U.S. Department of Homeland Security

FY 2023 Annual Evaluation Plan

With honor and integrity, we will safeguard the American people, our homeland, and our values.

About this Report

The *Foundations for Evidence-Based Policymaking Act of 2018* (Evidence Act) requires that the U.S. Department of Homeland Security (DHS) issue an Annual Evaluation Plan concurrent with the Department's annual performance plan describing "significant" evaluations and the related information collections and acquisitions planned for the subsequent fiscal year.

The *DHS FY 2023 Annual Evaluation Plan* describes a subset of the Department's evaluation work for the next fiscal year. These evaluations, designated as significant, are shared with the American public and receive additional resources to ensure successful completion.

As required, the *DHS FY 2023 Annual Evaluation Plan* is published at the DHS public website and at Evaluation.gov with the Department's other Evidence Act plans and reports.

DHS invites feedback on the *DHS FY 2023 Annual Evaluation Plan* and continued collaboration from relevant communities on potential priority questions, data, methods, and analytic approaches that could guide these and future DHS evidence building activities. Public feedback and input may be submitted to: dhslearningagenda@hq.dhs.gov.

Contact Information

For more information, contact:

Michael Stough, Evaluation Officer
Department of Homeland Security
Office of the Chief Financial Officer
Program Analysis and Evaluation Division
245 Murray Lane SW
Mailstop 200
Washington, DC 20528

Contents

Overview

The U.S. Department of Homeland Security (DHS) has a diverse and complex mission to prevent attacks and mitigate threats against the United States and our allies, respond to national emergencies of all kinds, and advance American prosperity and economic security. Since DHS was established from its predecessor agencies in 2003, the Department has continued to expand and mature capabilities to build and use evidence in shaping strategy and operations. DHS generates and uses rigorous evidence from evaluations to inform decisions about programs, policies, regulations, and organizations, better enabling the Department to achieve the most effective U.S. homeland security outcomes and greater accountability to our primary stakeholders, the American people.

The *DHS FY 2023 Annual Evaluation Plan* describes a subset of the Department's evaluation work for the next fiscal year. These evaluations, designated as significant, are shared with the American public and receive additional resources to ensure successful completion. New evaluations are identified annually through systematic consultation with DHS Components, developed with the assistance of and reviewed by internal program evaluators, and coordinated with external stakeholders, including OMB. Many evaluations are designed to address priority questions identified in the *DHS FY 2022-2026 Learning Agenda*. As such, they intend to empower Department decisionmakers to achieve their objectives while fostering organizational learning. DHS evaluations are consistent with relevant legal authorities and privacy, civil rights, and civil liberties protections.

Exhibit 1 lists and describes the criteria DHS considers when making the significant evaluation designation. Evaluations included on in the Annual Evaluation Plan meet one or more of the listed criteria.

Exhibit 1. DHS Criteria for Significant Evaluations

Criteria	Description
Supports the DHS learning agenda	The evaluation is identified as a strategic priority to support decision making in the Department learning agenda
Aligns with leadership priorities	The evaluation addresses leadership priorities at the Component, DHS, or Administration levels
Responds to a mandate	The evaluation responds to requirements or recommendations of the Administration, OMB, Congress, Government Accountability Office (GAO), or the DHS Office of the Inspector General (OIG)
Has potential for agency-wide impact or engagement	The scope of the evaluation activity or the resulting learning affects multiple Components, the entire Department, federal agencies, or other external partners
Has potential for high financial impact	The subject of the evaluation, or evaluand, requires substantial Department funding and/or may pose higher financial risk
Has potential for high stakeholder impact	The subject of the evaluation, or evaluand, affects a large number of entities, including the potential risk for differential or inequitable impacts that should be assessed

FY 2023 Evaluations

Department evaluations follow DHS Directive 069-03, Rev 00, *Program, Policy, and Organizational Evaluations*[1] and may include a range of evaluation types to best answer the questions proposed. To ensure credibility and quality of evidence for learning and decision making, DHS evaluations follow the principles of relevance and utility, rigor, independence and objectivity, ethics, and transparency. These principles align with published Federal evaluation standards.[2]

This *DHS FY 2023 Annual Evaluation Plan* includes evaluations of the following activities and operations:

1. CISA Stakeholder Engagement Division Critical Infrastructure Partnership Advisory Council (CIPAC) National Convening Activities
2. All-Hazards Communications Unit Position-Specific Training and Stakeholder Communication Unit Program
3. DHS Targeted Violence and Terrorism Prevention Grant Program
4. Implementation of DHS Directive 026-06, Rev 02, *Test and Evaluation*, 01 October 2020.

---

[1] DHS Directive 069-03, Rev. 00 *Program, Policy, and Organizational Evaluations* (DHS, 2021)
[2] Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices, M-20-12 (OMB, 2020)

CISA Stakeholder Engagement Division Critical Infrastructure Partnership Advisory Council (CIPAC) National Convening Activities

Lead Organization
Cybersecurity and Infrastructure Security Agency (CISA) Program Analysis and Evaluation (PA&E)

Program Description
Trusted, sustained, and effective partnerships between government and the private sector are the foundation of our collective effort to protect the Nation's critical infrastructure. CISA's Stakeholder Engagement Division (SED) applies a range of legislative and policy-mandated convening authorities and relationship management capabilities to expand, mature, and sustain partnerships with federal, state, local, tribal, territorial, private sector, and international stakeholders to enhance threat and vulnerability information sharing and collaboration to address risks to the Nation's infrastructure.

SED executes CISA's mandated functions as the Executive Secretariat for the President's National Security Telecommunications Advisory Committee, the National Infrastructure Advisory Council, the Cybersecurity Advisory Committee, the Cyber Safety Review Board, and the Critical Infrastructure Partnership Advisory Council (CIPAC) and associated committees described in the National Infrastructure Protection Plan. These forums are the formal structures through which CISA, the broader Federal interagency community, and the President receive guidance, advice, and recommendations to advance national cybersecurity, infrastructure security, and emergency communications goals, and to address systemic risks and issues impacting the national critical infrastructure ecosystem. Further, the CIPAC framework defines how government and private sector entities, organized as coordinating councils, can jointly engage in a broad spectrum of activities to enhance critical infrastructure information sharing, security, and resilience efforts.

SED leads efforts on behalf of the Agency to apply the convening authorities associated with the CIPAC framework to promote collaboration with sector partners. The division uses these convening authorities and associated forums to share threat, vulnerability, and risk information with sector partners, and to guide and influence sector partners to implement and use CISA products and services to manage and reduce risks to the Nation's critical infrastructure.

Purpose and Scope
The evaluation will assess the extent to which CISA's CIPAC convening activities are effective in enabling the sharing of timely, accurate, and useful threat, vulnerability, and risk information between CISA and strategic stakeholder groups; the extent to which these convening authorities increase the access and use of CISA products and services among strategic stakeholder groups; and the extent to which stakeholders achieve measurable reductions in risk within their respective sectors as a result of the information, products, and/or services shared or accessed as a result of CISA's convening activities. CISA will assess the overall perceived reduction in risk from CIPAC participation using retrospective post-participation assessment and collection of activity-level data, beginning in FY 2022.

Resources

The evaluation will be conducted externally through an existing contract mechanism, with execution expected to begin in FY 2023. CISA currently estimates a 12-month period of performance.

Questions

This study addresses *DHS FY 2022-2026 Learning Agenda* Question **G3-Q2**. The evaluation will address the following key questions:

1. Do stakeholders participate in and perceive value and utility of national convening forums and activities as mechanisms of information exchange about security and risk resilience?
    a. To what extent does CISA's SED offer unique opportunities to exchange information with sectors and across sectors within the agency?
    b. To what extent do national convening activities increase awareness and use of CISA's products and services?
    c. What actions have stakeholders and partner organizations taken to manage risks to critical infrastructure?
    d. Which types of actions are widely undertaken by stakeholders?
    e. Which stakeholders were likely to undertake these actions?
    f. What benefits and unintended consequences have these actions had on reducing risk at organizational and national levels?
    g. What barriers to actions have been identified by those who have not implemented any actions to manage risks?
2. What gaps and unmet needs, if any, exist?

Information Needed

Existing data may include: Number of convening activities conducted and stakeholder participation in these activities.

New data to collect or acquire may include:

- Sector stakeholder satisfaction with the CIPAC convening activities as a forum for threat, vulnerability, and risk information sharing
- Stakeholder satisfaction with the utility of the administrative/operational data shared via the convening activity
- CISA staff and stakeholders' perspectives on the value and utility of national convening activities as mechanisms for relationship-building and information exchange across sectors
- How information obtained from the convening activities has been used by stakeholders to improve protective efforts, reduce risk to assets and/or systems, or reduce the costs of existing protection and risk reduction efforts, or by CISA staff to adjust activities to meet identified stakeholder needs

- How information obtained from the convening activities, CISA stakeholder responses, and CISA staff improves CISA programs, improves protective efforts, and/or mitigates critical infrastructure risk
- Unmet needs that could be addressed with improvement to the convenings or other CISA programs.

Design and Methods

The planned study is an outcome evaluation, using a non-experimental retrospective post-participation design, of CISA's CIPAC convening activities. The evaluation may include process evaluation questions but will primarily examine the extent to which participating sector stakeholders believe that the CIPAC convening activities (1) provide threat, vulnerability, and risk information that is timely, accurate, and useful in guiding investment decisions and actions within their sector, (2) provide useful information on relevant CISA threat, vulnerability, and/or risk reducing products and services for their sector, and (3) provide information that results in measurable reduction in risk in their sector.

Primary data sources include CISA staff and stakeholders participating in CIPAC convening activities. Secondary data sources include SED administrative/operational data and artifacts from CIPAC convening activities. Methods of data collection may include quantitative surveys and qualitative interviews/focus groups, with federal government and sector stakeholders and CISA staff as the population of interest. Additional assessment will be conducted during the detailed design phase to determine any applicable sub-group analysis or other sampling considerations necessary to ensure that the diversity of subject matter experts within CIPAC's membership is adequately represented.

Descriptive statistical analysis will be used for quantitative survey data and program administrative data. Qualitative analysis, including subgroup analysis and theme identification, will be used for qualitative data collected through surveys, interviews, and focus groups.

Anticipated Challenges and Limitations

Challenges include recruiting and scheduling interview participants, survey response rates, and insufficient or variable resources for evidence building.

Evidence Use and Dissemination

Understanding the extent to which CIPAC convening activities offer meaningful information exchange and contribute to enhanced trust between CISA and its stakeholders will allow CISA to improve both the execution of CIPAC convening activities and the adoption, use, and impact of CISA products and services. Specifically, the evaluation will be used to identify areas for improvement in the planning and execution of convening activities, in how CISA collaborates and interacts with stakeholders through these convening activities, how CISA shares information through the activities, and how CISA uses the outcomes of the convening activities to prioritize the delivery of products and services that best support information exchange within and across sectors. Evidence building may engage, or be used by, CISA SED and CISA operational divisions

(Cybersecurity, Infrastructure Security, Emergency Communications, Integrated Operations, National Risk Management Center).

All-Hazards Communications Unit Position-Specific Training and Stakeholder Communication Unit Program

Lead Organization
Cybersecurity and Infrastructure Security Agency (CISA) Program Analysis and Evaluation (PA&E)

Program Description
CISA's Emergency Communications Division (ECD) leads the effort, on behalf of Department, to assure vital information exchange, situational awareness, and resource prioritization among emergency managers, first responders, and across public and private sectors is both reliable and resilient during catastrophic incidents and events. DHS provides targeted training, technical assistance, and strategic advice and guidance to prepare and support a diverse cadre to fill a variety of critical emergency communications response roles in federal and State, local, Tribal, and territorial (SLTT) partner organizations. The intended result is a qualified cadre of COMU[3]-credentialed responders that can enable continuity of communications during a catastrophic incident or event.

DHS has conducted prior research and reporting, including the Nationwide Communications Baseline Assessment and the SAFECOM Nationwide Survey. The Nationwide Communications Baseline Assessment seeks to improve understanding across all levels of government on the capabilities needed and in use by today's emergency response providers to establish and sustain communications operability, interoperability, and continuity). The SAFECOM Nationwide Survey is a nationwide data collection effort to obtain actionable and critical data that drives the nation's emergency communication policies, programs, and funding.

Purpose and Scope
The evaluation will assess the extent to which the all-hazards communications unit position-specific training and stakeholder communication unit program created, maintained, and deployed qualified people to fulfill roles in partnering organizations. This evaluation will look at already existing data from the past five years as well as collect new data from stakeholders who were involved between 2019-2023.

---

[3] The term COMU is an umbrella term for the Communications Unit Program. Trained COMU personnel are qualified in one or more of eight position specific specialties: Telecommunications Emergency Response Team (TERT), Incident Commander (INCM), Incident Tactical Dispatcher (INTD), Communications Technician (COMT), Radio Operator (RADO), Communications Leader (COML), Information Technology Service Unit Leader (ISTL), and Auxiliary Communications (AuxCom).

Resources

The evaluation will be conducted externally through an existing contract mechanism, with execution expected to begin in FY 2023. CISA currently estimates a 12-month period of performance.

Questions

This study addresses *DHS FY 2022-2026 Learning Agenda* Question **G3-Q3**. The evaluation will address the following key questions:

1. To what extent has the all-hazards communications unit position-specific training and stakeholder communication unit program activity contributed to enhanced interoperability through the creation, deployment, and support for a qualified cadre to fulfill Communications Unit (COMU) roles in federal, State, local, Tribal, and territorial (SLTT) partner organizations, compared to other factors?
2. To what extent is the all-hazards COMU position-specific training and stakeholder communication unit program activity associated with enhanced and integrated collaborative communications planning, compared to other factors?
3. How effective is the current COMU training plan, and its associated processes and activities, in contributing to officially qualifying/credentialing those SLTT stakeholders who are interested in qualifying?
4. To what extent did training and program attendees apply delivered principles after receiving training to enhance interoperability or integrated, collaborative communications planning?

Information Needed

The evaluation will require the following information for which data is available:

- Number and type of personnel receiving the COMU training and the number of COMU-credentialed responders from training course administrative/operational data
- Rate of states and territories accomplishing national interoperable emergency communications goals
- Percentage of states and territories reporting improvement in developing a formal COMU Process, including formal policies and procedures in place to qualify public safety personnel for COMU positions
- Percentage of states and territories reporting an increase in public safety agencies having access to trained COMU personnel
- Number of jurisdictional conflicts

The evaluation will require the following information for which new data collection is necessary:

- Satisfaction and self-reported learning from trainees
- Self-reported changes in skills and post-participation application of knowledge and skills from trainees

- Incident communications position utilization, activities performed, and communications and IT management structure for incidents and planned events from administrative/operational data (Incident Communications Activity Reports) from state, local, tribal, and territorial responders

Design and Methods

The planned study is an outcome evaluation with an embedded process evaluation study of COMU role-specific training effectiveness and opportunities for improvement. The evaluation will use a non-experimental retrospective post-participation design.

Primary data sources include trainees. Secondary data sources include training course administrative/operational data and jurisdictional Incident Communications Activity Reports. Methods of data collection may include quantitative and qualitative surveys and qualitative interviews/focus groups. An additional assessment will be conducted during the detailed design phase to determine any applicable sub-group analysis or other sampling considerations necessary to ensure that the diversity of subject matter experts is adequately represented.

Descriptive statistical analysis will be used for quantitative survey data and program administrative data. Qualitative analysis, including cross-site analysis of states with individualized credentialing procedures and theme identification, will be used for program administrative data and qualitative data collected through surveys, interviews, and focus groups. Cross-site analysis of states with individualized credentialing procedures may be selected for deeper analysis based on specific regional and/or population criteria to ensure a representative sample.

Anticipated Challenges and Limitations

Challenges include gaps in data access or quality, insufficient or variable resources for evidence building, low response rates due to survey fatigue of incident leaders, and pandemic, disaster, and other emergent interruptions.

Evidence Use and Dissemination

The evaluation will be used by the CISA Emergency Communications Division, Interoperable Communications Technical Assistance Program to engage the Federal Emergency Management Agency (FEMA), federal and SLTT partners, and developers and providers of training and technical assistance to improve plans and build and sustain capabilities to maintain responder entity readiness. External stakeholders include SAFECOM and the National Council of Statewide Interoperability Coordinators (NCSWIC). SAFECOM is a stakeholder-supported public safety communications program administered by CISA that works with federal government entities to improve emergency response providers' inter-jurisdictional and interdisciplinary emergency communications interoperability across local, regional, tribal, state, territorial and international borders. NCSWIC is comprised of Statewide Interoperability Coordinators and their alternates from the 56 States and territories. NCSWIC works with public safety responders in their respective regions, state or territory to create governance structures that promote and enhance interoperable emergency communications.

DHS Targeted Violence and Terrorism Prevention Grant Program Evaluation

Lead Organization

Science & Technology Directorate (S&T) Social Science Technology Centers

Program Description

Starting in December 2015, the Department has made millions of dollars available through the Targeted Violence and Terrorism Prevention (TVTP) Grant Program to help communities across our country develop innovative capabilities to combat terrorism and targeted violence. In FY 2020, through a competitive process, DHS selected 29 projects that provided $10 million in grant funds to SLTT governments, nonprofits, and institutions of higher education for enhancing capabilities of locally based targeted violence and terrorism prevention programs through a whole-of-society approach.

DHS Science and Technology (S&T) Directorate works closely with the DHS Center for Prevention Programs and Partnerships (CP3), who administers these grants, to conduct evaluation on selected grants through an independent contractor. The goal of these evaluations is to:

1. Provide assessment of program implementation and goal achievement among grantees,
2. Conduct outcome and impact evaluations to measure objectives grantees aimed to accomplish at the onset of the grant, and
3. Develop recommendations for DHS so that the Department can measure effectiveness of various locally based TVTP programs while continually improving and understanding the outcomes, impacts, and unintended consequences of future investments in TVTP programs.

This objective assessment ensures the Department and other locally based TVTP programs can continually improve and build upon the lessons learned derived from these evaluations, consistent with the Constitution and other applicable law and policy, in close coordination with DHS privacy, civil rights, and civil liberties experts. S&T plans to support CP3 in and beyond FY 2022 to continue independent evaluations of recipients of future award in support of the DHS CP3 Grants Program.

Purpose and Scope

The purpose of this study is to conduct a portfolio of evaluations for selected grants funded by the TVTP Grant Program with the goal of assessing the program's implementation and goal achievement among grantees while measuring how successful they were in reaching their intended objectives. Depending on the selected grantee and the data they capture, evaluations are likely to be an outcome, impact, or formative evaluation of the locally based TVTP program. These evaluations will allow S&T to provide recommendations to CP3 on how to better define effective programs, create useful tools, and understand the impacts stemming from TVTP programs.

S&T will support evaluations of seven FY 2020 TVTP Grant Program awards. The selection of grantees to evaluate is determined by S&T in collaboration with CP3, based on criteria aimed at

distributing the efforts across their goals, types of outcomes proposed by grantees, and whether an activity or outcome is new or has inherent risk that requires evaluation.

Resources

S&T will award these independent evaluations via contracts, cooperative agreements, or grants to evaluators who are able to design and develop evaluation approaches, protocols, and measurement instruments that will be used with the grantees to measure their overall effectiveness in achieving the goals stated at the onset of the grant program. S&T will initiate evaluations with FY 2022 funds and will continue evaluation efforts and reporting in FY 2023 as funding permits. This effort was not previously included in the *DHS FY 2022 Annual Evaluation Plan.*

Questions

This study addresses *DHS FY 2022-2026 Learning Agenda* Question **G1-Q1.** The overall evaluation addresses the following key questions:

1. What are the priorities and expectations of S&T, CP3, and selected grantees?
2. How effective are the CP3 grantees in achieving their stated objectives?
3. What barriers and/or facilitators to implementation impacted the grantee programs?
4. What additional tools/resources can DHS provide to address implementation barriers or promote facilitators for CP3 grantees?
5. What are the recommendations for policy, research, and programming from this evaluation?
6. How can we better build upon lessons learned from this evaluation?

Information Needed

The evaluation will require information for which data are available, such as recent research findings on community-based prevention programs and indicators of grant outputs (project services and products), participation, partnerships, networking, project-defined outcomes and progress toward achieving them.

The evaluation will require information for which new data collection is necessary, including but not limited to how practitioners are completing their work, what challenges they face, how grantees have adjusted to overcome obstacles, key points of interaction between agency staff members, cross-agency collaborations, and salient community factors. In addition, data may be collected for indicators of short- and mid-term outcomes related to the DHS TVTP program goals of strengthened community resilience through adaptive social capabilities, improved likelihood of referral and self-referrals to community-based interventions prior to criminal conduct and prevented recruitment and deradicalization of at-risk individuals.

Design and Methods

The selection of outcome, impact, or formative evaluation for each of the seven evaluations will depend on the grantee's program design, maturity, and the data they have or can readily collect. All evaluations, regardless of design, will use multiple methods to collect and analyze qualitative

and quantitative data to ensure overall evaluation questions (1-6) are addressed. The following will inform selection of evaluation type:

- Outcome evaluation will be used when the grantee has or can collect outcome data from program participants only.
- Impact evaluation will be used when grantee has or can collect outcome data from program participants and non-participants at two time points such that non-participants serve as a comparison group (for quasi-experimental design) or control group (for experimental design).
- Formative evaluation will be used to support grantees conducting pilot, demonstration, or early-stage activities to determine if they are appropriate, acceptable, and feasible before fully implementing or scaling their activities. These evaluations may examine both process and outcomes, with the purpose of learning and improvement.

Primary data sources include staff and key stakeholders of the seven FY 2020 TVTP Grants Program grants selected for this evaluation, including staff members responsible for cross-agency collaboration, supervision of program staff, and agency-setting within agencies. Secondary data sources include DHS TVTP Program administrative data provided by grantees through the FEMA Non-Disaster Grants Management System, including proposals, logic models, project implementation and evaluation plans, output databases, other performance reporting, and policy documents. Grantees will also provide evaluators with access to their local program administrative data for use in outcome and impact evaluations.

An initial design and development phase will collect grant process information through monthly check-ins, qualitative interviews, quantitative surveys, and observations (made during virtual and on-site visits) of program staff and stakeholders, as well as analysis of DHS TVTP Program and grantees' own administrative data. S&T and the evaluators will determine whether, what, and how more detailed data describing outcomes or impacts can be collected from each of the sites. Evaluators will develop data collection templates, as needed, to request data relevant to each site as well as a set of cross-site measures. Data collected will be used to assess grantee progress and achievement of objectives and will be compiled into a dataset suitable for quantitative analyses to assess progress towards completion of key program goals. Multiple site visits will provide opportunities to directly observe key points of interaction between agency staff members, conduct in-depth interviews with the agency staff members responsible for cross-agency collaborations and with people in supervisory or agenda setting roles within agencies, and document salient community factors (e.g., number, characteristics, and roles of participating community-based organizations).

Descriptive and inferential statistics, as appropriate, will be used to analyze quantitative survey data and administrative data provided by grantees (if applicable). Qualitative data analysis will be applied to narrative or text-based data obtained from administrative data, monthly check-ins, qualitative interviews, and observations.

Anticipated Challenges and Limitations

There are challenges internal and external to DHS for this study. First, given the potential for privacy sensitive activities that may present novel and substantial privacy considerations, the S&T evaluators will closely coordinate with the DHS and S&T privacy offices and perform advance planning to build privacy protections into data collection. S&T will manage the evaluation to ensure timely completion of statutory Paperwork Reduction Act requirements and clearance for new information collections conducted on behalf of DHS. These are necessary compliance requirements, but they can impact S&T's ability to align the beginning of evaluation efforts to the grantee's operational timeline and may limit the study's ability to effectively and thoroughly collect the necessary data. In addition, successful outcome and impact evaluations depend on, and previous evaluations have been hampered by, grantees' ability to align their program logic model with their theory of change and then collect the outcome data needed to assess that theory of change (evaluability). Finally, evaluators rely on but are often not provided sufficient access to grantee administrative data when it has been collected.

As an emerging field of practice, terrorism prevention remains in a state of under-development with ambiguity on definitions, a vagueness on purposes, and a lack of clearly defined measurable goals related to preventing or intervening in violent extremism. While communities have signaled their willingness to grapple with the weighty challenge of our time (e.g., radicalization to violence, terrorism), this is profoundly challenging work. They are experimenting with potential solutions, reflecting on processes and outcomes, and revising programs continuously. This evaluation should be placed in the context of these limitations of the field of terrorism prevention in general.

Evidence Use and Dissemination

Understanding the effectiveness of targeted violence and terrorism prevention practices on outcomes will allow DHS, its federal partners, and grant recipients to better shape policies and programs to implement the national strategic terrorism and targeted prevention strategy. Evaluation findings, suggestions about the effectiveness of the programs, tools to conduct future evaluation activities, and recommendations for future policy, research, and programs will be shared with CP3, federal partners, practitioners, policy makers, and other researchers in the homeland security enterprise. Findings may also be made publicly available through final evaluations reports, peer-reviewed and academic journals, DHS websites, conferences, webinars, and other audience-appropriate dissemination channels. These products have the potential to reach state, local and tribal partners, academia, and even international audiences.

Implementation of DHS Directive 026-06, Rev 02, *Test and Evaluation*, 01 October 2020

Lead Organization

Science and Technology Directorate (S&T), Operations and Requirements Analysis (ORA)

Program Description

DHS Directive 026-06, Rev 02, *Test and Evaluation*,[4] requires that major acquisition programs utilize test and evaluation (T&E) to assist with risk management during the development, procurement, deployment, and operation of products, technologies, and systems. T&E is the process of logically and independently collecting, verifying, assembling, and analyzing data about the products of an acquisition program at key points in the acquisition lifecycle and comparing the data to the program's expected performance parameters. The fundamental purpose of T&E is to provide timely information to senior department management regarding an acquisition product's progress towards meeting mission needs along three dimensions: operational effectiveness, suitability, and resilience.

T&E is delivered through the ITA program described in DHS Directive 026-06 *Test and Evaluation*. At the conclusion of key T&E milestones the DHS Director, Office of Test and Evaluation (DOT&E) provides acquisition decision makers a written assessment of the "adequacy" of an acquisition program's T&E, along with the results of the T&E analysis. In this context, adequacy is the determination that the T&E conducted was of sufficient rigor and objectivity to provide acquisition program leadership with a valid, reliable analysis of the status of the acquisition product. The DOT&E assesses and reports the adequacy of the three dimensions of T&E separately.

At a minimum, a test and evaluation master plan, independent test agent (ITA), operational test and evaluation plans, and letters of assessment (LOAs) are required. LOAs, prepared by the DOT&E, assess the adequacy of independent testing and provide an independent evaluation of operational effectiveness, suitability, and resilience.

Purpose and Scope

The primary purpose of this evaluation plan is to determine if the level of ITA program participation throughout the acquisition lifecycle correlates with the adequacy of T&E as defined in the "Program Description" above: the determination that the T&E conducted was of sufficient rigor and objectivity to provide acquisition program leadership with a valid, reliable analysis of the status of the acquisition product.

The evaluation will consider ITA program support to major acquisition programs and will focus on the adequacy of the independent T&E to inform acquisition program leadership. It will not assess correlations or causality between test results and the impact of the acquired system's performance on supported DHS missions.

Resources

The evaluation will be conducted independently by ORA. ORA staff may be augmented by contractor support. ORA currently estimates a 24-month period of performance in FY 2023 and FY 2024 to determine if there is a correlation between ITA levels of support and the adequacy of

---

[4] DHS Directive 026-06, Rev 02, *Test and Evaluation* (DHS, 2020)

independent T&E. Additional time may be required (FY 2025 through FY 2026) to closely examine specific acquisition programs (cases) with T&E levels of interest.

Questions

This study addresses *DHS FY 2022-2026 Learning Agenda* Question **G6-Q1**. The evaluation addresses the following key questions:

1. How do levels of ITA support differ across DHS major acquisition programs?
2. Why and where does variation in the level of ITA support occur?
3. How do differences in levels of and type of ITA support (independent variables), controlling for other independent characteristics of acquisition programs (independent variables), correlate with ITA assessment adequacy (dependent variable)?

Information Needed

The evaluation will require the following information for which data are available:

Independent variables, including characteristics that differentiate levels of ITA support, and the differentiating characteristics of the specific acquisition programs

Dependent variables, including the DOT&E's finding of adequacy for each of the three dimensions of T&E, operational effectiveness, suitability, and resilience, as reported in the LOA delivered to the acquisition program after key T&E milestone events.

Design and Methods

The planned study is a process/implementation evaluation that will establish the relationship between the level of input resources and the adequacy of T&E processes employed by the ITA. A non-experimental design will be used to explore the correlation between ITA program adequacy (dependent variable), and the characteristics of resources applied to those ITA assessment efforts for a variety of completed acquisition program assessments (independent variables). As appropriate, complementary case study analysis may be performed to deepen understanding of the results of the initial correlational analysis for a selective sample of ITA assessment efforts. The study population includes all acquisition programs that were initiated after the revised ITA process was implemented in DHS Directive 026-06, Rev 01 (May 2017).[5]

Existing data sources include program descriptions, budgets, and timelines of ITA supported acquisition programs, and a series of documents that describe how T&E was planned and implemented for supported acquisition programs. These data sources should be sufficient to describe both the independent and dependent variables in the initial correlational analysis. No new data collection is anticipated at this time; however, additional data collection, including qualitative interviews/focus groups and quantitative surveys, may be conducted if a case study analysis is appropriate.

---

[5] DHS Directive 026-06, Rev 01, *Test and Evaluation* (DHS, 2017)

Qualitative data analysis will be required to translate secondary data (mostly textual data) into variables that can be analyzed quantitatively and to support case study development, if needed. Descriptive statistics will be used to characterize the data for each quantitative variable. Inferential statistics (e.g., chi-square, statistical regression) will be used to examine the correlation between independent and dependent variables. If case study analysis is performed of a selected set of ITA engagements the results will be used to deepen and elaborate descriptive and statistical findings.

Anticipated Challenges and Limitations

Access to sufficient secondary data to distinguish levels of ITA support may be a challenge. If the information is not available in the listed secondary data sources, the evaluators may need to develop proxy measures from other existing data and/or collect new data.

The findings of this study are limited to determining correlation alone, and not the causal relationship, if any between the independent and dependent variables. The study will answer the first-level question of whether sufficient correlation exists to justify more extensive and costly evaluation designs that would support causal analyses and linking ITA outputs to acquisition outcomes, such as cost, schedule, and performance.

Evidence Use and Dissemination

Understanding whether and how ITA participation across the acquisition life cycle correlates with the adequacy of test and evaluation to inform acquisition decisions via the LOA will inform allocation of resources and other improvements to ITA implementation across DHS. Evidence building may engage or be used by the S&T Test & Evaluation Division, the DHS Office of the Chief Acquisition Officer, Component Acquisition Executive Offices, the Joint Requirements Council, Program Accountability and Risk Management office, and Independent Test Agents. Budget and program descriptions may have restricted disclosure and dissemination (either For Official Use Only or Law Enforcement Sensitive).

Appendix A. Abbreviations and Acronyms

CFO	Chief Financial Officer
CI	Critical Infrastructure
CIPAC	Critical Infrastructure Partnership Advisory Council
CISA	Cybersecurity and Infrastructure Security Agency
COMU	Communications Unit
CP3	DHS Center for Prevention Programs and Partnerships
DHS	U.S. Department of Homeland Security
DOT&E	DHS Director of the Office of Test and Evaluation
EC	Emergency Communications
FY	Fiscal Year
ITA	Independent Test Agent
LOA	Letter of Assessment
MGMT	DHS Management Directorate
NCSWIC	National Council of Statewide Interoperability Coordinators
OMB	U.S. Office of Management and Budget
ORA	Operations and Requirements Analysis
QED	Quasi-Experimental Design
RCT	Randomized Control Trial
SED	CISA Stakeholder Engagement Division
SLTT	State, local, Tribal, and territorial
S&T	DHS Science & Technology Directorate
T&E	Test and Evaluation
TVTP	Targeted Violence and Terrorism Prevention

Appendix B. Glossary

Terms used in the evaluation plans are defined below.

Case studies – A case study is an in-depth, qualitative analysis of a single subject or small group of subjects, such as an individual, group (e.g., organization, community, or "site") or event. The analysis integrates data collected through several methods, such as quantitative surveys, qualitative interviews/focus groups, observations, and documents to draw conclusions only about the studied subject(s) and within the given context. Although case studies cannot be used to infer causality or to measure effectiveness, they are often valuable for theory building and developing awareness of factors that affect outcomes.

Descriptive statistics – A set of methods for tabulating summary statistics that characterize cases in a sample data set. Descriptive statistics often focus on quantifying the proportions of various characteristics, major subgroups in the sample, and the shape of the distribution.

Formative evaluation – Formative evaluation assesses whether a program, policy, regulation, or organization approach (or some aspect of these) is feasible, appropriate, and acceptable before it is fully implemented. It may include process and/or outcome measures. However, it focuses on learning and improvement and does not aim to answer questions of overall effectiveness. It can help answer the questions, "Is the program, policy, regulation, or organization appropriate for this context," "Does it feasibly address the identified needs," and "Can it be implemented as designed?"

Impact evaluation – Often used for summative purposes, impact evaluation assesses the causal effect or impact of a program on outcomes by estimating what would have happened in the absence of the program or aspect of the program. This estimation requires the use of experimental/randomized control trial (RCT) designs or quasi-experimental designs (QED) in which another group is compared to program participants. Experimental/RCT designs randomly assign (e.g., lottery draw) persons to either a treatment group that receives the program or policy intervention or to a control group that does not. Quasi-experimental groups identify a program or policy intervention group and comparison group from pre-existing or self-selected groups and not through random assignment. Impact evaluation can help answer the question, "Does the program, policy, regulation, or organization work, or did it lead to the observed outcomes?"

Inferential statistics – A set of methods for drawing conclusions that extend beyond simply summarizing the characteristics of the immediate data. Inferential statistics may specify under what circumstances a sample represents the population (population estimates and confidence intervals). Inferential statistics may also be used to identify statistical relationships by testing hypotheses to determine if differences between two or more groups, changes over time, or associations between two or more variables are not likely to occur randomly.

Observation – An immersive qualitative method for collecting data about people, processes, and cultures, but may be entirely or partially structured (quantitative) or unstructured (qualitative). Structured observations systematically classify behaviors into distinct categories using numbers or letters to describe a characteristic or use of a scale to measure behavior intensity. Unstructured observation records all relevant behavior without a system.

Outcome evaluation – Used for summative purposes, outcome evaluation assesses the extent to which a program, policy, regulation, or organization approach has achieved certain objectives, and how it achieved these objectives. Outcome evaluations use non-experimental designs characterized by the absence of a control or comparison group. Unlike impact evaluation, outcome evaluation cannot discern that outcomes result from or are a causal effect of the program. It can help answer the question, "Were the intended outcomes achieved?"

Primary data sources – Individuals, groups, or organizations from which new data collection is expected, designed specifically for the evaluation.

Process/implementation evaluation – Process/implementation evaluation assesses the extent to which essential elements of a program, policy, regulation, or operation are in place; conform to requirements, program design, professional standards, or customer expectations; and are capable of delivering positive outcomes. t can help answer the questions, "Was the program, policy, regulation, or organization implemented as intended?" or "How is it operating in practice?" In the learning agenda, several evaluations study process-related questions to understand underlying mechanisms of outcomes achievement.

Quantitative surveys – Surveys are predetermined set of questions, often with set response options administered to samples or panels of respondents to cost-effectively compile statistical information about individuals, households, and organizations. DHS uses surveys in different ways. DHS uses surveys to track variables of longer-term interest, as well as to obtain reliable information about conditions through shorter-term studies. DHS conducts low-burden Customer Experience (CX) surveys to gather near real-time impressions of customers' touchpoint(s) or transaction(s) with a government service in terms of trust, overall satisfaction, and experience drivers (e.g., service quality, process, and people, when applicable). DHS also uses surveys of participants in program evaluations to determine their baseline conditions and subsequent outcomes.

Qualitative data analysis – A flexible set of approaches to examine patterns in communicated information. Content analysis may focus on the presence and frequency of concepts—typically words, phrases, or images— or show how concepts are related to each other and the context in which they exist. Thematic Framework Analysis identifies patterns of meaning, or themes. Themes may be determined deductively (themes selected from existing research or theory) or inductively (themes built from the data) to develop patterns. The analysis may examine explicit content of data or examine subtext or assumptions from the data.

Qualitative interviews/focus groups – These qualitative data collections use primarily open-ended questions to converse with an individual respondent or with a small group of respondents simultaneously to collect narrative information about a subject, circumstance, or event. DHS uses this method across evidence-building activities to understand the way people think, their motivation, and their attitudes toward the topic or experience. Although qualitative interviews/focus groups cannot be used to infer causality or to measure effectiveness, they are often valuable tools for theory building and developing awareness of factors that affect outcomes. As such they often complement other evidence building such as surveys, economic analysis, and different types of program evaluation.

Secondary data sources – Existing data, or data collected for purposes other than the specific evidence building activity.