more detail. Border Patrol developed the SBTP with a two-step process. First, teams of analysts identified the types of technology to be used for the overall plan and aligned them with 13 different sectors along the border that were most fitting by terrain. Second, the analysts narrowed the project based on the quantities of each technology type needed in each sector by factoring operational conditions of traffic patterns, weather, infrastructure, and vegetation.[80] This section provides a brief overview of the seven primary surveillance systems used in the SBTP (see Figures 1–7).

---

78 Office of Inspector General, *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry*, 12.

79 Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 8.

80 Gambler, 9.

Figure 1. Integrated Fixed Tower[81]

Integrated Fixed Tower (IFT): The IFT is an 80-to 160-foot-tall fixed tower that includes a mounted radar with day and night cameras. The tower is capable of covering large areas of rugged terrain with a clear picture of what is being detected to prevent the need for sending Border Patrol agents to livestock movements or other false alarms. The radar and cameras send information wirelessly through microwave links to a central hub station where the data is monitored by Border Patrol agents that determine appropriate responses. The monitoring agents are able to "detect a single walking average-sized adult at up to 7.5 miles in both daylight and darkness."[82] The high resolution video produced is so specific that it can determine what the detected individual is carrying, whether it be a long-arm weapon or a backpack. This level of situational awareness improves the responding Border Patrol agent's operational capability by informing them on what the hazards are before they arrive, allowing the agents to be more effective, efficient, and safe,

[81] Source: Gambler, 11.

[82] Mitch Moxley, "Better Than a Wall: A New Detection System Can Help Monitor the U.S.-Mexico Border," *Popular Mechanics*, January 28, 2016, 4, https://www.popularmechanics.com/technology/security/a18622/border-control-integrated-towers-system-invisible-wall/.

which has not always been the case.[83] The IFT towers are manufactured by an Israeli based company called Elbit Systems and designed to be much more rugged than previously deployed systems, and more capable of handling the harsh dessert environments. Israel has deployed similar IFT towers along hundreds of miles of its Palestinian, Gaza, and Egypt borders in recent years.[84]

Figure 2. Remote Video Surveillance System[85]

Remote Video Surveillance System (RVSS): This system is much like the IFT but does not include a radar system and can only be mounted up to 120 feet. Highly trafficked and populated areas tend to over saturate radar systems like that of the IFT, making the RVSS the preferred option in these specific area types. While RVSS towers are currently in use at both the northern and southern U.S. borders, the systems are being upgraded with

[83] Eric Blum, "Further Reflection," Department of Homeland Security, 1, accessed August 26, 2019, https://www.cbp.gov/frontline/frontline-june-az-technology.

[84] Moxley, "Better Than a Wall: A New Detection System Can Help Monitor the U.S.-Mexico Border," 4.

[85] Source: Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness,* 11.

new cameras fit for long, medium, and short-rang surveillance.[86] Supplemental power is generated by local solar panels while the transfer of data is sent to a command and control center through a microwave link much like the IFT system.[87] There is a relocatable variant that is mounted to an 80-foot-tall tower on a platform trailer that provides Border Patrol agents an opportunity to strategically locate the tower for a shorter period of time.[88]

Figure 3. Unattended Ground Sensors and Imaging Sensors[89]

Unattended Ground Sensors and Imaging Sensors (UGS and I-UGS): In her testimony, Clair Grady, the DHS Under Secretary describes UGS as "remotely monitored surveillance systems that detect, identify and track activity and subjects in areas not easy

86 *Bang for the Border Security Buck: What Do We Get for $33 Billion?*, 115 Cong. (2018) (statement of Claire Grady, MGMT Under Secretary). https://www.dhs.gov/news/2018/03/15/written-testimony-mgmt-under-secretary-and-cbp-house-homeland-security-subcommittee.

87 Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 11.

88 Blum, "Further Reflection," 2.

89 Source: Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 11.

to access or monitor with other technology."[90] Grady further states that these sensors are stationary when installed, but easily relocatable and concealed when necessary; and can track and identify the movement of humans, animals, and vehicles with the ability to differentiate them from each other. The information is sent to a command and control center as well as directly to Border Patrol agents who are carrying hand held monitors in the field.[91] The detection capabilities provide a wide range possibilities from seismic, magnetic, acoustic, infrared, radar, and microwave sensors. The Imaging variant called the I-UGS, provides photo or video verification of the detections and enables the agents to perform image analyses of the data upon receipt.

Figure 4. Agent Portable Surveillance System[92]

Agent Portable Surveillance System (APSS): This system contains daylight and infrared cameras, radar, and a laser illuminator. The APSS provides medium-range mobile

[90] Grady, testimony on *Bang for the Border Security Buck.*

[91] Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 11.

[92] Source: Gambler, 11.

surveillance and is portable by two or three agents for use in areas where more permanent systems are not capable of reaching.[93] While a direct link is provided to agents, there is no link to a command and control center.[94] These suitcase-based camera and radar systems provide Border Patrol agents improved visibility at key vantage points inaccessible by larger truck-mounted systems.[95]

Figure 5. Thermal Imaging Device[96]

Thermal Imaging Device (TID): This system contains a portable handheld infrared camera that enables border patrol agents to see up to 5 miles in dim lighting or total darkness in varying weather conditions from rain to dense fog and blowing dust.[97] Much

[93] Grady, testimony on *Bang for the Border Security Buck.*

[94] Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 11.

[95] Blum, "Further Reflection," 3.

[96] Source: Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness,* 11.

[97] Gambler, 11.

like the APSS, these rugged portable systems provide Border Patrol agents with an advantage at key observation points inaccessible to larger vehicles.

Figure 6. Mobile Surveillance Capability[98]

Mobile Surveillance Capability (MSC): This truck-mounted system consists of daylight and infrared cameras, radar, a laser illuminator, and a laser range finder all mounted to a retractable 25-foot tower for long-range surveillance.[99] Information is sent to the crew inside of the truck through multiple monitors, but no data is sent to a command and control center.[100] The control room for the deployed agents is in the backseat, the passenger seat faces the rear to provide the agents access to monitors for the radar and camera, with a keyboard and track pad to control them. The MSCs are on a ruggedized vehicle capable of traversing rough terrain and maintaining a location for up to a week when deployed.[101]

---

98 Source: Gambler, 11.

99 Grady, testimony on *Bang for the Border Security Buck.*

100 Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 11.

101 Blum, "Further Reflection," 3.

Figure 7. Mobile Video Surveillance System[102]

Mobile Video Surveillance System (MVSS): This is a trimmed down version of the MSC variant and does not include radar capability. Where the MSC is designed for long-range surveillance, the MVSS performs short and medium-range surveillance in areas with higher levels of activity.[103] The information is sent to monitors within the cab of the truck and not to a command and control center.[104] The MVSS provides CBP an affordable option to outfit standard pickup trucks with a mobile camera platform that can remain operational for up to 72 hours without charge or be removed from the truck and temporarily deployed with a solar panel for power. One truck is capable of placing and monitoring multiple MVSS systems at a time as long as it remains within a few miles of the mobile camera systems therefore drastically increasing the coverage capability of each border patrol team.[105]

---

[102] Source: Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness,* 11.

[103] Grady, testimony on *Bang for the Border Security Buck.*

[104] Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 11.

[105] "The Eagle MVSS," Tactical Micro, 2, accessed September 9, 2019, http://www.tacticalmicro.com/products/MVSS_CATALOG.pdf.

These systems make up an extensive and highly capable network of sensors designed to detect and track people and vehicles approaching or crossing the border. In the CBP article "Further Reflections" a Border Patrol agent explains that these surveillance technologies not only improve officer safety with each response but they drastically reduce the time required for each incident in the field; where interdictions previously took between 8–10 hours, they now are typically resolved in less than an hour.[106] However, there are still unanswered questions; what is the ultimate impact of these systems—and how can that outcome be measured? These issues will be discussed in the following sections…

B. RECENT MEASURES

There have been many recommendations for performance measurements of border security to DHS in the last few years, and DHS has adopted or refined many of them. This section describes a few recommended methods as well as the methods DHS is moving forward with in the 2018 DHS Metrics Report.

In 2011, RAND suggested four "promising methods" to relieve DHS from relying on the number of apprehensions: capture-recapture methods that tag apprehended individuals and determine immigration flow by their recaptures over time; sampling border segments by placing assets in areas with low, medium, and high flow and calculating apprehensions; community surveys with questions that promote honesty; and synthetic modeling based on migrant risk and costs of "coyote" services.[107]

In 2015, The Bipartisan Policy Center recommended improving methodology on data that is already collected by Border Patrol. The first recommendation was known-flow data, which combines the apprehension rate, individuals who give up while attempting to cross, and those known to get away. The second method is an analysis of the recidivism rate, which is the percentage of border crossers caught more than once within the same

[106] Blum, "Further Reflection," 4.

[107] Morral, Willis, and Brownell, *Measuring Illegal Border Crossing Between Ports of Entry*, 11.

fiscal year. The third and final method is migrant surveys focused on asking how many times individuals have been apprehended and how many attempts it took to cross.[108]

1. Known Flow Data

The known-flow data, also known as the effectiveness rate, has been collected since the 1990s to analyze the performance of enforcement operations in preventing illegal entry.[109] The known-flow data as shown in Figure 8, is calculated by combining apprehensions with turn backs, then dividing the number by the total of apprehensions, turn backs, and got-aways.[110] DHS defines apprehensions as removable aliens arrested by USBP; got-aways are "those that make an illegal entry, are not turned back or apprehended, and are no longer pursued by USBP"; while turn backs are subjects who attempt illegal entry but give up, return to their origin, and are not apprehended.[111]

$$\frac{Apprehensions + Turn\ Backs}{Apprehensions + Turn\ Backs + Got\ Aways}$$

Figure 8. Known Flow Data/The Effectiveness Rate[112]

The known-flow method has a few drawbacks, however, starting with the fact that it excludes undetected entry, which exaggerates the effectiveness of enforcement.[113] Secondly, since the data is not collected by individual biometrics and is instead based on the event of capturing, there is a chance of double counting.[114] Finally, due to the variation

[108] Roberts, *Measuring the Metrics*, xi.

[109] Roberts, *Measuring the Metrics*, 25.

[110] Argueta, *Border Security Metrics Between Ports of Entry*, 8.

[111] Department of Homeland Security, *Department of Homeland Security Border Security Metrics Report*, 11.

[112] Source: Argueta, *Border Security Metrics Between Ports of Entry*, 8.

[113] Alden, *Measuring the Effectiveness of Border Enforcement*, 7.

[114] Argueta, *Border Security Metrics Between Ports of Entry*, 9.

of reporting got away and turn-back data between sectors, the rate cannot be compared among them.[115] The most recent DHS report recognizes the shortcomings of the effectiveness rate but states that it is the only tool available for analysis of sector-level security until further developments are made.[116]

2. The Recidivism Rate

The recidivism rate also began in the 1990s and has been used by DHS with improvements in accuracy enabled by the introduction of biometric systems.[117] The intent behind the recidivism rate is that border effectiveness shows improvement when the rate goes down, implying that individuals who are caught are less likely to try again. The rate, as shown in Figure 9, is calculated by dividing the number of individual subjects apprehended multiple times by the total number of individual subjects apprehended within the same year.[118]

$$\frac{\textit{Unique Subjects Apprehended Multiple Times}}{\textit{Total Unique Subjects Apprehended}}$$

Figure 9. The Recidivism Rate[119]

The issues with the recidivism rate as a measure of performance come from the multiple factors that are not accounted for. The first is the sheer distance that different migrants have to travel. Economic drivers that cause migration from countries other than Mexico require a much longer distance to travel and will drop the recidivism rate even

---

115 Argueta, 9.

116 Department of Homeland Security, *Department of Homeland Security Border Security Metrics Report*, 18.

117 Roberts, *Measuring the Metrics*, 24.

118 Department of Homeland Security, *Department of Homeland Security Border Security Metrics Report*, 24.

119 Source: Argueta, 7.

though security is unchanged.[120] Second, since apprehension is required to gather data, a decrease in the recidivism rate could technically be caused by a decrease in the apprehension rate.[121] Finally, the increased intensity in enforcement over time could increase both apprehension and deterrence rates, which would muddy the results since the offset would cause no change in the recidivism rate.[122] DHS recognizes these mentioned issues as well as recommendations by GAO for improvement, but states in the most recent metrics report that the annual recidivism rate is a useful measure of performance.[123]

3. Migrant Surveys

Migrant surveys have been conducted by outside agencies regarding border security since 1987 through surveys conducted by the Mexican Migration Project (MMP) which is a binational effort led by scholars in Mexico and the United States.[124] DHS uses survey data to calculate deterrence as well as the cost of coyote services as a measure of performance.[125] According to DHS, deterrence is the estimated portion of immigrants who have unsuccessfully attempted entry, are from that point deterred from reattempting, and either return to their place of origin or remain in Mexico.[126] Agents gather data from deportees at reparation facilities and ask them "about their intentions to return to the United States within the next 7–90 days"; the responses correlate to a level of changes in deterrence over a period of time.[127] The data for the cost of smuggling services, also known as coyote fees, is gathered from surveys and interviews conducted by USBP.[128] Survey data has many limitations regarding the sample pools and trustworthiness of the

---

120 Argueta, *Border Security Metrics Between Ports of Entry*, 7.

121 Roberts, *Measuring the Metrics*, 25.

122 Argueta, *Border Security Metrics Between Ports of Entry*, 7.

123 Department of Homeland Security, *Department of Homeland Security Border Security Metrics Report*, 24.

124 Alden, *Measuring the Effectiveness of Border Enforcement*, 6.

125 Department of Homeland Security, *Efforts by DHS*, 10–13.

126 Department of Homeland Security, *Department of Homeland Security Border Security Metrics Report*, 46.

127 Department of Homeland Security, 47.

128 Department of Homeland Security, *Efforts by DHS*, 13.

interviewees. When the sample only includes apprehended individuals, Mexican nationals, or those already living in the United States, the information will be skewed.[129] Those surveyed about coyote services may be reluctant to mention the use of the service or provide details on the payment process.[130] DHS currently uses surveys and interviews for the deterrence rate as the only way to measure the intentions of immigrants to make future illegal entry attempts.[131]

4. Asset Assists

In 2014, GAO recommended that CBP develop a performance metric capable of displaying the contributions surveillance technologies have on border security, called asset assists.[132] Asset assists are documented when technology or other assets such as working dog teams or sensor systems assist with the capture of an individual attempting illegal entry.[133] The intent for asset assist data is to provide decision-makers with a clear understanding of changes in apprehensions before and after the placement of technology.[134] If this plan were to be implemented correctly, it would ease the concern of huge enforcement expenditures authorized by Congress without properly accounting for the effectiveness of the resources.[135]

However, after GAO analyzed asset assist data from 2014 to 2017, they discovered numerous discrepancies. First, agents repeatedly input asset assist data as "other" or even attributed assists to technology that was not even installed in the region, such as IFT towers in the Rio Grande Valley. After leadership was notified of the discrepancies in 2016,

---

129 Roberts, *Measuring the Metrics*, 27.

130 Department of Homeland Security, *Department of Homeland Security Border Security Metrics Report*, 48.

131 Department of Homeland Security, 46.

132 Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 29.

133 Argueta, *Border Security Metrics Between Ports of Entry*, 6.

134 Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 29.

135 Alden, *Measuring the Effectiveness of Border Enforcement*, 10.

further data analysis showed the same issue was occurring in 2017.[136] Second, Border Patrol agents were not briefed on how to input data, nor were they briefed on the importance of why the data was even being collected. The results of these findings led GAO to label the data unreliable; DHS officials stated the collection was only meant to satisfy external agency requests and never meant to affect budgeting, planning, or performance measurements.[137]

136 Gambler, *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*, 17.

137 Gambler, 32.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. INTRODUCING RED TEAMING

Though the name "red teaming" is new, the concept has been used by NATO for years under the name of "alternative analysis," and by the U.S. Naval War College under the title of "war gaming" as far back as 1923.[138] Along with the recommendation of increased red teaming within the DoD, the 2003 Defense Science Board requested the development of a guide and coursework to demonstrate best practices.[139] The United States released such a guide in 2005 for the University of Foreign Military and Cultural Studies (UFMCS), and the United Kingdom Ministry of Defence produced a similar guide. The guides mostly act as an academic introduction for individuals new to red teaming and do not necessarily provide technical guidance for the use of red teaming. The guides assist with understanding the history of red teaming, the successful conditions of red teaming, how to apply red teaming techniques, and highlighting red teaming's potential benefits.

The Ministry of Defence guide defines red teaming as "the independent application of a range of structured, creative, and critical thinking techniques to assist the end user in making a better-informed decision or produce a more robust product" allowing for a focus on system testing.[140] The UFMCS defines red teaming with an intellectually focused approach as a "function to avoid groupthink, mirror imaging, cultural missteps, and tunnel vision in plans and operations; to help staffs avoid making poor assumptions and account for the complexity inherent in the Operational Environment."[141] With less of a focus on the specific management side that the UFMCS has, the 2003 DoD Science Board report and the U.K. guides provide the better outline for a productive red team, one that DHS can learn from.

---

138 Department of Defense, *The Role and Status of DoD Red Teaming Activities*, 31.

139 Department of Defense, 1.

140 Ministry of Defence, *Red Teaming Guide*, 1–3.

141 University of Foreign Military and Cultural Studies, *Red Team Handbook* (Fort Leavenworth, KS, 2012), 1, http://www.au.af.mil/au/awc/awcgate/army/ufmcs_red_team_handbook_apr2012.pdf.

A. EFFECTIVE RED TEAMING

Effective red teaming requires the end user to fully support the team and be open to using the products in future decision making.[142] Red teaming only works in an atmosphere that accepts and values critiques for improvement.[143] According to Zenko, institutions that are unable to accept or utilize a red team's findings are better off not performing a test in the first place. He believes the top levels of leadership need to provide proper direction, offer adequate resources, and ensure that the rest of the organization values the red team as well; if this does not occur, the entire process will more than likely be ignored.[144] It is expected that issues raised will not be welcome to the organization, and leadership top cover is required to guarantee that red teams have the required level of independence and that their outputs are seriously considered.[145] The end user of red team outputs establishes the parameters that the team should stay within from the beginning, not as an afterthought.[146]

1. Effective Red Teaming Development

There are three steps in developing effective red teaming. The first step is to identify the specific task the team needs to accomplish at an initial state. Red teams are typically used after problems have already occurred or when there has already been a heavy investment in repairing issues, and earlier use could have made changing directions easier.[147]

The second step for an end user is to identify the appropriate red team and team leader who possess the skills necessary to accomplish the task.[148] Team member quality, the team synergy, and a shared vision are the most important factors in red team

---

142 Ministry of Defence, *Red Teaming Guide*, 2–1.

143 Department of Defense, *The Role and Status of DoD Red Teaming Activities*, 6.

144 Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 235.

145 Department of Defense, *The Role and Status of DoD Red Teaming Activities*, 6.

146 Ministry of Defence, *Red Teaming Guide*, 2–2.

147 Department of Defense, *The Role and Status of DoD Red Teaming Activities*, 7.

148 Ministry of Defence, *Red Teaming Guide*, 2–3.

performance.[149] According to Zenko, red teamers need to be: 1.) creative, confident, open-minded, and a "little odd," while still able to communicate well with the targeted organization; and 2.) "Have a big bag of tricks," as variety is essential to red teaming. When methods used by the team are predictable or already well known within the organization, they have little to no effect. Red teamers need to be able to adapt quickly and be ready to use new tactics and techniques.[150] The Ministry of Defence guide states that red teams should be composed of "critical thinkers, subject matter experts, analysts, cultural advisors, and role players with a team size that appropriately matches the assigned task."[151]

The third step is to task and provide freedom of decision making to the red team leader. The red team leader needs to be at least semi-independent to perform assessments effectively.[152] More specifically, the leader needs the freedom to run the team with techniques he or she deems appropriate that may include attacking the given issue from angles not originally identified by management.[153]

2. Ineffective Red Teaming

According to the DoD, typical causes of failures include the red team not taking its assignment seriously, usually due to not being provided a clear objective; losing independence by just trying to meet the end user's personal goal; destroying the integrity of the process by leaking information during the test phase; not performing the role of an adversary adequately by mirror imaging abilities; and simply not providing interesting challenges to the blue side, whether due to lack of skill or overbearing constraints.[154]

The amount of red teaming an organization requires is not easily defined. Red teaming should be performed often enough to detect and address emerging vulnerabilities

---

[149] Michael J. Skroch, "Modeling and Simulation of Red Teaming," U.S. Department of Energy Office of Scientific and Technical Information, 2009, 4, https://www.osti.gov/biblio/972439-modeling-simulation-red-teaming-part-why-red-team.

[150] Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 235.

[151] Ministry of Defence, *Red Teaming Guide*, 2–3.

[152] Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 235.

[153] Department of Defense, *The Role and Status of DoD Red Teaming Activities*, 2–3.

[154] Department of Defense, 6.

but not so often that it is disruptive to the organization and its employees, or does not allow enough time to make adjustments to previous findings between tests.[155] The three "golden rules" of successful red teaming according to the U.K. guide are 1.) timeliness, to be useful to the end user, which also meets the aforementioned performance measure; 2.) quality, to retain the red team's credibility and make the final report useful, also meeting a performance measure; and 3.) access, as findings should be presented to the correct level to influence proper decision making.[156]

B. RED TEAMING IN PRACTICE: PHYSICAL PENETRATION

By implementing physical and simulated penetration testing at the border both before and after deployment of technology, DHS can provide immediate data to policymakers that demonstrates its effectiveness, as well as war game potential adversarial tactics provided by the intelligence community.

Physical penetration testing is conducted in four phases. First, the team begins by simply scoping out the engagement area or targeted institution. Second, information is gathered by active reconnaissance of the building or, in the case of the border, a targeted region. Third, the team conducts the actual penetration of the targeted area. Finally, the team presents the findings and a prioritized list of recommendations to the organization's leadership.[157] Physical penetration tests are used to prove that an organization's security measures have inadequacies that can be bypassed through challenging untested assumptions, finding strategic blind spots, and uncovering security weaknesses.[158]

The FAA initiated physical penetration testing by red teams in 1991 as a response to the 1988 bombing of Pan Am Flight 103 to identify shortfalls in airport security.[159] Most of the penetration tests were conducted by a red team member, known only to management, attempting to circumvent security by smuggling dangerous components

---

[155] Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 236.

[156] Ministry of Defence, *Red Teaming Guide*, 2–7.

[157] Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 197.

[158] Zenko, 206.

[159] Zenko, 117.

through airport security. If a team member was detected, he or she could quickly provide documentation of the test and be cleared. Throughout the 1990s, the teams discovered a wide range of security deficiencies at every airport they tested, including the Frankfurt International Airport where 44 of 44 of their smuggling attempts were successful in 1996, the same airport the Pan Am 103 flight originated from.[160]

Tests were continued after 9/11 by the GAO, which conducted multiple tests at airports, border points of entry, and government buildings directed by Congress under the authority of the comptroller general.[161] Using the information available to the general public, the GAO was able to smuggle radioactive material through ports of entry at the border in 2006, though it was deemed a nearly immeasurable amount.[162] The GAO team was also able to smuggle bomb components through 19 airports in 2009, into 10 of 10 federal buildings in 2009, and a major seaport in 2011.[163] Further tests by GAO between 2003 and 2007 showed weaknesses at CBP's ports of entry, where the teams used fraudulent documents to gain access to the United States, but the team rarely tested security between them.[164]

What DHS can learn from the physical penetration tests is the importance of top cover and responses to findings. A red team is expected to raise concerns that are unwelcome but require support from top-level management to be implemented.[165] The FAA red team accomplished its mission by annotating over 400 pages' worth of discrepancies, but very few suggestions were implemented until after 9/11.[166]

The lack of implementation was not due to an absence of notification attempts by the red team leadership. Notifications of findings originated through the FAA chain of

---

160 Zenko, 120.

161 Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 111.

162 Eric Lipton, "Testers Slip Radioactive Materials Over Borders," *New York Times*, 2006.

163 Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 111.

164 Zenko, 112.

165 Department of Defense, Defense Science Board, *The Role and Status of DoD Red Teaming Activities*, 2003, 6.

166 Zenko, 123.

command up to the FAA administrator, and eventually the secretary of transportation in 1998. After further recognition that no action was taking place, the red team leadership briefed the DOT inspector general, GAO investigators, and Congressional staffers throughout 1999 and 2000. Finally, the red team leadership resorted to threatening their own employer with formalized a whistleblower disclosure to the Office of Special Council that the FAA was a threat to public safety.[167]

C. RED TEAMING IN PRACTICE: MODELING AND SIMULATION

Not all red teaming is done by actual teams in the field. Often, red teaming is performed using modeling and simulation tools that can be cheaper and yet still effective in determining the effectiveness of security systems.

1. Computational Red Teaming

One form of model and simulation is Computational Red Teaming (CRT), which takes the concept of physical penetration tests with humans and builds it into a computer program capable of creating similar results through repeated simulations. A study performed at Curtin University in Perth, Australia on computational red teaming described the motivations for CRT as the need to discover vulnerabilities, reveal biases, learn about competitors, create a database for events, and unlearn previous practices to learn new ones.[168] The computing and engineering students developed a simulation wherein a red team would attempt physical entry of a building protected by security systems and a set number of blue team guards in order to perform a physical security assessment of the building. The results from each battle are analyzed and used for improvements in both defensive and offensive capabilities.[169]

The model building was designed after a real building, much like DHS would use real terrain features in their model. The simulated red team was given a set number of tools

---

[167] Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 123.

[168] Terence Tan et al., "Computational Red Teaming for Physical Security Assessment," in *The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems*, 2014, 259, https://doi.org/10.1109/CYBER.2014.6917471.

[169] Tan et al., 1.

to bypass the physical barriers, ranging from rocks to electric drills and explosives.[170] The simulation was based on the attack-defend model wherein the red team attempts the first attack, and the blue team adjusts its defenses accordingly before the next round. The blue team made adjustments such as hardening certain walls, repositioning surveillance systems and even hiring more guards to counter each red team move.[171] One surprising outcome of the tests was the ability of the red team to find new paths unseen by the blue team each time the blue team attempted to adapt to previous attacks that did seem obvious. This removal of biases that enables the discovery of possibilities and vulnerabilities while providing a useful database for future events are key reasons why CRTs have been applied in both military and non-military capacities.[172]

The measures of effectiveness from the simulations were derived from the amount of time the red team was delayed on each attempt by the changes that the blue team made to security.[173] Instead of showing delays as the measure of effectiveness, DHS could quickly determine effectiveness through detection rates in a similar model. The results of the test also traced the specific attack paths taken by the red team, which would assist in explaining the need for surveillance technology in a given area. When the system was set to run continual tests without human involvement, the outcomes consistently tapped into a space of unexplored possibilities and discovered multiple vulnerabilities in the physical security of the facility.[174]

2. The Joint Conflict and Tactical Simulation Software

A prominent system used in red team simulations is the Joint Conflict and Tactical Simulation (JCATS) software. JCATS was introduced by the Lawrence Livermore National Laboratory (LLNL) and is the primary ground maneuver simulation model used

[170] Tan et al., 261.

[171] Tan et al., 260.

[172] Tan et al., 1.

[173] Tan et al., 261.

[174] 262.

by the U.S. Army and NATO.[175] While typically used as a training simulator and in military exercises, JCATS has been used for simulating facility and border security scenarios; emergency management response scenarios; and testing new technology effectiveness with integration into military tactics.[176] JCATS accurately simulates all sensor and weapon systems, as well as ground, sea, and air vehicles, including those found along the United States southern border.[177]

A key advantage for DHS using JCATS as a simulation model is its ability to create a playfield based on actual terrain. The software pulls terrain data from the National Geospatial-Intelligence Agency and develops a 3D field that is capable of being viewed accurately down to a one-meter level.[178] The different soil types, vegetation, and bodies of water change the characteristics of the map as well as how a simulated player moves, sees, and shoots.[179] Player visibility is further simulated by the inclusion of day and night inputs, moon or no moon, and tunnels, all of which significantly impact surveillance technology and Border Patrol agent detection capabilities.

D. PHYSICAL PENETRATION VS. MODELING AND SIMULATION

As shown in their 2009 red team study, Sandia National Laboratories is an advocate for both physical and simulated red teaming events, but a cost-benefit analysis is necessary to determine the correct method.[180] Since a red team is typically charged with covering a broad environment and emulating all likely attacks, two limiting factors are time and funding. According to Sandia, the answer to this issue as well as improving red team

---

175 Lawrence Livermore National Laboratory, *Joint Conflict and Tactical Simulation Capabilities Brief*, LLNL-PRES648472 (Livermore, CA: Lawrence Livermore National Laboratory, 2018), 1, https://csl.llnl.gov/content/assets/docs/JCATS_Capabilities_Brief-Update-May2018.pdf.

176 Mark Piscotty and Erica Burleson, *Conflict Simulation Laboratory Quarterly Review* (Livermore, CA: Lawrence Livermore National Laboratory, 2018), 8.

177 Lawrence Livermore National Laboratory, *Joint Conflict and Tactical Simulation Capabilities Brief*, 2.

178 Lawrence Livermore National Laboratory, 3.

179 Lawrence Livermore National Laboratory, 3.

180 Skroch, "Modeling and Simulation of Red Teaming," 2.

effectiveness is modeling and simulation.[181] Modeling and simulation simulates force-on-force interplay by incorporating complex and adaptive human behaviors into a simulated 3D environment and is a tool that can be used by red teams to be more effective when developing plans.[182] When comparing live red teams to red team models and simulation, the pros and cons of each end up making the team more effective when used together. The live red teams perform better in areas such as the breadth of knowledge and creativity, while the models and simulations perform better with wide ranges of possibilities, measurable results, and potentially costs. The final recommendations of the study led to the decision that red team model and simulation 1.) should not substitute human red teams but instead augment them by providing new capabilities and improved analysis; 2.) should be used to capture human red team information and utilize it more broadly at less of an expense; and 3.) use the broad possibilities to direct human red teams where further testing is needed.[183]

Other academic research from Operational Research students has been conducted using simulations to improve border security as well. In his Naval Postgraduate School thesis, Bahri Yildiz used a simulation tool called Map Aware Non-Uniform Automata (MANA) to specifically test the improvements small unmanned aerial vehicles (SUAS) made to border security.[184] Much like what can be done in JCATS, Yildiz replaced the Border Patrol agent kill range to signify a capture and utilized interlinked sensors in the play area to communicate the detection of an alien and initiate the agent's movement to an area.[185] In his final comments, Yildiz found that SUAS technology did improve agent detection and apprehension as well as determine the most effective placement of the assets.[186] The findings also proved that even with surveillance technology improvements,

---

[181] Skroch, 2.

[182] Skroch, 3.

[183] Skroch, 7.

[184] Bahri Yildiz, "Exploration of the Use of Unmanned Aerial Vehicles along with Other Assets to Enhance Border Protection" (Master's thesis, Naval Postgraduate School, 2009).

[185] Yildiz, 39.

[186] Yildiz, 86.

the most crucial asset in border protection is the availability and mobility of individual agents.[187] While detection is important, physical apprehension is what completes the mission. When Turkish Army Operational Research students performed simulation models to analyze Turkey's border, they obtained similar results. The intention for their simulation was to find the most cost-effective border security system that integrated surveillance technology and border patrols. The findings showed that technology such as thermal imaging improved patrol capability, but the most critical factor was effective communication and the ability to respond with mobile patrols.[188]

187 Yildiz, 85.

188 Gökhan Çelik and İhsan Sabuncuoğlu, "Simulation Modelling and Analysis of a Border Security System," *European Journal of Operational Research* 180, no. 3 (August 2007): 1409, https://doi.org/10.1016/j.ejor.2006.04.040.

V. FAA LESSONS LEARNED FOR SIMULATIONS

The following chapter provides a recommendation based on a pre-9/11 FAA red team case study for other red teams to follow in the development, management, and use of outcomes. As briefly described in the previous chapter, the FAA red team was established in 1991 in response to the 1988 bombing of Pan Am Flight 103 that killed 270 people.[189] While their findings were often not employed, the planning, techniques, and team makeup provide a basic outline for a red team to be used at the U.S. southern border.

A. RED TEAM DEVELOPMENT

Red team development starts with choosing the correct players for the team and ensuring they are capable of handling the positions challenges. The selected red team members must also be able to develop a character profiles independently to perform as the backbone of real-world simulations.

1. Choosing the Players

When feasible, red team participants should not be recruited from within the ranks of the organization being tested for two reasons. One, having too much insider knowledge on operations will alter the realistic decision making of the red team; and two, the likelihood of not providing sensitive simulation information to colleagues is low due to the individual loyalties still being tied to performance of their own agency. Individuals with a background in law enforcement or military operations share the same values as those trying to capture them and are less likely to act in the same manner a person trying to cross the border illegally would.[190] In rare cases such as the U.S. Navy Red Cell made up of SEAL team members in the mid-1980s, military members can be successful red teamers. Though the team completed the role of penetrating U.S. naval bases and spreading destruction for

189 Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 116.

190 Stephen Sloan, *Red Teams and Counterterrorism Training* (Norman: University of Oklahoma Press, 2011), 76.

training purposes, internal issues and political fallout led to the program's end in 1992.[191] To back up this theory, the original FAA red team was completely independent of the FAA's regulatory oversight of airport and airline security and was able to plan most of their actions independently.[192] More recently, the inspection teams are directly tied to TSA under the Office of Inspection (OI) and have been labeled as "testing regiments that fit within the confines of bureaucratic needs."[193]

The individuals selected for the red team are also expected to work under challenging conditions. Quite often, red team members are required to work well beyond the 9–5 schedule and must be prepared to go long periods without a resupply of food and water. The teams remain small and are not expected to exceed 12 members at a time, depending on the simulation. The most effective red teams adapt to these challenges and try to feel and think like adversaries by taking transformative steps to take on characteristics of the adversary and develop skills to work as a team.[194] The FAA red team averaged no more than four to five elite agents at a time.[195]

2. Character Development

Current simulation techniques need to focus on the needs of red team creation and how to get red team members into their roles for that team to mirror the behavior, actions, and specific levels of sophistication of a threat group for simulation purposes. Role players in red team operations are often required to write out their character history and ideological views that help them get into character.[196] The FAA red team members received routine top-secret intelligence assessments of terror groups from both the FBI and CIA and used the information as a basis for emulating terrorist activities. The team was mandated to

---

191 Sloan, 77.

192 Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 125.

193 Zenko, 125.

194 Sloan, *Red Teams and Counterterrorism Training*, 91.

195 Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 117.

196 Sloan, *Red Teams and Counterterrorism Training*, 13.

simulate how terrorists were believed to operate and to do covert testing of the security procedures the airline industry was supposed to implement.[197]

To plan effectively for red team role-playing, preparation is essential to understand the feelings and thoughts of the adversary fully. For the character portrayal to seem real, each team member must work to develop their characters independently.[198] By developing individual characters, the red team members avoid the common pitfall of generalizing a role with no specific characteristics and are portrayed as unreal.[199]

Two fundamental principles of character development are observing other people who are similar to the desired character profile and analyzing your character by asking and answering specific questions.[200] Examples of observing people are merely mimicking the posture, gestures, and expressions, as well as the way the person walks, sits and eats. If the red team members neglect this portion of characterizing and are noticed before even moving into position, the validity of the scenario results would quickly be reduced. Much like any law enforcement agency that sees a scenario developing before the official announcement, the Border Patrol agents will soon know that something out of the ordinary is happening and begin preparing in advance. To take the character development even further, red team members need to answer specific questions that effectively build the role. A few example features to question include physical traits, such as posture, gait, appearance and gestures; social characteristics, such as economic status, habitual behaviors, and friendships; as well as psychological traits, such as attitude, motivations, and dislikes; and finally intentions, and how the character will go about achieving the intended goals.[201]

To complete the character profile, highly detailed questions should be addressed as well; examples of these questions are in Table 1.

---

197 Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 118.

198 Sloan, *Red Teams and Counterterrorism Training*, 91.

199 Sloan, 92.

200 Sloan, 92.

201 Sloan, 95.

Table 1. Specific Questions in Character Development[202]

How old am I?	Where was I raised and educated?
What were the social and economic conditions of my family?	What are my hobbies, pastimes, and recreations?
What might I habitually carry with me?	Am I married, and do I have children?
If I work, what is my employment?	What are my skills and training?
How much does my character hide or reveal about myself?	What is my knowledge of weapons and tactics?
How committed am I to my cause?	Have I served prison time?
Am I mentally stable?	If religious, how devout am I?
Has my upbringing, training, or influences caused a cultural hatred?	What are some items indicative of my culture?
What is my native language?	What are the cultural norms to be aware of?

After character development is complete, the red team members move into the practice phase of characterization with improvisation. The following techniques enable individuals with minimal acting experience to become a group of believable characters and therefore create a realistic simulation:

1. Do not contradict the flow of the scenario and remain within the context.

2. Do not contradict another character in the situation. Instead, give up your ideas and go along with what is happening. Failure to do this only prevents forward momentum.

3. Do not break character for any reason other than safety reasons or formal ending of the scenario.[203] By practicing these rules as a group before the

[202] Adapted from: Sloan, 95.

[203] Sloan, 98.

actual scenarios, the red team will become a convincing whole instead of a group of individuals.[204]

3. Conducting the Simulation

The next section discusses simulations from the development phase to the documentation of findings. Simulations are complicated by nature in that they test players and equipment in a manner that is as close to reality as possible; the sense of reality as described in the following FAA scenarios are what make them different than similar exercises.

a. Simulation vs. Exercise

Red team simulations are different than typical exercises in that they ensure a level of realism that mimics real events instead of an intellectual challenge. Exercises are designed to place emotional or mental strain that goes beyond standard functions and are meant to test the participant's capabilities. When these added stressors come into play, it is difficult to evaluate the results of the exercise.[205] In contrast, real-world simulations with red teams create an environment that occurs just as it should, and therefore, the results are not less debatable and thrown into the realm of unrealistic.[206] Exercises require a certain level of scripting, which leads to a higher level of routine actions and reactions, particularly in those who have previously participated in the given exercise. In unscripted simulations, the participant's uncertainty benefits the scenario and better simulates an actual incident.[207]

When it comes to simulations, smaller is better. Largescale exercises quickly lose focus on the original intent of testing and correcting, and move towards long planning phases that lose focus. In many instances, the planning phases are bureaucratized and focus on the responding forces rather than the opposing force and their ability to defeat the

---

[204] Sloan, 101.

[205] Sloan, 17.

[206] Sloan, 17.

[207] Sloan, 19.

countermeasures in place.[208] Planning for the FAA simulations consisted of the red team developing a fifteen to twenty-page operational plan that detailed the movements, timeline, and objectives of the participating red team members.[209]

The higher level of monitoring required by formal exercises is another issue. The presence of too many observers, and the intentional or unintentional control, removes valuable findings from the outcome and takes the control away from the participants. There are few things more disruptive to a scenario than several controllers wearing reflective vests or insignias standing around the environment. The visual interference brought on by the controllers blurs the line between actuality and training so much that reality is essentially eliminated.[210]

Exercises are typically preannounced, sometimes even with a safety briefing and initiated by a declaration of a scenario such as a hostage situation or a bomb being detected. When a simulation begins, there is no preannouncement, and the responding players are required to assess the situation on their own instead of having it predetermined. This test of assessment is not always performed in exercises, and it is common for participants not to know how to react and initiate the remaining phases of response, such as calling for reinforcements. A poor management call for the FAA red team was the inability to self-task and decide where to conduct vulnerability assessments. To complete an evaluation, the team had to receive written permission from senior officials within the FAA. The team was supposed to operate with no-notice inspections per the 1996 requirement, but in reality, they were required to notify U.S. embassies if operating overseas and were never allowed to interfere with daily airport operations.[211] Though the simulations were intended to be unannounced, there were instances of the tests becoming corrupted by FAA administrators tipping off local FAA security managers about upcoming inspections. Written documentation proved instances where equipment such as CTX explosive machines was

---

[208] Sloan, 22.

[209] Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 119.

[210] Sloan, *Red Teams and Counterterrorism Training*, 23.

[211] Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 118.

turned on and operated only during the red team's tests.[212] Other issues occurred, such as FAA administrators interfering with the red team reporting honestly about the reliability of new imaging software programs.

Mandatory exercises tend to focus on "checking the box" to fulfill a requirement set forth by the given agency. Having this mentality prevents the red team from fulfilling its role in playing out highly realistic scenarios. The most essential point Sloan makes is that simulations should be as close as possible to the organizational, physical, emotional, and tactical demands that a real attack has.[213]

b. FAA Red Team Simulations

The FAA red team conducted simulations ranging from simple bag-match violations to high-risk smuggling operations. In the bag-violation simulations, a red team member would check two bags for a flight and never actually board it while documenting the results. Other simple tests would involve the member walking around on the tarmac and waiting for a security guard or ground crew to notice while timing the response. The riskier tests would include smuggling fake bombs, weapons, and even unauthorized personnel onto airplanes.

One example of a bomb-smuggling simulation by the FAA red team was Operation Marco Polo in 1996. The red team planned and conducted 44 bomb-smuggling attempts at the Frankfurt International Airport, and not one was detected. The scenario played out with red team member A placing bagged bomb components onto the x-ray conveyor belt within eyesight of red team member B. As the bag made its way through the x-ray, member B would call member A to walk by the screen as the pack went across. While the bomb components were visible, the distraction took the x-ray attendants' eyes off the screen every time. If any of the bombs were to be detected, the red team members were to provide credentials and inform the baggage handlers that they were part of an assessment. As no bomb components were ever detected, this action was never necessary. The findings from

---

212 Zenko, 121.

213 Sloan, *Red Teams and Counterterrorism Training*, 31.

those scenarios proved that the lack of detection was not necessarily a lack of technological sophistication but simply that the baggage screeners were not watching the monitors.[214]

In another simulation, the red team performed a test for a local Fox news affiliate wherein prohibited weapons passed through Terminal B of the Boston Logan International Airport. The simulation aired on May 6, 2001, and the findings were hand-delivered to the office of Senator John Kerry of Massachusetts, but the red team leadership never received a response. Just a few months later, Mohamed Atta of the 9/11 hijackers performed a surveillance run of the Logan International Airport. On September 11, United Airlines 175 and American Airlines 11 departed Terminal B at Logan International Airport and flew into the World Trade Center towers.[215]

4. Guiding the Simulation

From the time of initiation in 1991, the FAA red team struggled to make an impact on airport and airline security due to a lack of acceptance from the FAA. The group was founded in response to a clear need but had no mission statement or guidance document to provide conduct of the operations, the scope of activities, or the expected use of the findings.[216] After approximately five years, however, the requirement for the FAA administrator to "conduct periodic and unannounced inspections of security systems of airports to determine the effectiveness of such systems" was signed into law in 1996. While this requirement was still vague, it provided a formal need for the FAA red team. The remainder of this section builds on these lessons from the FAA team's experience to suggest practices that red teams should attempt to follow

---

214 Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 120.

215 Zenko, 123.

216 Zenko, 117.

a. Simulation Mission and Intent

The scenario should be created comparable to what the common threat is in the given jurisdiction. Therefore, the target selection, OPFOR group size, tools used, and attack methods need to be formed as realistically as possible.[217]

One question to ask before executing the scenario is what part of the timetable is the simulation focused on: pre-incident, trans-incident, or post-incident?[218] If the red teamers are detected during the planning phase and questioned by the blue team before attempting the desired action, that information may go in the after action report (AAR) as a positive remark, but will not actually provide any results for other measures of effectiveness. Another question is, will the scenario come to a close after the capture or kill has taken place? Or will those who make it to a predetermined goal call in for the end of the scenario? In the case of this the FAA red team, the simulations were only meant to test the screening capability and security of the airport, therefore ending the tests immediately after passing through security.

Another question involves the area of focus in which the scenario will take place. While the details of the play area should be briefed to the red team, other aspects of the test should remain within as tight a circle as possible. For safety concerns, it is essential to pre-brief respondents that there will likely be a scenario, but the exact timeline is not necessary. Merely stating that there will be a test of the region within the month of the initial testing phase may be enough detail rather than providing an exact day for the test.

b. Simulation Organizational Requirements

While over planning is an issue, there is still a requirement for at least some planning. Meetings should be minimal for simple scenarios and occur over the course of a few hours instead of days.[219]

---

[217] Sloan, *Red Teams and Counterterrorism Training*, 33.

[218] Sloan, 34.

[219] Sloan, 38.

Once the operation is planned at the management levels above the border patrol agents on the ground, the simulation operations order must be secure from that point forward. Only the liaison officers finalizing the scenario should know the details; even the red team should be kept unaware of certain information. The actual red team members should not be provided information beyond what their scenario personas would know about the targeted venue. Insider knowledge beyond what a typical adversary would have significantly dictates the red team's actions during both during the planning phase and the actual simulation execution.[220]

Once the simulation plan (SimPlan) is created, the plan must be treated as law enforcement sensitive because of the information contained within it. If a well-devised plan and the results of the scenario were to fall into the wrong hands, the capabilities could be exploited by real-world adversaries. The standard SimPlan contains the following categories described in Table 2.

Table 2. SimPlan Categories[221]

Simulation Information	**Time, date, location, and duration**
Goals	Training to take place and the benefits
Objectives	The delineation of what will actually be evaluated or assessed in the training
Narrative	The general storyline that the simulation will follow (the less rigid the better)
Participants	Who is involved: coordinators, liaisons, red team members, observers, etc.
Command and control	Description of information flow and who has authority in simulation components
Red team plan	Red team mission and how it will be attained
Blue team plan	A declaration from the responding department or standard operating procedures already in place

[220] Sloan, 39.

[221] Adapted from: Sloan, 41.

Simulation Information	Time, date, location, and duration
Equipment list	Checklist of items used in the simulation
Safety and security measures	Description of how the participants and local residents will be protected from harm as well as protocols for emergencies
Evaluation method	The specified feedback loop to critique the outcome of the simulation and pass on findings through the designated channels
Ad Hoc	A recommendation list of final planning considerations for participants that satisfy human needs such as food, water, batteries, first aid kits, etc.

c. Three Phases of Simulation

All simulations will go through three primary phases that should not be skipped or intentionally overlooked: pre-simulation, trans-simulation, and post-simulation.[222]

The pre-simulation phase consists of getting all of the participants in their correct places and in the appropriate mindset. A brief is typically given to provide the basic guidelines and intent behind the simulation. Though the simulation is meant to demonstrate realism, it is essential to incorporate safety protocols, simulation boundaries, actions upon a family emergency, and equipment checks. The participants must also develop an agreed-upon method to start and stop the simulation before moving into positions.[223]

Upon initiation of the simulation, the trans-simulation phase begins. All of the deliberate actions take place during this phase and are usually closely monitored by the scenario controllers. To meet the intent of keeping the simulation as realistic as possible, it may be necessary to remove the ground controllers and have monitors for emergency response only within radio or cell phone contact.

The post simulation begins when the previously agreed upon stand-down order is given, or there is a serious safety incident. The after-action brief should be given as soon

[222] Sloan, 46.

[223] Sloan, 46.

as possible after the conclusion of the simulation. Lessons learned and candid performance critiques are proven to make learning experiences more valuable; as more time passes, the participants get out of character and begin to intellectualize their responses making them less realistic.[224]

d. Safety Measures

The middle ground between a safe scenario and a real scenario is a fine line. Efforts should be made to avoid unnecessary danger without creating an overly bureaucratic checklist approach. A few considerations to brief the red team in a border security scenario are as follows: 1.) No use of booby trap techniques such as fishing line or any weapons to counter the border patrol agents; 2.) Upon capture, you will provide the predetermined documentation of red team affiliation, and there will be no attempt to resist arrest or employ force to escape; 3.) Do not fake an injury as a means of being released; if there is a real-world emergency, further actions must be taken seriously.

e. Measuring the Outcome

Simulations tend to become politicized where the law enforcement officers or military members want to appear well prepared to higher-level authorities and flexible, unpredictable scenarios are not practiced to allow the good guys to win consistently.[225] In some cases, hot washes and lessons-learned meetings are not held with candid feedback. After-action reports must be developed and acted upon with sensitivity to rank and politics left out as much as possible.[226] All participants should be allowed to explain their impressions of the scenario. The debrief should not turn into a session of congratulations on how well the scenario went but focus on the lessons learned to annotate the real outcomes.[227]

---

[224] Sloan, 47.

[225] Sloan, 15.

[226] Sloan, 15.

[227] Sloan, 28.

f. Documentation of Findings

Documentation of findings after completed FAA simulations were shared in the form of written reports to the associate administrator for Civil Aviation Security (CAS). The data within the reports would be shared with the CAS field units responsible for implementing the follow-on remedial actions that should be taken. However, communication was limited, red team findings were never shared directly with airport officials, and the red team was never given responses to the reports.[228]

Where the FAA failed in using the red team was a response to the findings. Many of the exact problems found during the early 1990s were found in the same airports as late as 2001; nothing was being done.[229] A couple of reasons the FAA red team's findings were not being implemented come from 1.) not having a system in place to adequately disseminate and track simulation results and 2.) the administrator of CAS deliberately suppressing or covering up the findings from the tests. As Zenko says, "The pre-9/11 FAA red team is a cautionary tale of the extreme peril of failing to heed a red team's findings."[230]

---

[228] Zenko, *Red Team: How to Succeed by Thinking Like the Enemy*, 118.

[229] Zenko, 120.

[230] Zenko, 127.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSION

This chapter provides a summary overview of the thesis, recommendations for CBP to use for red teaming as an added measure of surveillance technology effectiveness, recommendations for further research, and a conclusion of the research.

This thesis asked the following questions: What technologies are currently utilized for border security and how can their effectiveness be measured? How can red teaming be used to improve existing measures of effectiveness? The answer to the first question is fulfilled by Chapters I–III, which discuss the current technology at the border and the history of methods used to measure them—for example, the recidivism and apprehension rates. The answer to the second question is demonstrated in Chapters IV–V by describing red teaming, the previous uses of red teaming as measures of effectiveness, and through the case study of the pre- 9/11 red team. To answer this question for CBP specifically, this chapter will provide a recommendation based on the research of red teaming discussed in the previous two chapters.

A. SUMMARY

Chapter II provides the background of the immigration enforcement system and the significant changes that have occurred to make it as large as it is today. To demonstrate the need for measures of effectiveness, the chapter describes the ever-changing and long history of measurements of effectiveness at the border beginning as early as 1892. The majority of current measurement methods have been developed over the last two decades, but the unreliable apprehension rate is the one most commonly used. Finally, the chapter describes the importance of measurements and the key components that should be a part of useful measures.

Chapter III provides a contemporary depiction of the southern border by discussing the variety of programs implemented to improve technology in recent history and the primary methods DHS uses to measure effectiveness. The most current program, the Southwest Border Technology Plan, is described in further detail to provide the reader with an understanding of the technology in question. Finally, the chapter describes the methods

of measurement and analyzes the use of their outputs as proper measures of security technology effectiveness.

Chapter IV introduces red teaming as another option for measuring effectiveness and describes its brief history, essential elements, as well as its variety of uses, whether they be physical penetration or computer-generated simulations. While physical penetration provides the more useable results, its expensive and complicated nature make running computer-generated simulations desirable as well, especially when the simulations can be run hundreds of times at little to no extra cost. Analysis of the two categories of uses showed that a combination of the two might provide the most efficient method.

Chapter V goes into further detail on red teaming by describing the creation and implementation of real-world simulations. With the pre-9/11 red team as a case study, this chapter provides details on important aspects of developing a simulation, executing the simulation, and briefing the results to policy makers capable of using the data correctly.

B. RECOMMENDATIONS

As stated in Chapter IV, the steps to effective red teaming include identifying the specific task, building an appropriately empowered team for the task, conducting the simulation, and delivering the findings. This section uses the simulation techniques and case study findings from earlier chapters to develop a concept, based on these three steps, for how DHS could use a red team for measuring border security technology effectiveness.

1. Identifying the Specific Task

The primary goal of red teaming the security technology at the southern border is to provide measurements of the technology's effectiveness in detecting unauthorized movement between the ports of entry and the technology's ability to improve the U.S. Border Patrol Agents apprehension capability. While the most straightforward measurements would come from taking measures both before and after deployment of the technology, this will not always be possible since areas of the border already have various components of the Southwest Border Technology Plan installed, particularly in Arizona.

The primary measurements should be acquired through physical penetration tests; for example, apprehension improvements can be calculated by dividing the number of red team members apprehended during a simulation by the total number of members who attempted entry. Likewise, detection improvements can be calculated by dividing the total number of detected red team members by the total number of members who attempted entry in the designated area and time of the scenario. However, due to the cost and risks involved with physical penetration, further measurements of effectiveness should continue to be performed by using modeling and simulation software in the given area. The ability to generate accurate maps and input specific surveillance capabilities while simulating hundreds of red team attempts can assist DHS in sensor placement or honing in on uncovered areas found by the red teams.

2. The CBP Red Team

The red team would work directly for the Chief of U.S. Border Patrol and consist of 5 to 8 individuals with a designated red team leader position. As mentioned in Chapter V, these select individuals should not come from directly within the CBP organization. The primary reasons for this are to prevent bias toward U.S. Border Patrol and their mission as well as to prevent the red team members from having an unrealistic amount of insider knowledge that would factor into the overall realism of the simulation. However, since the red team will have access to sensitive information, operate on both sides of the southern border, and work in high-risk environments, it will still need to fall within government employee status under DHS or GAO, not a private company.

To prevent corruption of the test results, the red team members should only be provided information that is in line with the expected level of knowledge that an individual attempting a border crossing would have, such as impassable areas to avoid and preferred times of travel. The red team members will be required to perform character development to the level described in Chapter V, which will not only add to the realism of the simulation but also prevent the players from blowing cover before the actual test can begin. Trained Border Patrol agents will quickly spot an out-of-place individual and be able to inform other agents that a simulation is likely to occur or they will follow the individual until they

attempt to cross the border; both options would throw off realistic levels of agent attentiveness and corrupt the test results.

3. Conducting the Simulation

The primary goals of the red team simulations are to maintain a realistic setting that prevents the dismissal of findings due to a debate on realism while also maintaining the safety of the players. Simulations inherently have little to no script and minimal oversight as discussed in Chapter V but with teams operating on both sides of a country's border, extensive coordination with Mexican officials will be required. These simulations will have no controllers to direct scenario phases, and there will be no preannouncement that a simulation is going to take place below the chief of U.S. Border Patrol. These actions are designed to prevent information leakage as happened with the FAA red team simulations. Border Patrol agents will have already been provided a blanket safety brief for the handling of red team operatives and know that the members carry predetermined credentials to verify their status.

The safety of the players will be kept in line with a thorough SimPlan that provides specific details of each simulation that includes designated boundaries and prior coordinated emergency protocols. With no simulation monitors on site, it will be necessary to have emergency services on standby within radio or cell phone contact to respond when needed. Other safety concerns such as the use of improvised weapons to aid in the capture of or escape by the players will be covered in the blanket brief as well.

4. The Findings

To keep the results of the simulations as accurate as possible, the findings should not be swayed by the political nature trying to appear better than the actual findings. To ensure the conclusions stay true to form, the final AAR should initially be reviewed by only the chief of U.S. Border Patrol and the Deputy Commissioner of DHS before moving up the chain of command to Congress. Having a designated routing chain will help prevent the FAA issues of deliberate suppression of results and the inability to adequately disseminate them. The results of the red team simulations will provide apprehension and detection rates that compare the same regions before and after the deployment of

surveillance technology where possible and compare them to findings in areas that already have deployed the technology.

C. RECOMMENDATIONS FOR FURTHER RESEARCH

To further develop the usefulness of red teams and their impact on surveillance technology at the southern border, this thesis recommends further research be conducted into the modelling and simulation side of red teaming and the positive impact it could have on improving the capabilities of the technology. Previous NPS theses on red teaming at the southern border by the NPS Operations Research department students are either outdated due to changes in technology or only focus on one piece of a technology program such as small unmanned aerial vehicles. By inputting the data gathered from this research for the Southwest Border Technology Plan in Chapter III and combining it with a more modern simulator such as JCATS, I believe researchers will uncover improved layouts for the surveillance systems and further improve their effectiveness at the southern border.

D. CONCLUSION

With the consistent failures and the inability of DHS to accurately measure performance with indirect methods, the more direct method of red teaming requires consideration. Red teaming is an approach used extensively in military exercises to find gaps and vulnerabilities just as it is with developing resilience in critical infrastructure. Challenging new technologies with dedicated red teams capable of discovering new ways of penetration is vital and a recognized practice by many organizations to test security technology and measures.[231] By developing scenarios designed to test technology capabilities at the border, DHS may be able to effectively demonstrate areas needing improvement and provide substance for resource requests in those regions. To ensure new technology performs over time, red teaming provides designers information on how technology under development can be circumvented, allowing designers to adapt in the

[231] Brian A Jackson, *Developing Robust Border Security Technologies to Protect Against Diverse and Adaptive Threats* (Santa Monica, CA: RAND Corporation, 2007), 7, https://www.rand.org/content/dam/rand/pubs/testimonies/2007/RAND_CT294.pdf.

process.[232] There is always the potential for adversaries to defeat technology soon after its implementation, and testing the systems before large-scale procurement is a necessity.[233] Aggressive red teams are capable of challenging operational concepts and discovering weaknesses before real threats do.[234]

By developing scenarios designed to test technology capabilities at the border, DHS can more effectively demonstrate areas needing improvement and provide substance for resource requests in the regions that need further development. Red teams offer the user an ability to improve plans and make decisions through quantitative factors such as specific technology and geography, as well as qualitative factors of perspectives and reactions.[235] According to the U.K. Ministry of Defence, red teaming provides benefits in understanding the operational environment; testing a system, plan, or view through the eyes of an adversary; measuring impacts of external influences to adversaries; assessing security and technology through identifying vulnerabilities, risks, and threats; and most importantly, finding additional or enhanced measures of effectiveness.[236] The direct nature of red teaming is also more capable than extensive data collection with meeting the widely accepted performance measurement requirements of being meaningful, clear, and readily understandable; timely and actionable; and stable over time.[237]

---

232 Jackson, 7.

233 Jackson, 9.

234 Department of Defense, *The Role and Status of DoD Red Teaming Activities*, 1.

235 Ministry of Defence, *Red Teaming Guide*, 1–6.

236 Ministry of Defence, 1–6.

237 BPC, viii.

LIST OF REFERENCES

9/11 Commission. *The 9/11 Commission Report*. Washington, DC: Government Printing Office, 2004. http://govinfo.library.unt.edu/911/report/911Report.pdf.

Abrahamson, Reed. "Fixing the Net: The Fall of SBInet, the Rise of Integrated Fixed Towers." *Georgetown Immigration Law Journal* 25, no. 3 (2011): 743–49.

Alden, Edward. *Measuring the Effectiveness of Border Enforcement*. Washington, DC: Council on Foreign Relation, 2013. https://cfrd8-files.cfr.org/sites/default/files/pdf/2013/03/Alden_Border_Security_Testimony_03-14-13%20-%20Final.pdf.

Argueta, Carla N. *Border Security Metrics Between Ports of Entry*. CRS Report No. R44386. Washington, DC: Congressional Research Service, 2016. https://fas.org/sgp/crs/homesec/R44386.pdf.

Blum, Eric. "Further Reflection." Department of Homeland Security. Accessed August 26, 2019. https://www.cbp.gov/frontline/frontline-june-az-technology.

Çelik, Gökhan, and İhsan Sabuncuoğlu. "Simulation Modelling and Analysis of a Border Security System." *European Journal of Operational Research* 180, no. 3 (August 2007): 1394–1410. https://doi.org/10.1016/j.ejor.2006.04.040.

Department of Defense. *The Role and Status of DoD Red Teaming Activities*. Washington, DC: Department of Defense, 2003. https://fas.org/irp/agency/dod/dsb/redteam.pdf.

Department of Homeland Security. *Department of Homeland Security Border Security Metrics Report*. Washington, DC: Department of Homeland Security, 2018. https://www.dhs.gov/sites/default/files/publications/BSMR_OIS_2016.pdf.

———. *Efforts by DHS to Estimate Southwest Border Security between Ports of Entry*. Washington, DC: Department of Homeland Security, 2017. https://www.dhs.gov/sites/default/files/publications/17_0914_estimates-of-border-security.pdf.

Gambler, Rebecca. *Southwest Border Security: Border Patrol Is Deploying Surveillance Technologies but Needs to Improve Data Quality and Assess Effectiveness*. GAO-18-119. Washington, DC: Government Accountability Office, 2017. https://www.gao.gov/assets/690/688666.pdf.

Jackson, Brian A. *Developing Robust Border Security Technologies to Protect Against Diverse and Adaptive Threats*. Santa Monica, CA: RAND, 2007. https://www.rand.org/content/dam/rand/pubs/testimonies/2007/RAND_CT294.pdf.

———. *Developing Robust Border Security Technologies to Protect Against Diverse and Adaptive Threats*. Santa Monica, CA: RAND Corporation, 2007. https://www.rand.org/content/dam/rand/pubs/testimonies/2007/RAND_CT294.pdf.

Lawrence Livermore National Laboratory. *Joint Conflict and Tactical Simulation Capabilities Brief*. LLNL-PRES648472. Livermore, CA: Lawrence Livermore National Laboratory, 2018. https://csl.llnl.gov/content/assets/docs/JCATS_Capabilities_Brief-Update-May2018.pdf.

Lipton, Eric. "Testers Slip Radioactive Materials Over Borders." *New York Times*. 2006.

Ministry of Defence. *Red Teaming Guide*. 2nd ed. Wiltshire, UK: Ministry of Defence, 2012. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/142533/20130301_red_teaming_ed2.pdf.

Morral, Andrew, Henry Willis, and Peter Brownell. *Measuring Illegal Border Crossing Between Ports of Entry*. Santa Monica, CA: RAND Corporation, 2011. https://www.rand.org/pubs/occasional_papers/OP328.html.

Moxley, Mitch. "Better Than a Wall: A New Detection System Can Help Monitor the U.S.-Mexico Border." *Popular Mechanics*, January 28, 2016. https://www.popularmechanics.com/technology/security/a18622/border-control-integrated-towers-system-invisible-wall/.

Office of Inspector General. *CBP's Border Security Efforts: An Analysis of Southwest Border Security Between the Ports of Entry*. Washington, DC: Office of Inspector General DHS, 2017. https://www.oig.dhs.gov/sites/default/files/assets/2017/OIG-17-39-Feb17.pdf.

Orchard, Paulina. *The 2014 Quadrennial Homeland Security Review*. Washington, DC: Department of Homeland Security, 2014. https://www.dhs.gov/sites/default/files/publications/2014-qhsr-final-508.pdf.

Piscotty, Mark, and Erica Burleson. *Conflict Simulation Laboratory Quarterly Review*. Livermore, CA: Lawrence Livermore National Laboratory, 2018.

Poister, Theodore H. *Measuring Performance in Public and Nonprofit Organizations*. San Francisco, CA: Jossey-Bass, 2003. http://www.untag-smd.ac.id/files/Perpustakaan_Digital_2/NON%20PROFIT%20ORGANIZATION%20Measuring%20Performance%20in%20Public%20and%20Nonprofit%20Organizations.pdf.

Roberts, Bryan. *Measuring the Metrics: Grading the Government on Immigration Enforcement*. Washington, DC: Bipartisan Policy Center, 2015. https://bipartisanpolicy.org/wp-content/uploads/2015/02/BPC_Immigration_MeasuringEnforcement.pdf.

Rosenblum, Marc R. *Border Security: Immigration Enforcement Between Ports of Entry*. CRS Report No. R42138. Washington, DC: Congressional Research Service, 2013. https://securityassistance.org/sites/default/files/R42138.pdf.

Schroeder, Robert D. *Holding the Line in the 21st Century*. Washington, DC: U.S. Customs and Border Protection. Accessed May 14, 2019. https://www.cbp.gov/sites/default/files/documents/Holding%20the%20Line_TRILOGY.pdf.

Skroch, Michael J. "Modeling and Simulation of Red Teaming." U.S. Department of Energy Office of Scientific and Technical Information, 2009. https://www.osti.gov/biblio/972439-modeling-simulation-red-teaming-part-why-red-team.

Sloan, Stephen. *Red Teams and Counterterrorism Training*. Norman: University of Oklahoma Press, 2011.

Tan, Terence, Stuart Porter, Tele Tan, and Geoff West. "Computational Red Teaming for Physical Security Assessment." In *The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems*, 258–63, 2014. https://doi.org/10.1109/CYBER.2014.6917471.

Tactical Micro. "The Eagle MVSS." Accessed September 9, 2019. http://www.tacticalmicro.com/products/MVSS_CATALOG.pdf.

University of Foreign Military and Cultural Studies. *Red Team Handbook*. Fort Leavenworth, KS, 2012. http://www.au.af.mil/au/awc/awcgate/army/ufmcs_red_team_handbook_apr2012.pdf.

Yildiz, Bahri. "Exploration of the Use of Unmanned Aerial Vehicles along with Other Assets to Enhance Border Protection." Master's thesis, Naval Postgraduate School, 2009.

Zenko, Micah. *Red Team: How to Succeed by Thinking Like the Enemy*. New York: Basic Books, 2015.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California