Transportation Security Administration

Test and Evaluation Guidebook

Revision 4

December 4, 2019

APPROVAL:

Terry Caughran

Terry Caughran
Acting Director, Test and Evaluation Division
Acquisition Program Management
Transportation Security Administration

TABLE OF CONTENTS

LIST OF TABLES

TABLE NO. **PAGE NO.**

LIST OF FIGURES

FIGURE NO. **PAGE NO.**

1.0 PURPOSE

The Transportation Security Administration (TSA) Test and Evaluation (T&E) Guidebook (hereafter referred to as the guidebook) describes how TSA T&E organizations, in collaboration with stakeholders, plan and execute an effective and efficient integrated T&E strategy.

The guidebook aligns with TSA Acquisition Qualification Policy, December 14, 2017, and TSA T&E Policy Revision 3, August, 13, 2018. This guidebook meets the spirit and intent of overarching DHS policy as described in the following documents and supporting instructions where appropriate:

- Department of Homeland Security (DHS) Acquisition Management Directive (MD) 102-01, Change 1, May 03, 2019
- DHS MD 026-06 Test and Evaluation, May 05, 2017

The T&E activities described in this guidebook support the DHS Acquisition Life Cycle Framework (ALF). The ALF, depicted in Figure 1-1, identifies the four phases required for system acquisition at DHS and the associated Acquisition Decision Events (ADEs).

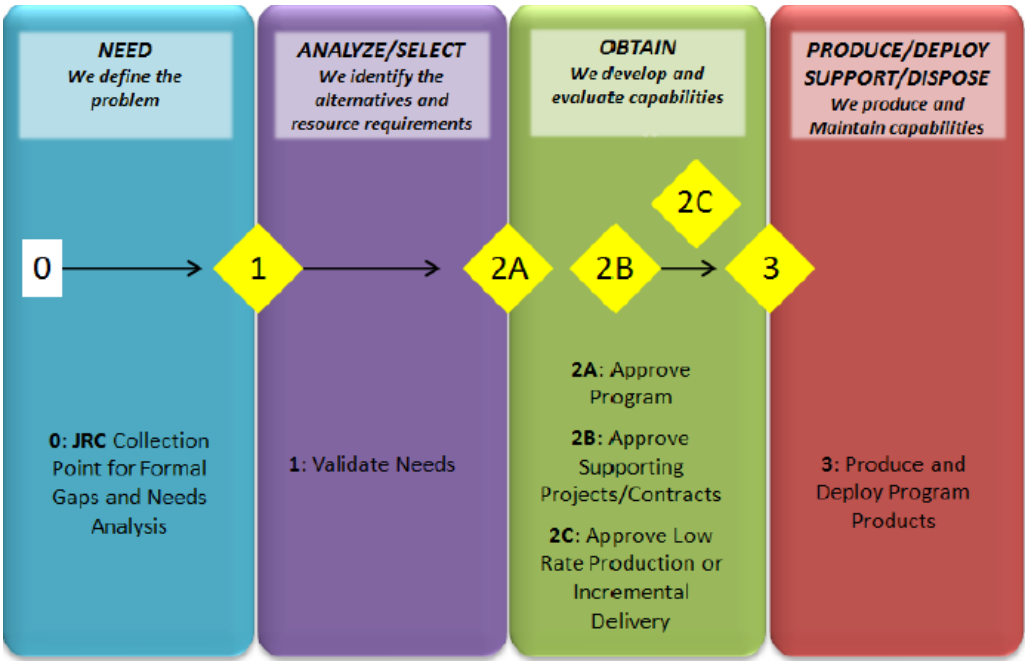

Figure 1-1: ALF

A successful ADE reflects approval to continue to the next ALF phase as well as the commitment of effort and funds. T&E provides data to inform decision makers during the Acquisition Review Board (ARB) and assist the Acquisition Decision Authority (ADA) determine whether to proceed to the next phase of the acquisition life cycle.

1.1 Test and Evaluation Overview

The TSA Acquisition Program Management (APM) performs T&E to determine whether a system, inclusive of technology, operator, and procedures, meets operational and functional requirements. TSA T&E organizations support the T&E of a wide variety of systems from checked baggage and passenger screening Transportation Security Equipment (TSE) to Information Technology (IT) systems. This revision of the guidebook focuses on the T&E procedures for the assessment of TSE.

Directive 102-01 applies to Acquisition Level 1 and Level 2 programs and programs designated as special interest or oversight by the Undersecretary of Management (USM). Typically, T&E in support of TSE acquisition decisions fall into one of these categories. TSA T&E organizations may also assess capabilities not directly associated with an acquisition program (e.g., technical demonstration, proof of concept). The guidebook is primarily focused on the T&E processes associated with acquisition programs. Any explanation of processes not associated with acquisition program T&E is clearly identified when presented.

TSA follows a structured qualification process in support of an acquisition program T&E. The qualification process depicted in Figure 1-2 provides an example of a potential system acquisition path. The program Test and Evaluation Master Plan (TEMP) describes the qualification process, for a particular system, which may be tailored by the Program Office from the illustration in Figure 1-2. Generally, Qualification Testing (QT) and Operational Testing (OT) are required test events for a system evaluation.

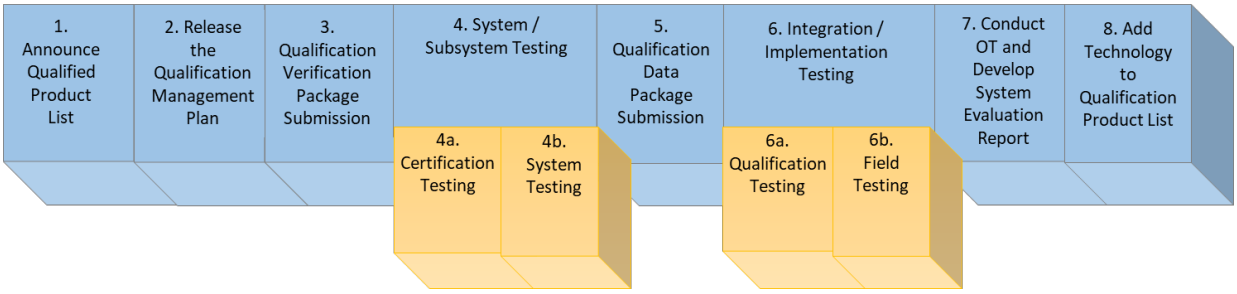

Figure 1-2: TSA Qualification Process for TSE

Throughout this guidebook, there are recurring references to Operational Effectiveness, Suitability, and Cyber Resilience. These terms are critical to the overall goals of T&E and are defined as follows:

- *Operational Effectiveness* is the overall degree of mission accomplishment of a system when used by representative personnel in the operational environment. Operational Effectiveness components typically include, but are not limited to, detection, false alarm rate, and throughput.
- *Operational Suitability* is the overall degree to which a system can be satisfactorily placed in field use with consideration given to interoperability, Reliability, Maintainability, and Availability (RMA), safety, Human-Systems Integration (HSI),

manpower supportability, logistics supportability, training, and other suitability requirements.

- *Operational Cyber Resilience* is the overall ability of the system to protect, detect, counter react to threats, and restore capabilities, to include exploitation potential, and mission impact.

1.2 Test and Evaluation Organization and External Data Sources

The TSA Operational Test Agent (OTA), led by the APM T&E Division Director (hereafter referred to as the T&E Director, OTA), performs the independent evaluation of Operational Effectiveness, Suitability, and Cyber Resilience. The OTA is comprised of the APM T&E Division as appointed by the Director, Office of Test and Evaluation (DOT&E) in its *Designation of Operational Test Agent (OTA) memorandum*, August 18, 2015. The APM T&E Division consists of the branches listed below.

- The Evaluation and Quality Assurance Branch (EQAB) is responsible for the overall evaluation of a system and the final determination of Operational Effectiveness, Suitability, and Cyber Resilience.
- The Operational Test Branch (OTB) plans and conducts OT in support of the EQAB overall system evaluation.
- The TSA Systems Integration Facility (TSIF) Test Branch plans and conducts QT at the TSIF to verify system technical capability in support of the overall system evaluation. The TSIF Test Branch generally verifies non-detection-related functional requirements however, may verify detection-related requirements when needed.
- The Test Infrastructure and Operations Branch (TIOB) supports TSIF Test Branch and OTB test activities by providing facility, test article, Configuration Management (CM), and other related logistical and infrastructure support.
- The Acceptance Test and System Assessment (ATSA) Branch is responsible for providing Acceptance Testing (AT) support activities for T&E events.

The OTA leverages the Transportation Security Laboratory (TSL) Independent Test and Evaluation (IT&E) Division, when needed, to perform Certification Testing (CERT) and provides results as inputs to the Operational Effectiveness determination. The OTA may also leverage data from other data sources in its evaluation under certain conditions. These conditions include instances where the OTA has pre-approved the external data source organization and test plan, and has validated the test results. Additional details on the OTA's use of external data sources are found in Section 3.4 of this guidebook.

Collectively, EQAB utilizes an integrated T&E strategy that leverages test data collected by the supporting test organizations to verify operational requirements.

2.0 TEST AND EVALUATION ROLES AND RESPONSIBILITIES

The primary functional roles responsible for T&E at TSA are listed in the TSA T&E Policy and align to the roles and responsibilities from DHS MD 102-01 and DHS MD 026-06. To minimize redundancy, these roles and responsibilities are not repeated in this guidebook. The following are TSA organizations that provide additional input and support to T&E efforts.

2.1 Requirements and Capabilities Analysis

Requirements and Capabilities Analysis (RCA) supports T&E by providing analysis, threat and risk assessment, and technical support for systems throughout their life cycle. RCA serves as the requirements development organization. As such, RCA personnel develop pre-ADE 2A acquisition documentation such as the Mission Needs Statement (MNS), Concept of Operations (CONOPs), and Operational Requirements Document (ORD).

2.2 Component Acquisition Executive Staff

The APM Component Acquisition Executive (CAE) Staff represents the CAE by providing acquisition assistance to the Program Management Office (PMO) and T&E organizations. The CAE staff guides the PMO through the acquisition process to successfully navigate all ADEs.

2.3 Contracting and Procurement

Contracting and Procurement (C&P), together with the PMO, serves as the liaison between TSA and the vendor. They are responsible for managing the procurement process, coordinating industry days, and the release of qualification documentation.

2.4 User Organization

The user organization is responsible for approving and or concurring with acquisition documentation such as the MNS, CONOPs, ORD, and TEMP with the final approval belonging to the DHS USM. The user organization provides the operator's perspective throughout the T&E process, to include providing input to any observed vulnerabilities, and is consulted upon when determining test articles and any threat simulants used during testing. The user organization at the time of this guidebook revision is the RCA Capabilities Management Division (CMD).

2.5 Information Technology

Information Technology (IT) personnel provide expertise in areas associated with IT and cybersecurity. IT personnel also assist with items such as installation/network connectivity at operational sites and the verification of IT requirements. The IT Information Assurance and Cybersecurity Division (IAD) is responsible for providing the primary Operational Cyber Resilience input to the System Evaluator. This includes activities such as granting the Authority to Operate (ATO) for systems requiring network access, development of penetration test plans, execution of vulnerability scans, execution of the penetration test/adversarial assessment, results analysis, and development of reports.

2.6 Strategy, Communications and Public Affairs

Strategy, Communications and Public Affairs (SCPA) advises the T&E Integrated Product Team (IPT) on any approvals, public/passenger notifications, or other activities that may be required prior to conducting OT. The Operational Test Director ensures completion of these activities/criteria during test planning.

2.7 Chief Counsel

Chief Counsel (CC) provides legal support as needed during the course of an acquisition and when coordinating directly with vendors prior to and after contract award.

2.8 Occupational Safety, Health, and Environment

Occupational Safety, Health, and Environment (OSHE) provides support to the T&E organizations by performing, or supporting the performance of, safety assessments for candidate TSE. OSHE provides findings and recommendations related to any issues discovered.

2.9 Acquisition Program Management Deployment and Logistics Division

The Deployment and Logistics Division (DLD) provides site characteristics and relevant historical data to inform the site selection process and supports the installation and integration of systems at operational sites as needed. DLD is also responsible for coordinating with the PMO and T&E organizations to develop the maintenance concept prior to T&E activities and works with the PMO to develop a TSE deployment strategy.

2.10 Intelligence and Analysis

Intelligence and Analysis (I&A) provides the PMO, requirements sponsor, and T&E organizations with intelligence-based threat information (i.e., threat actors, vulnerabilities, adversary tendencies and preferences) in support of the system threat assessment, overarching T&E strategy, and Threat Inject (TI) testing.

2.11 Training and Development

Training and Development (T&D) is responsible for coordinating with the vendor and the T&E IPT to prepare training curriculum and job aides prior to test execution. T&D personnel review vendor documentation and TSA Standard Operating Procedures (SOPs), develops the curriculum, coordinates with the user organization and each test site to determine training needs, and administers the training.

3.0 EVALUATION STRATEGY AND PLANNING

The following sections describe the procedures for the development of the T&E strategy and evaluation planning in support of system qualification. Efforts in support of non-qualification activities (e.g., technical demonstrations) typically do not include the development of a T&E strategy and proceed directly to test event planning.

The PMO leads the development of the T&E strategy early in the acquisition process once requirement development commences and prior to ADE 2A. The TEMP describes the overarching T&E strategy and is a key reference point for the System Evaluator in their development of the System Evaluation Plan (SEP). Figure 3-1 illustrates TSA T&E strategy and evaluation planning activities between acquisition milestones 2A and 2B.

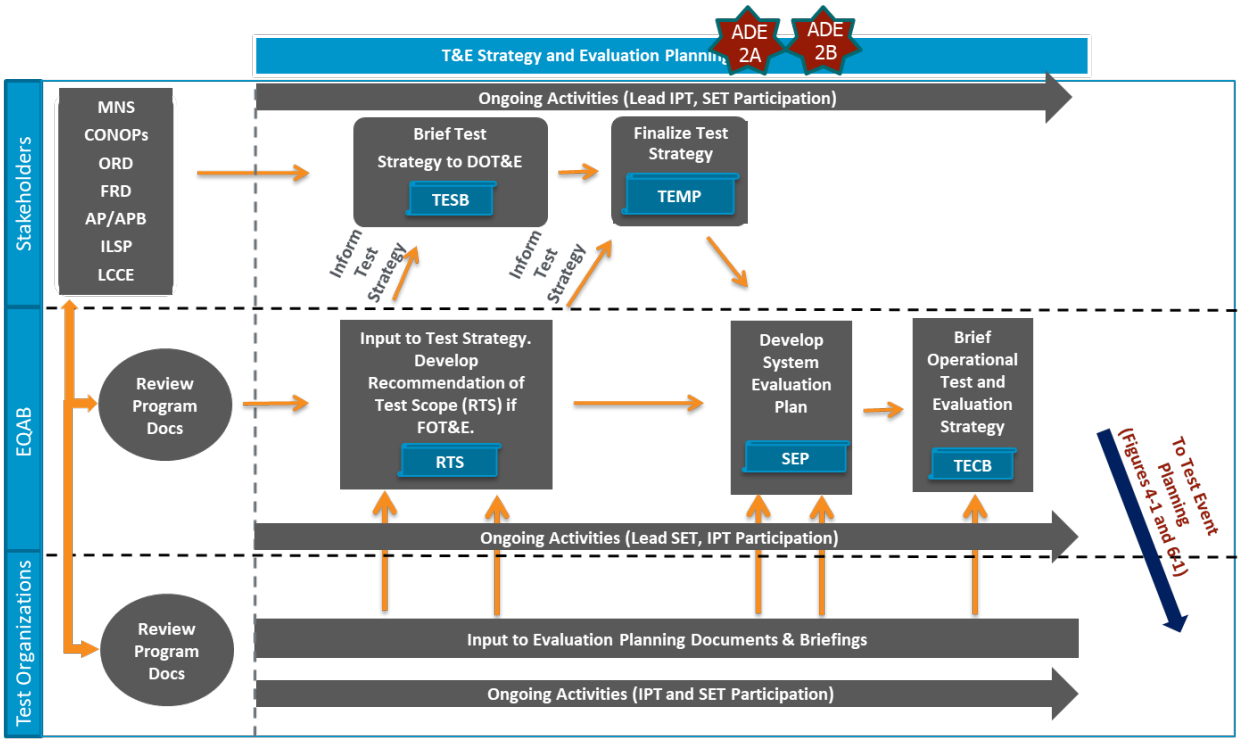

Figure 3-1: T&E Strategy and Evaluation Planning Process Flow

3.1 Test and Evaluation Integrated Product Team

The overall T&E mission begins with the PMO's formation of a T&E IPT that includes representation from the user community, OTA, and other T&E stakeholders. The Program Manager charters the T&E IPT and has the overall responsibility for coordination of T&E activities. The Program T&E Manager leads the T&E IPT and is responsible for working with the T&E IPT to develop the TEMP.

3.2 System Evaluation Team

The primary objective of the System Evaluation Team (SET) is to provide the coordination necessary to execute T&E activities per the TEMP. The System Evaluator is the SET Chair and forms the SET shortly after the PMO forms the T&E IPT or upon successful completion of ADE 2A. The System Evaluator reports the status of SET activities to the T&E IPT. The SET's individual members should:

- Be empowered by their parent organizations to make decisions, and keep their organization informed of issues discussed by the SET.
- Plan, execute, and report test activity results, progress, and risks to the SET chair.
- Provide input, feedback, and concurrence/non-concurrence as required for all T&E documentation requiring SET approval.
- Provide input in support of site selection.
- Provide input and complete tasks as listed in the Qualification Test Readiness Review (QTRR) and Operational Test Readiness Review (OTRR).

SET participating organizations will vary based on the T&E strategy in the TEMP, however, will typically consist of the PMO, EQAB, OTB, TSL IT&E Division, TSIF Test Branch, the user organization, and as needed, from DOT&E, DLD, and IT. The System Evaluator will expand SET membership as needed to ensure engagement of all necessary parties.

3.3 Reviewing and Assessing Requirements

Working within the T&E IPT, the EQAB, OTB, TSL IT&E Division, and TSIF Test Branch provide early and continuous feedback to the PM and RCA on the testability and measurability of operational requirements, to include the Key Performance Parameters (KPPs) as documented in the ORD and technical requirements as documented in the FRD. The test organizations also comment on other requirement documents such as Interface Control Documents (ICDs) which specify functionality for system interoperability.

When reviewing operational and functional requirements, the EQAB, OTB, TSL IT&E Division, and TSIF Test Branch consider the following:

- Whether requirements are clearly defined, achievable, testable, and measurable.
- Available test capabilities and limitations as it relates to the evaluation of requirements.
- Major operational changes or impacts.
- Acquisition strategies and issues that may impact the T&E process.
- Safety and security risks.

3.4 Use of External Data Sources in a System Evaluation

The System Evaluator may consider external data sources (e.g., vendor test data, third party test data) for use in a system evaluation to reduce future testing performed by the OTA (generally QT) and/or add additional confidence to test results.

The OTA's decision to accept an external data source to compliment or possibly reduce the OTA's level of testing is separate from the PMO's decision on whether to accept a vendor into the TSA acquisition process based on external data sources. The OTA first reviews vendor-provided data source information and test plans in the Qualification Verification Package (QVP) and then reviews test data in the Qualification Data Package (QDP) as illustrated in Figure 3-2.

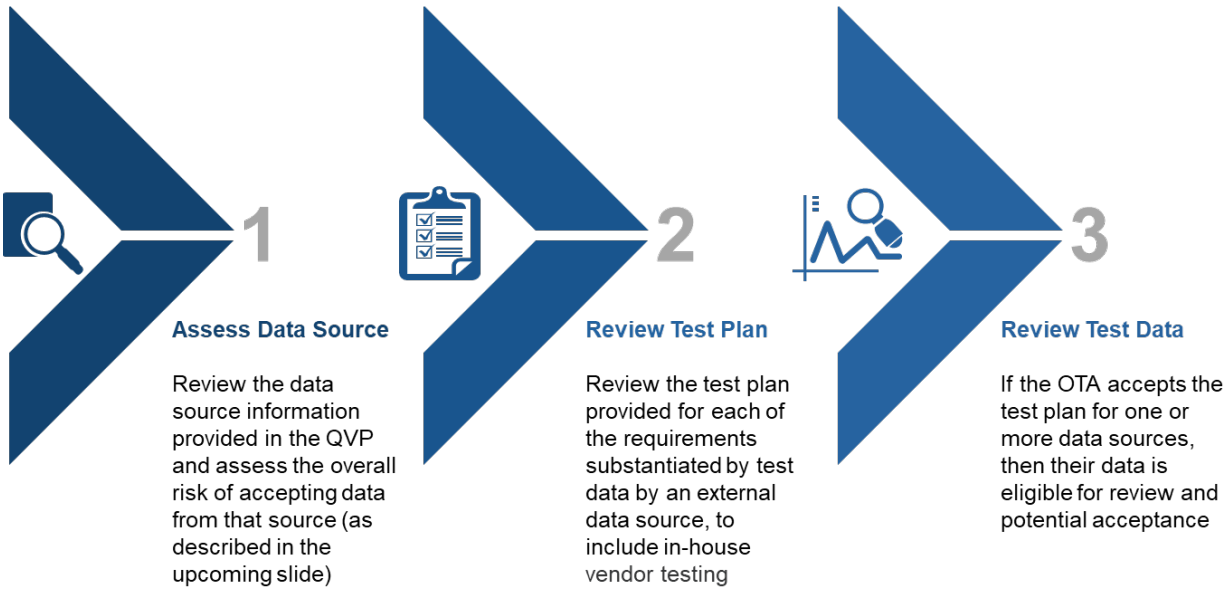

Figure 3-2: OTA Review of External Data

3.4.1 Qualification Verification Package Review

Upon delivery of a vendor-developed QVP, the System Evaluator performs an assessment, as described in the bullets below, of each vendor-proposed external data source for potential use of the data as a primary data source for requirements verification. In the event the PMO does not require a QVP, the following information must be included in the QDP.

- Identifies the external data sources and the criteria for assessing the overall risk of each data source. These criteria include items such as evidence of process maturity, knowledgeable staff, representative operating conditions, relevance of data (e.g., volume of data, spread across sites, machines, conditions, time, etc.), appropriate facilities, appropriate test tools and articles, sufficient and appropriate infrastructure.
- Scores each external data source against the identified criteria. The scoring produces an overall risk score for each data source and an associated Low, Medium, or High rating.

- Reviews the external data source test plan for each requirement and determines whether low or medium risk data source submissions are adequate for verifying requirements. High risk data sources are ineligible to be used for system evaluation purposes.
- Documents a *QVP Data Source Risk Memorandum* that provides the risk rating and establishes what data sources may be used to verify or validate specific requirements. The OTA provides this memorandum to the PMO.

In addition to the above QVP data source assessment, the OTA supports the PMO in the review of any T&E planning, configuration management, or other areas as requested by the PMO.

3.4.2 Qualification Data Package Review

Upon vendor submission of a QDP, the OTA reviews and evaluates the vendor-provided test data and reports. The OTA performs the following activities:

- Reviews the QDP for deviations from the test plans submitted in the QVP, configuration management discrepancies (e.g., testing performed on a non-representative configuration), and any non-compliances that could invalidate the data provided.
- Determines whether the data provided adequately assess the requirement and demonstrate successful system performance/functionality.
- Documents a *QDP Data Source Acceptance Memorandum* that provides the data source risk rating, the requirements with adequate data (if any), and final OTA determination on whether any of the data can be used in support of the system evaluation.

The OTA provides the *QDP Data Source Acceptance Memorandum* to the PMO as an input to the development of the test strategy.

3.4.3 Using the Data Source Acceptance Worksheet

The OTA uses the *Data Source Acceptance Worksheet* to assess external data sources, test plans, and test data. Appendix F includes instructions for using this worksheet.

3.5 Developing Evaluation Criteria

The evaluation of Operational Effectiveness, Suitability, and Cyber Resilience is dependent on the resolution of Critical Operational Issues (COIs) and corresponding Critical Operational Issues and Criteria (COICs). COIs are introduced in the ORD and represent:

- Operational Effectiveness, Suitability, and Cyber Resilience issues that must be evaluated through T&E to determine the system capability to perform the mission.
- Key operational concerns of the user representative, with standards of performance that if met, signify that the system is operationally ready to proceed at ADE 3.

Upon review of the ORD, the System Evaluator may designate one or more Additional Issues (AIs). AIs are key operational concerns that have not been explicitly designated by the user representative as a COI. Each COI/AI is associated with one or more COICs. COICs are the

standards of performance (criteria) for each COI that must be answered by the system evaluation to determine if the system is ready to enter full-rate production.

For each COIC, the System Evaluator derives supporting measures used to assess the system's ability to satisfy the COIC and in turn, the COI/AI. The System Evaluator uses KPPs as criteria that directly support the evaluation of a COI. Characteristics of good COICs are:

- Measurable and can be evaluated
- Avoid terms that could be misinterpreted during the analysis and/or system evaluation
- Associated with one or more measures that reflect the minimal system acceptable performance for entry into procurement (i.e., threshold values)

Figure 3-3 depicts examples of COICs (phrased as questions) and derived measures.

<table>
<tr><th colspan="4">Operational Effectiveness</th></tr>
<tr><td colspan="2">Does the system perform effectively in the target environment?</td><td colspan="2">• Probability of detection
• False alarm rate
• Throughput</td></tr>
<tr><th colspan="4">Operational Suitability</th></tr>
<tr><td>Is the system interoperable with other systems and protocols?</td><td>• System interface and data requirements
• Installation and integration ease</td><td rowspan="2">Will the system be reliable, maintainable and available in the target environment?</td><td rowspan="2">• Operational Availability
• Mean time between failures
• Mean time between critical failures
• Mean time to repair
• Mean down time
• Mean time between maintenance actions</td></tr>
<tr><td>Are Human Systems Integration attributes appropriate for the target environment?</td><td>• Usability
• System design
• Training
• Staffing
• Safety</td></tr>
<tr><th colspan="4">Operational Cyber-Resilience</th></tr>
<tr><td colspan="2">Does the system provide sufficient information security to protect internal components and data in the target environment?</td><td colspan="2">• Privacy
• System integrity
• Accounts and access
• Vulnerability scan reports
• Operational penetration results</td></tr>
</table>

Figure 3-3: Sample Criteria and Derived Measures

Figure 3-4 shows the hierarchical relationship such that measures are used to evaluate the criteria for resolving a COI.

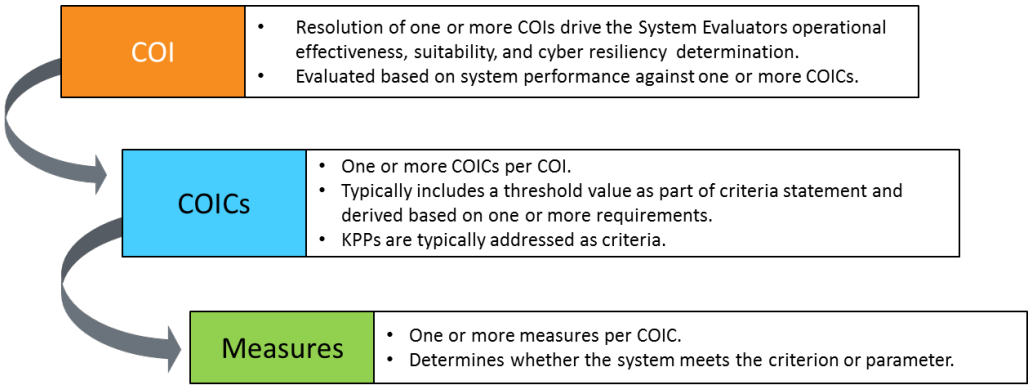

Figure 3-4: Relationship of Common Operational Issues, Criteria, and Measures

3.6 Test and Evaluation Master Plan Development

The PM, with support from the T&E Manager and the T&E IPT, develops a draft TEMP in conjunction with the development of requirements and obtains final approval before ADE 2A. The PM follows the guidance in *DHS Instruction Guide 026-06-001-01, Test and Evaluation Master Plan (TEMP)* when developing the TEMP.

The TEMP is the overarching T&E planning document approved by DOT&E. The TEMP describes the system-specific qualification process, an Integrated Evaluation Framework (IEF) that defines how system requirements will be evaluated, and the sequence of integrated test activities necessary to complete the evaluation. The OTA provides inputs to the TEMP with a specific focus on Section II (Evaluation Framework) and Section III (Integrated Test and Evaluation Activities). The TEMP also includes the plan for assessing Cyber Resilience as discussed in the *Procedures for Operational Test and Evaluation of Cybersecurity*. Figure 3-5 illustrates the general process and documentation hierarchy for a system, beginning with the TEMP and culminating in the development of the SER.

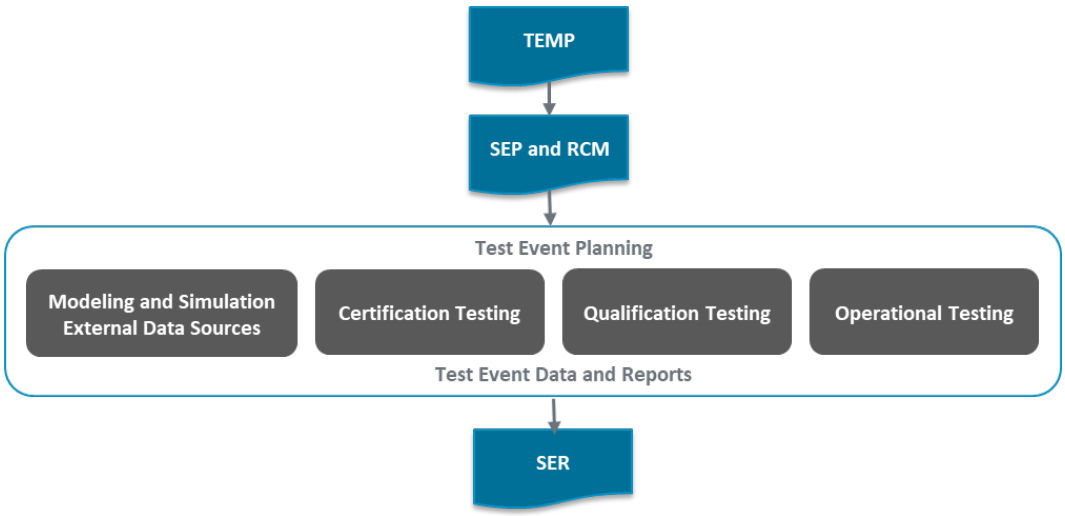

Figure 3-5: Test and Evaluation Master Plan to Reporting Hierarchy

The TEMP is updated as changes to the program (e.g., acquisition strategy, program mission, key requirements) necessitate a change in the overarching T&E strategy. TEMP revisions should receive the same endorsements and approvals as the initial version.

3.6.1 Integrated Evaluation Framework

The System Evaluator develops a Requirements Crosswalk Framework (RCM) and provides the RCM to the Program T&E Manager as an input to the formation of an overarching IEF. The Program T&E Manager then adds critical components to the RCM such as test conditions, Decision Support Questions (DSQs) and Vignettes to form the IEF.

The development of the RCM portion of the IEF begins by correlating FRD requirements to ORD requirements, to measures, to COICs, to COI by developing a Baseline Correlation Matrix (BCM). This traceability exercise enables the System Evaluator to produce a consistent, fully-justified set of operational measures as the foundation for the evaluation.

The System Evaluator, in conjunction with CERT, QT, and OT personnel, then complete the RCM portion of the IEF by identifying the primary (P) and secondary (S) data source for verifying each requirement. The System Evaluator considers all data sources when making the primary or secondary assignment. Data sources include, but are not limited to, OT, CERT, QT, Safety Assessments, vendor QDP, and IAD security assessments. Primary data is generally the source for determining whether a requirement is satisfied. Secondary data provide additional information that can influence the System Evaluator's judgment on whether a requirement has been satisfied.

3.6.2 Test and Evaluation Strategy Briefing and Test and Evaluation Master Plan Approval

The PM and Program T&E Manager brief the T&E Strategy to DOT&E in the form of a Test and Evaluation Strategy (TES) Briefing prior to ADE 2A. It is critical all relevant stakeholder groups (e.g., user, T&E organizations) be present for the TES to answer any direct questions. Upon resolution of any DOT&E identified concerns, the PM routes the TEMP for approval. The CAE concurs with the TEMP for Acquisition Level 1 and 2 programs and approves the TEMP for Level 3 (and delegated Level 1 and 2) programs.

3.6.3 Modeling and Simulation

Modeling and Simulation (M&S) planning typically begins during TEMP development to ensure the appropriate level of M&S throughout all phases of the acquisition life cycle. M&S can be used as an additional data source to supplement QT or OT data but should never replace QT or OT in the T&E strategy. During the planning process, the System Evaluators may use M&S for:

- Prediction of system performance.
- Stimulation or stressing of the System Under Test (SUT).
- Assessment of system capabilities in situations that cannot be field tested.

Prior to use, the System Evaluator should ensure the model is accredited per the guidance in the *TSA Verification, Validation, and Accreditation (VV&A) Policy*.

3.7 System Evaluation Plan

The System Evaluator begins development of the SEP after a successful TES Briefing to DOT&E. While the TEMP provides the overall scope and high-level strategy for T&E, the SEP dictates how the evaluation will be accomplished. The SEP details the measures and analysis methodology the System Evaluator will use to evaluate the system's Operational Effectiveness, Suitability, and Cyber Resilience. Additionally, the SEP:

- Describes how CERT, QT, OT, and external data sources support the evaluation of COICs and corresponding evaluation methodology.
- Contains the RCM portion of the IEF that correlates requirements to COICs and documents the data source(s) for each requirement.
- Includes T&E limitations. These limitations must not rationalize why the test team will not assess requirements, but rather, they present a suitable alternative test methodology.

The System Evaluator typically updates the SEP whenever the PM updates the TEMP and may initiate an update independent of the TEMP if significant program changes occur. The members of the SET provide feedback prior to final SEP approval. The System Evaluator gains concurrence on the SEP from the EQAB, OTB, and TSIF Test Branch Manager, and final approval from the T&E Director, OTA.

3.8 Test and Evaluation Concept Briefing

Once the TEMP is approved by DOT&E and the SEP has been drafted, the System Evaluator presents the evaluation plan to DOT&E in the form of a Test and Evaluation Concept Briefing (TECB). The briefing describes the system, the COICs, measures, test objectives, overall test methodology, CERT/QT/OT entrance and exit criteria, criteria for test sites, and other details that support OT. The System Evaluator presents the TECB to the DOT&E approximately 120-days prior to the start of OT (if schedules permit).

3.9 Determining the Scope of Testing After a System Change

The System Evaluator, and/or support staff, are participants on the Change Control Board (CCB) distribution and review summary communications detailing system changes. Upon vendor submission of a system change package and upon PMO request, the System Evaluator assesses the change to determine whether previously evaluated system performance has the potential to be impacted. The System Evaluator develops a Recommendation of Test Scope (RTS). The RTS is used to determine whether the results of a previous evaluation are impacted by the change and any resulting regression testing to include Follow-on Operational Test and Evaluation (FOT&E). During RTS development, the System Evaluator evaluates the changes against the IEF, specifically against each of the previously evaluated measures to determine if regression testing is required.

The System Evaluator drafts the RTS and distributes it to the SET for review and feedback. The System Evaluator then sends the RTS to the EQAB Manager for review and then to the T&E Director, OTA for approval. The System Evaluator then provides the RTS to the PMO to inform a program-level risk assessment.

3.10 Analysis and Reporting

Figure 3-6 illustrates the process for incorporating test event results, and developing evaluation level reports in support of an acquisition decision.

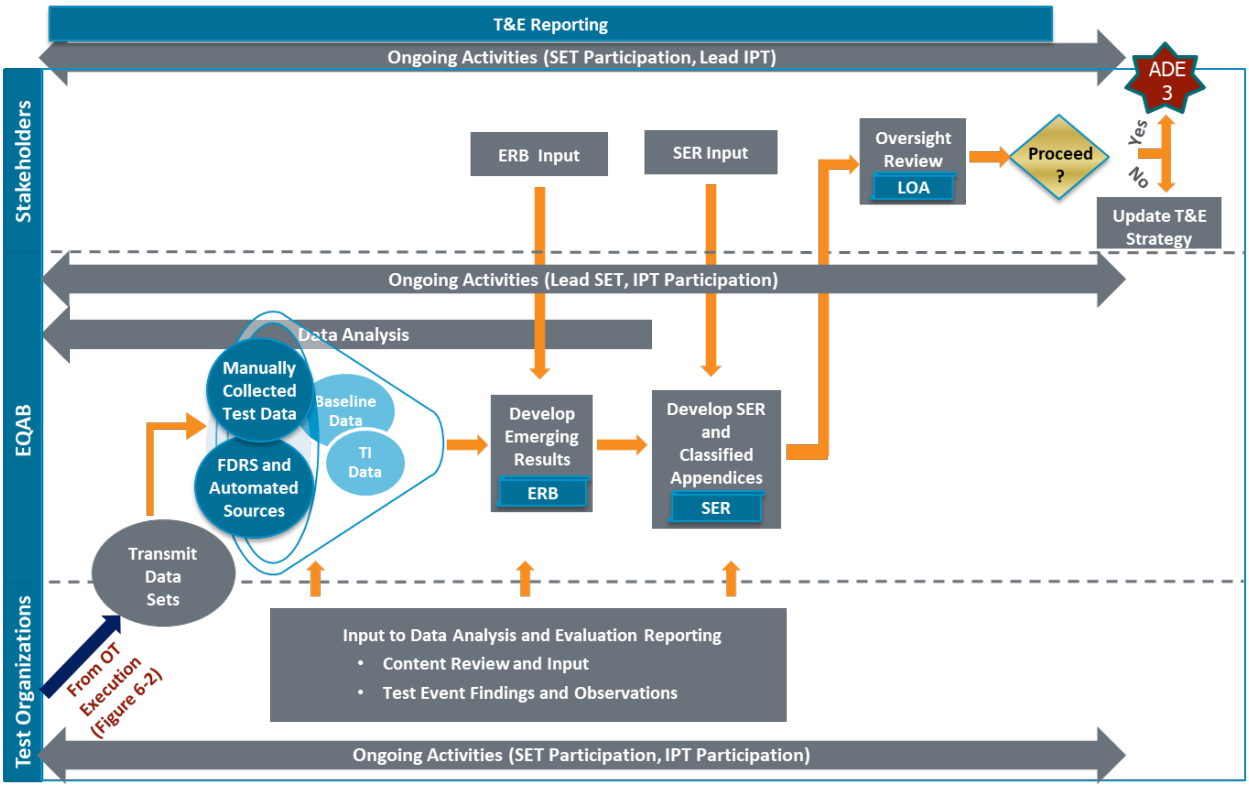

Figure 3-6: Evaluation Reporting Process Flow

3.10.1 Test Event Reporting in Support of the System Evaluation Report

Once test events have concluded, the appropriate TSIF Test Director or Operational Test Director sends the final data set, test findings, and any report created to the System Evaluator for reference in the SER.

3.10.2 Quick Look Report

The System Evaluator prepares an OT Quick Look Report (QLR) on a case-by-case basis as documented in the SEP, Operational Test Plan (OTP), and/or OAP, when information is required prior to a final SER in support of acquisition decisions. The OT QLR provides preliminary findings based on OT data. The data used to compute these results may not have been fully

analyzed. Therefore, prior to distribution, the System Evaluator ensures the OT QLR contains appropriate disclaimers, cautioning its use for making key decisions.

When required, the System Evaluator develops the OT QLR upon completion of OT with a delivery date of approximately 10 calendar days after the final DAG meeting. The System Evaluator gains concurrence from the EQAB and OTB Managers, and final approval from the T&E Director, OTA prior to final delivery of the OT QLR.

The TSIF Test Branch may also create a QT QLR as discussed in Section 4.3.7.

3.10.3 Emerging Results Briefing

The System Evaluator develops and delivers an ERB to DOT&E within 30 days of the last day of OT. The scope of the ERB is for OT only and does not include any information concerning QT or other data sources the System Evaluator will use in the final evaluation. The ERB includes unexpected test limitations, deviations from the approved OTP, and any major incidents encountered during OT. The System Evaluator should include the number of samples collected for each measure and compare quantity to the planned sample size. The ERB does not make final conclusions nor does it include any metrics or results.

The System Evaluator gains ERB concurrence from the EQAB and OTB Managers, and final approval by the T&E Director, OTA prior to delivery to DOT&E.

3.10.4 Modeling and Simulation Results

If M&S is required per the TEMP, the System Evaluator presents the M&S results in an appendix of the SER. A typical M&S results appendix may include a clear and defined modeling and analysis approach to include model validation, a current performance baseline, the predicted impact, and other associated findings.

3.10.5 System Evaluation Report

The System Evaluator delivers a signed SER 60 days after the final DAG meeting and 45 days prior to any scheduled ARB unless an alternative agreement has been reached with the Program Office and DOT&E. This allows DOT&E sufficient time for their Letter of Assessment to be issued prior to the ARB. The SER (referred to as an OT&E Report in DHS Directive 026-06) provides an independent evaluation of a system's Operational Effectiveness, Suitability, and Cyber Resilience and informs decision makers prior to an ADE 3 review. SER results are corroborated using data from all credible data sources. The SER addresses and answers the COIs, COICs, KPPs, and additional evaluation focus areas as described in the SEP. When addressing these areas, analysts generally leverage hypothesis testing. The confidence level used is typically 90 percent. Statistical methodologies used in data analysis are described in the *Test and Evaluation Statistical Methodology Policy for Probability of Detection and False Alarm Rate*.

During the evaluation of system performance, it is possible for a point estimate to meet a requirement threshold while the lower/upper confidence interval does not. As such, the System Evaluator will use the following language when evaluating performance against a threshold:

- *Requirement demonstrated with statistical confidence* - Point estimate and entire confidence interval meets or exceeds the threshold.
- *Requirement demonstrated without statistical confidence* - Point estimate meets or exceeds the threshold, however, the lower or upper bound confidence interval fails to meet the threshold.
- *Requirement not demonstrated with statistical confidence* - Point estimate and entire confidence interval fails to meet the threshold.
- *Requirement not demonstrated without statistical confidence* - Point estimate fails to meet the threshold, however, threshold is within the confidence interval.

The System Evaluator adjudicates SET feedback prior to finalizing the SER. The T&E Director, OTA approves the SER after concurrence by the EQAB and OTB Managers. When appropriate, the System Evaluator may also create a classified SER appendix detailing the results of TI testing. The System Evaluator creates the classified SER appendix and distributes to the same approving parties as the unclassified SER. Upon T&E Director, OTA approval, the System Evaluator forwards the SER to DOT&E for review. DOT&E reviews the SER and provides an LOA approximately 30 days after receipt of the SER documenting concurrence/non-concurrence with the SER ratings, conclusions, and an assessment of the overall conduct of the test.

3.10.6 Operational Assessment Report

The Operational Assessment Report (OAR) is the System Evaluator-prepared final report for Operational Assessments (OAs). The System Evaluator develops and delivers an OAR approximately 45 days after the final DAG meeting. The OAR provides an initial, potential rating of system Operational Effectiveness, Suitability, and Cyber Resilience prior to the conduct of an OT&E. The T&E Director, OTA approves the OAR after concurrence by the EQAB and OTB Managers. Upon T&E Director, OTA approval, the System Evaluator forwards the OAR to DOT&E for oversight review.

3.10.7 Field Data Collection Activity Report

The Field Data Collection Activity Report (DCAR) is the System Evaluator-prepared final report supporting Field Tests and Field Data Collection Activities (FDCAs). These reports provide system performance against the measures described in the Field Data Collection Activity Plan (DCAP). These reports do not include an Operational Effectiveness, Suitability, or Cyber Resilience rating. The System Evaluator considers SET input when finalizing DCARs. The System Evaluator sends DCARs to the OTB Manager for approval and to the EQAB Manager (when delegated by the T&E Director, OTA) for final approval.

4.0 CERTIFICATION AND QUALIFICATION TESTING

TSA performs CERT (as needed to verify system detection) and QT of final, vendor-submitted systems as described in the TEMP. These events provide a final, or near final, verification of system conformance to functional specifications, in a laboratory environment, prior to OT. These events are further described below:

- CERT verifies threat detection performance, such as probability of detection and false alarm rate. TSA systems designed to detect explosive materials typically undergo CERT at the TSL prior to QT to certify the systems meet minimum performance standards. Once successful, TSA certifies system detection standard compliance by issuing a certification letter to the vendor for the system configuration tested.
- QT provides confirmation that the system complies with the technical requirements defined in the FRD. QT typically verifies the FRD requirements not related to detection of explosive materials. QT reports must align with the TEMP and SEP and serve as an input to the SER.

Generally, the TSL IT&E Division performs CERT and the TSIF Test Branch performs QT; however, these events can be performed at other controlled environments as determined in the T&E strategy. The TSIF Test Branch may also provide detection-testing support for some TSE when non-explosive materials are used. A well-planned and executed CERT and QT program supports the determination of whether a system meets user needs and can proceed to OT.

This section of the T&E Guidebook primarily describes QT processes and procedures performed by the TSIF Test Branch.

Additional objectives of QT include:

- Perform engineering analysis on vendor-submitted change requests for requirement and CM compliance practices. This information is used to establish the as-tested baseline configuration leading to the Government-approved Master Configuration Item List (MCIL).
- Identify and describe system design, usability, and other related risks.
- Working with OSHE, assess the safety of the system to ensure its safe operational use. The intent is to ensure that systems are sufficiently free of hazards to prevent harm or injury to test personnel, typical users and the public in operational environments.
- Support assessments to verify system-of-systems interoperability. This includes coordinating with other organizations to facilitate assessments. For example, the TSIF Test Branch would coordinate with IT IAD to perform Cyber Resilience scans and assessments.

4.1 Qualification Test Events Overview

QT events support system qualification as described in the TEMP. TSIF Data Collection Activities support all other requests for data and results. These test events do not directly inform a system acquisition/qualification. These events are listed in Table 4-1. Note, for all tables listing

T&E evaluation areas, planning levels are based on the complexity of requirements, stakeholders involved in the assessment, test conditions and resources, and/or lengthy duration.

Table 4-1: Qualification Test Events

Test Type/Purpose	Additional Detail	Sponsor	Approximate Test Length
QT Events			
QT The formal verification of system performance against FRD requirements and provides confidence that the system design will satisfy the desired capabilities in an operational environment.	• Provides input for an acquisition decision • Confirm compliance with technical requirements • Provides performance metrics to advise on system readiness to enter OT	PMO	2 weeks – 2 months (Scope Dependent)
Regression QT Regression Testing verifies vendor implementation of system updates to ensure the change does not impact performance from previously completed QT and/or OT.	• Configuration change requests made by the vendor • Determine adequacy for fielding to operational fleet or to enter a FOT&E	PMO	1 week – 2 months (Scope Dependent)
Non-QT Events			
TSIF Data Collection Activities Test activities that support any requests for data collection outside of a formal QT. These can include integration tests, proof of concept testing, and ad-hoc requests to determine capabilities and limitations.	• Assess/validate CONOPs • Forensic testing • Provide insight into the human factors, usability, and design characteristics of a system • Provide input into SOP development	PMO or other requesting organization	1 day to 1 month (Scope Dependent)

4.2 Qualification Test Planning

The TSIF Test Branch performs a number of planning activities to prepare for a QT event. Central to these planning activities is the review of vendor-submitted QVPs and QDPs and the development of the QTP. QDP acceptance is a pre-requisite to the commencement of QT at the TSIF. Figure 6-1 illustrates the various activities taking place during QT planning by various stakeholders.

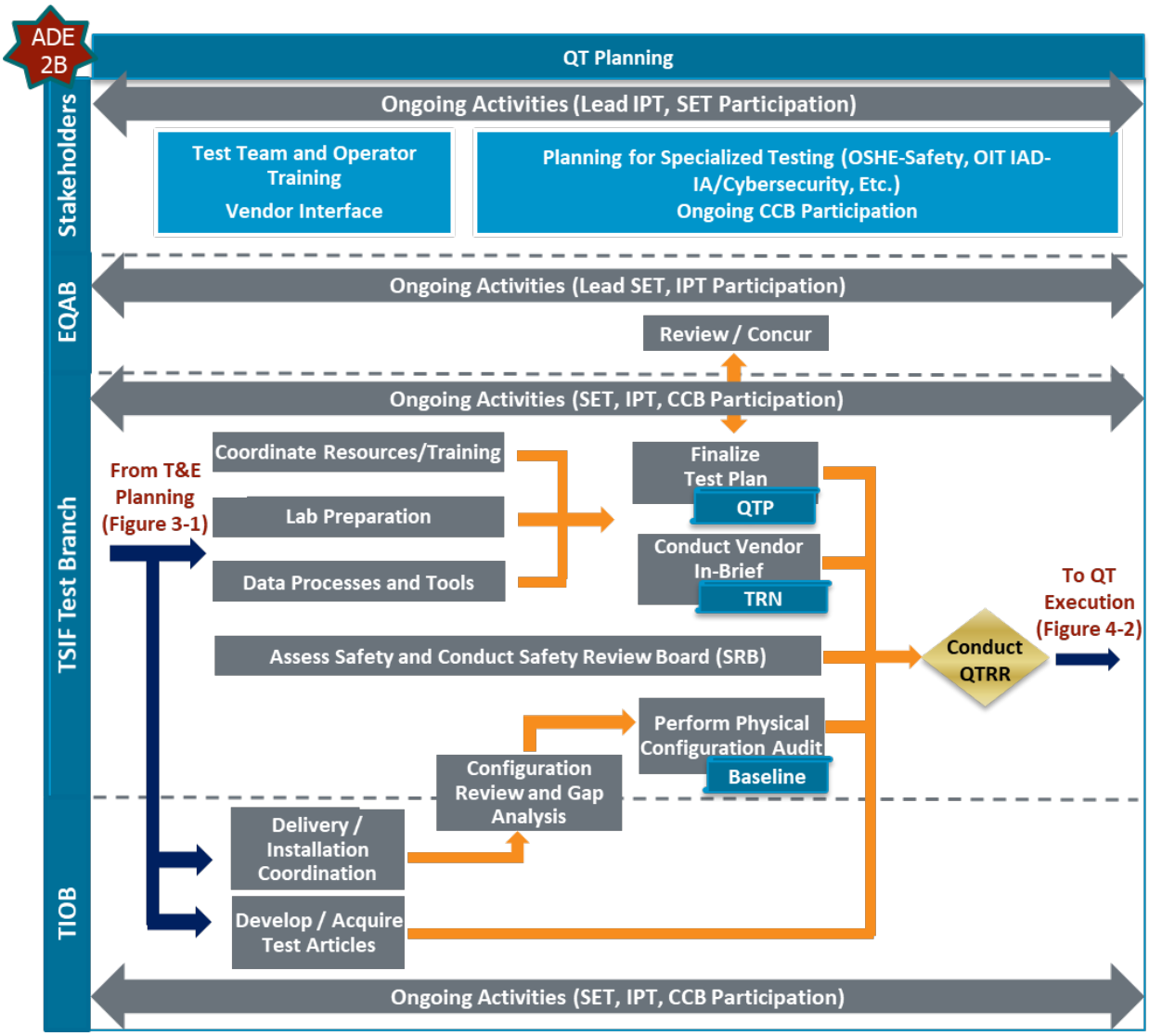

Figure 4-1: Qualification Test Planning Process Flow

4.2.1 Planning Documentation

The TSIF Test Director creates a QTP based on the TEMP, SEP and requirements allocated to the QT organization in the IEF as a primary or secondary data source. The System Evaluator primarily allocates FRD requirements to the TSIF Test Branch; however, ORD requirements may be assigned for verification on a case-by-case basis. Dependent on the scope of the QT, FRD requirements may be allocated to specialized groups (e.g., IT IAD, OSHE). The planning documentation for the activities performed by the specialized groups is typically developed under separate cover by those groups. The QTP describes the details of the test to include a schedule of events, safety considerations, training required, test article details, required personnel, test scenarios, data collection tools and test scripts. The TSIF Test Director gains concurrence from the System Evaluator, affirming the QTP will adequately support the

evaluation prior to initial approval by the TSIF Test Branch Manager and final approval by the T&E Director, OTA.

Unless otherwise directed by the TSIF Test Branch Manager due to the significance of a test event, the TSIF Test Director approves the TSIF DCAP.

4.2.2 Establishing the Configuration Baseline Prior to Qualification Testing

The TSIF Test Branch CM Lead, with support from the TIOB CM Specialist, obtains contractual CM documentation that details critical components of the SUT. If the TSL IT&E test team certified the TSE prior to QT, an electronic copy of the TSL IT&E test team's approved Developmental Baseline (DBL) and CM audit results are provided.

The TIOB CM Specialist provides full copies of all Developmental Engineering Proposals (DEPs) approved since the final date of certification for TSIF test team review. DEPs are TSA required documentation that a vendor completes to fully document any system changes. The review of these DEPs primarily entails assessing the impact of the changes on the requirements.

The TSIF Test Branch CM Lead, with support from the TIOB CM Specialist, establishes the As-Tested Configuration Baseline (ATCB) prior to QT. The TSIF Test Branch CM Lead compares TSA requirements against the Test CIs that may be in the vendor's system design and determines critical functions by considering the requirements, and the potential consequences if any are compromised. The TSIF Test Branch identifies Test CIs for the critical components of the system, including hardware, firmware, software, databases, manuals, and settings. The test entities must ensure the CIs adequately reflect the technical, functional, and operational requirements and document instances where gaps exist in test reports. Differences between "as built" and "as documented" configurations are noted and provided to the Contracting Officer Representative (COR) to resolve with the vendor.

The TIOB CM Specialist utilizes the ATCB to establish the Developmental Product Baseline Configuration Audit (DPBCA) list. The PMO is the final approval authority for the Developmental Product Configuration Baseline. The TIOB CM Specialist maintains this baseline throughout the developmental process with all future Developmental Change Requests processed as described in Section 4.3.2.

The TSA APM Configuration Management Plan (CMP) contains additional information on the establishment of the CM Baseline.

4.2.3 Qualification Testing Readiness Notification

The QT Test Readiness Notification (TRN) describes the ground rules for vendor presence and participation during a QT event. The TRN serves as an acknowledgement between TSA and the vendor that the system is ready for QT for the scheduled test period. The following provides additional details of the TSIF Test Branch TRN.

The TRN includes the rules of engagement for a vendor with respect to system access and maintenance to ensure test integrity. Typically, the vendor signs the TRN immediately following system installation once the system is operating properly as intended by the vendor. There are several activities during QT where a vendor's presence may be required. Prior to the start of test,

the TSIF Test Director, through coordination with the PMO, may authorize the vendor to support one or all of the following activities:

- *Training* – Vendor onsite to provide training to the TSIF Test Branch.
- *Configuration Audit Support* – Vendor onsite with the TSIF Test Branch, TIOB CM Specialist, and/or TSIF Test CM Lead to support the applicable configuration audit or test characterization check.
- *TSE Installation* – In coordination with the COR, PMO, and TIOB, vendor onsite to support installation and integration.
- *Verification Support* – Vendor onsite to provide requirement demonstrations for compliance verification.

A vendor agent and TSIF Test Director or designee are required to formally sign the TRN. The TSIF Test Branch submits the TRN to the COR and vendor for record keeping.

4.2.4 Test Article Selection and Development

Test articles used during QT vary by technology, from baggage articles with explosive simulant threats to fraudulent and authentic identification documents. The TIOB and/or TSIF Test Branch (test-dependent) develops stream of commerce representative test articles in order to simulate operational conditions for scenarios that assess areas such as false alarm rate and throughput.

The TSIF Test Branch and TIOB meet with appropriate organizations to assist in the selection of threat articles to test any scenarios in the QTP that require these articles. These organizations can include the EQAB, user representative, intelligence agencies, other DHS agencies, and other expert groups within TSA (e.g., OIA analysts) that can provide an expert opinion on whether the articles are representative of the threat posed by the adversary.

When high fidelity test articles are required in support of QT, the TIOB team develops the test articles in accordance with the *TSA VV&A Policy* to ensure that the test articles are developed and configured to meet requirements (verification) and are an accurate representation of the real world (validation). The verification and validation is performed by the developing agency, typically the TIOB and/or TSIF Test Branch. Once verification and validation is complete, the System Evaluator determines whether the articles are acceptable for its intended use (accreditation decision).

4.2.5 Test Tool Development and Verification, Validation, and Accreditation

During the development of the QTP, the TSIF Test Branch determines the test tools and instrumentation needed to support T&E. These test tools can range from automated data collection and analysis software to robotics used to perform constant tasks to stress system use. The team first develops a test tool-specific requirements document to be approved by the product owner and sponsor. The team creates a validation plan to ensure the test tool will meet the criteria listed in the test tool document. The validation plan includes the number of samples needed and the type of data required to be collected (verification). Based on these attributes, the TSIF Test Branch develops the product, or acquires the necessary materials, and runs the tool

through a sample of data to determine result consistency and compliance with pre-approved criteria (validation). At the System Evaluator's discretion, EQAB would accredit the test tool if it is used to produce data in support of a critical requirement in the SER.

4.2.6 Coordination with Supporting Organizations

Some QT scenarios in the QTP may require a Transportation Security Officer (TSO) to operate the TSE to best represent the TSE's performance when used in its intended environment. It is the responsibility of the TSIF Test Director to begin coordination with organizations that will be providing operators for testing during development of the QTP. Although test dates may not be known, the initial coordination informs the organization of the need and lowers the risk that operators will not be available when the test dates are set. The TSIF Test Director informs the System Evaluator during the SET meeting of any issues in obtaining TSOs along with a mitigation plan.

The System Evaluator, working with the test organizations, may determine the need for other organizations to support the QT event. These organizations may include IT IAD for IA/Cyber Resilience requirements and OSHE for safety requirements. These assessments generally occur at the TSIF once the TSE is installed, however, may occur remotely dependent on the nature of the requirements. It is the responsibility of the TSIF Test Director to coordinate with the necessary organizations to ensure access is given to conduct the test.

4.2.7 Qualification Test Readiness Review

The QTRR is a forum to assess the system's readiness to enter QT. The TSIF Test Branch Manager, or designee, chairs the QTRR and is responsible for presenting the QTRR checklist and completion status for each readiness criteria. QTRR participation, at a minimum, includes the PMO representative (and/or T&E Manager), user representative, System Evaluator, TSIF Test Director, and other principal members of the SET and IPT.

During the QTRR, the TSIF Test Branch Manager or designee reviews all pre-start activities to determine if the entrance criteria are met, confirming test resource availability, safety, training, and that the SUT and instrumentation are in place to support the successful conduct of the test. The TSIF Test Branch Manager or designee also ensures that T&E planning documentation, system baseline configuration, and data management processes have been adequately addressed by the TSIF Test Director. The outcome of the QTRR is a determination by the TSIF Test Branch Manager or designee that QT is ready to begin. The TSIF Test Branch Manager or designee may determine additional items are necessary and delay the start of testing.

4.3 Qualification Test Execution and Reporting

During test execution, the TSIF Test Branch collects quantitative and qualitative data in accordance with the test scenarios in the QTP. The data collection may consist of recording results from vendor *demonstration*, *analyzing* vendor-provided data and documentation, physical *inspection* of hardware and/or software, and directly executing *test* scenarios as performed by the TSIF Test Branch or TSOs. The TSIF Test Branch may refer to the vendor-provided QDP documentation when resolving requirements. During test reporting, the TSIF Test Branch

presents preliminary findings to the user representative(s) for results adjudication. The TSIF Test Branch typically prepares a final QLR and QT Report with adjudicated results, formally describing findings, recommendations, additional issues, severity ratings, and acceptable workarounds as agreed to by the user representative. Section 4.3.6 provides additional information on results adjudication and severity level definitions.

Using QT results and findings as an input, the PMO convenes a System Readiness Review (SRR) to determine whether the system should proceed to further testing or whether remediation activities should be initiated. Alternatively, the PMO may leverage a satisfactory QLR in lieu of convening an SRR meeting. The SRR is an entry criterion to OT. Figure 4-2 describes key activities taking place during QT execution and reporting, as conducted by the TSIF Test Branch.

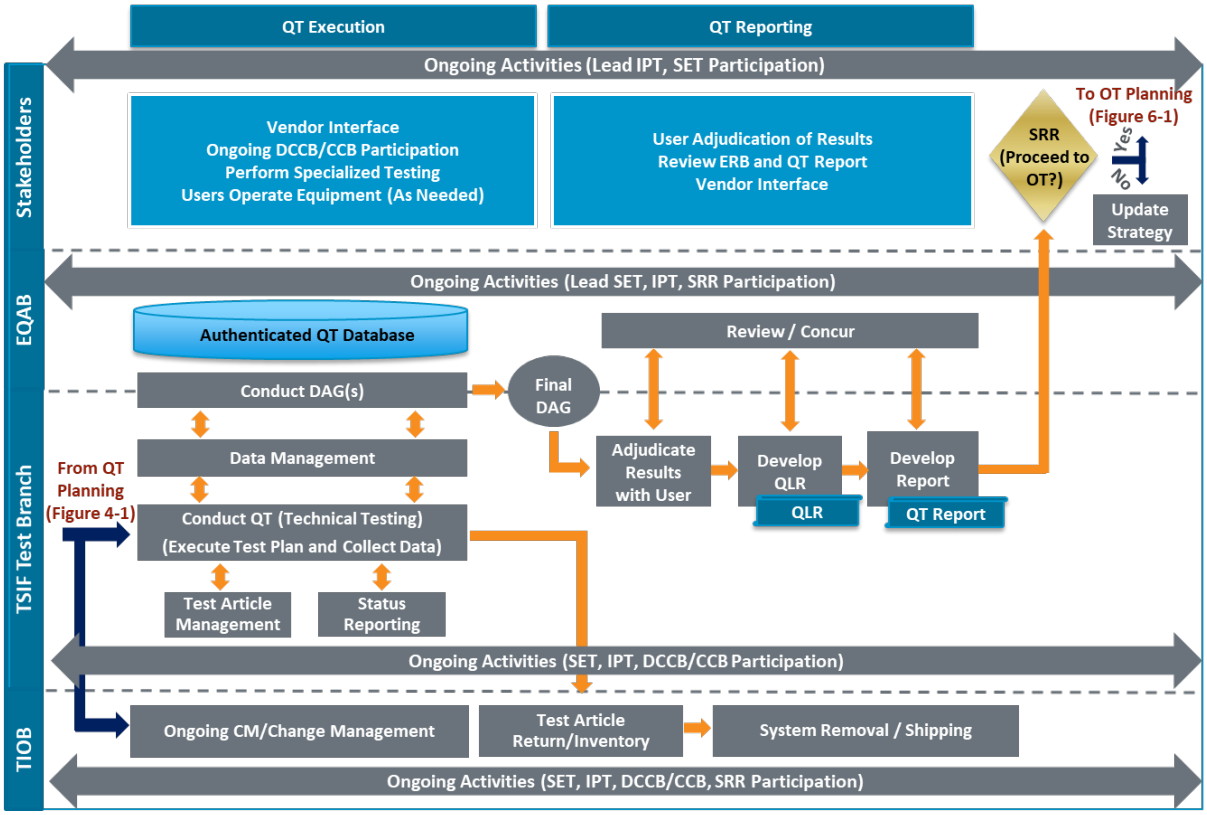

Figure 4-2: Qualification Test Execution Process Flow

4.3.1 Test Article Management

The TIOB Test Article Group develops, acquires, and controls QT articles. During test execution, the TSIF Test Branch manages the inventory, use, and control of test articles and returns the articles to the TIOB Test Article Group on completion of test activities. The TSIF Test Branch should ensure sensitive test articles are not visible to non-test team visitors (e.g., vendors, tours) to the test site. The TIOB team maintains a current inventory of all articles available or in use so that they can quickly respond to any inquiries.

4.3.2 Configuration Control during Test Execution

The TIOB CM Specialist manages the configuration baseline(s) against any vendor-submitted changes received after baseline approval. The TIOB CM Specialist receives the vendor's change request in the form of a Request for Deviation (RFD) or DEP and distributes it to the stakeholders for review and their recommended disposition in accordance with the APM CCB process. The TIOB CM Specialist presents Subject Matter Expert's findings to the APM CCB for final board disposition. Stakeholders should refer to the accepted, vendor proposed CI to FRD mapping in the QDP when assessing the change and making the approval decision. All stakeholder comments are consolidated and a disposition from the CCB is made. The CCB is the approval authority for technical changes. The PMO will route changes to contract scope, cost, or schedule to the CO for final adjudication.

During the change control process, the System Evaluator receives input from the TSIF Test Director and Operational Test Director to determine whether any changes to system configuration require regression testing and to socialize impacts to schedule and budget to the PMO prior to APM CCB approval. The TIOB CM Specialist maintains a history of approved changes for each SUT. The *TSA APM CMP* has additional information for configuration control during test execution.

4.3.3 Maintenance Activities

Maintenance activities during a QT event occur per the guidelines established in the TRN. Once the test event begins, the primary need for vendor onsite participation is preventive or corrective maintenance. Preventive maintenance activities are planned maintenance on the TSE often occurring at regular intervals (e.g., weekly, monthly). While some of these actions can be performed by the TSIF Test Branch, others require vendor involvement. TSA defines corrective maintenance as unanticipated requests for remediation due to an equipment malfunction.

Vendors must coordinate preventive maintenance activities with the TIOB Test Area Lead and the TSIF Test Director in the form of a schedule of activities prior to the test, as documented in the TRN. During the test event and prior to each maintenance activity, the vendor should coordinate with the TIOB Test Area Lead and TSIF Test Director to coordinate the visit. The TSIF Test Branch and/or TIOB Test Area Lead must monitor all preventive and corrective maintenance during the test event to ensure only the required maintenance is occurring. The TIOB Test Area Lead will document all actions performed by the service technician in a TSE Maintenance Checklist while the TSIF Test Branch records the maintenance actions in a Test Incident Report (TIR) as necessary.

4.3.4 Qualification Testing Data Authentication Group

Data authentication is performed at the discretion of the TSIF Test Director during test execution. Data authentication consists of a comprehensive review to ensure the data are complete, accurate, consistent, and representative of system performance within the tested environment. The DAG periodically reviews data for each test scenario, and provides feedback and recommendations to the test team. The DAG meeting is conducted to provide quality

assurance of the collected data. The output of the DAG is authenticated results for use in data analysis and reporting.

The TSIF Test Director chairs the DAG and is generally the final decision maker when authenticating data. The TSIF Test Branch Manager may choose to attend a DAG meeting and, if so, will serve as the final authority for data authentication. DAG participants may not release data without the prior permission of the DAG chair. Table 4-2 lists the minimum required DAG personnel and their role.

Table 4-2: Qualification Test Data Authentication Group Personnel

Test Role	DAG Role
TSIF Test Director	DAG Chair
TSIF Test Branch Manager (optional)	Attends DAG/Alternate Chair
QT Support Team Lead Engineer	Leads presentation of test results
QT Support Team Personnel	Provides details on collected data
System Evaluator	Participant
PMO Representative	Participant
User Representative	Participant

At the DAG meeting, the TSIF Test Director presents all data collected since the previous DAG data cutoff. The members of the DAG review the records and determine whether the record can be used for analysis and reporting. Any inquiries about the data records are directed to the TSIF Test Director, who either answers the question or conducts an investigation and provides the answer during the next DAG meeting. The TSIF Test Director considers all input and recommendations from the DAG to decide the most logical path for resolving each issue, and if additional data collection is required to resolve.

4.3.5 Early Termination of Qualification Test Activities

In addition to violating the restrictions listed in the TRN (Section 4.2.4), other events could also cause the suspension of a QT event. These include the following:

- Safety evaluation or occurrence during testing results in a critical deficiency
- TSE limitations or failures impede data collection
- The TSIF Test Branch discovers an unapproved system configuration discrepancy

4.3.6 Result Adjudication

At the completion of QT (or regression QT), the TSIF Test Director meets with the user representative, PMO, and others as needed to adjudicate test results. The TSIF Test Director

provides the stakeholders with an initial QLR that outlines pass/fail results and associated issues encountered during QT, to include substantiating data, and recommended severity ratings. The user representative may confirm or revise the severity rating for each failure presented using the definitions provided in Table 4-3. Representatives from IT adjudicate IT-related issues while IT IAD personnel adjudicate IA-related test results.

Table 4-3: Qualification Test Severity Ratings (Adjudicated)

Severity	Applies if an issue could:	Assessment
1	• Prevent the accomplishment of an operational or mission essential capability • Jeopardize safety, security, or other requirement designated as "critical" • Potentially cause a loss of life.	Do not proceed
2	• Adversely affect the accomplishment of an operational or mission essential capability and no work-around solution is known. • Adversely affect technical, cost, or schedule risks to the project or to life cycle support of the system, and no work-around solution is known.	
3	• Adversely affect the accomplishment of an operational or mission essential capability but a work-around accepted by the user rep. solution is known • Adversely affect technical, cost, or schedule risks to the project or to life cycle support of the system, and a work-around accepted by the user rep. solution is known.	Do not proceed if clustered
4	• Result in user/operator inconvenience or annoyance but does not affect a required operational or mission essential capability • Result in inconvenience or annoyance for development or support personnel, but does not prevent the accomplishment of those responsibilities.	Continue, pursue fix
5 (P)	• Any other effect or desired system changes to current baselines requirements • Anything not in the requirements but customer desires	Future Enhancement
P	• Indicates sufficient testing to attest to the ability of the system to meet the requirement with a high level of confidence	
ID	Indicates insufficient data.	
NA	Indicates item was not applicable.	
NT	Indicates item was not tested	

The TSIF Test Team will record results, by requirement, per the following severity table when results have not been adjudicated with the user and/or other subject matter expert. This approach

is typically utilized only for TSIF Data Collection Activities. A formal QT event may only utilize the severity ratings in Table 4-4 with approval by the TSIF Test Branch Manager.

Table 4-4: Data Collection Activity Severity Ratings

Severity	Applies if an issue could:	Assessment
Fail	• Prevent the accomplishment of an operational or mission capability. • Jeopardize safety, security, or other requirement designated as "critical". • Potentially cause a loss of life.	**Do not proceed**
Pass	Indicates sufficient testing to attest to the ability of the system to meet the requirement with a high level of confidence.	**Pass**
ID	Indicates Insufficient Data (ID) to form conclusions.	
NA	Indicates item was not applicable.	
NT	Indicates requirement was Not Tested (NT).	

The TSIF Test Director supports entry to OT if the following criteria are met:

- No open Severity 1 or 2 issues.
- No multiple Severity 3 issues (clusters) located in the same functional area of the system.
- Frozen system configuration baseline.
- All required test scenarios completed.

4.3.7 Data Analysis and Reporting

The TSIF Test Branch performs data analysis of QT data with respect to FRD requirement thresholds and any other success criteria from the QTP. In addition to computing results for each of the test scenarios, the TSIF Test Branch examines the data for trends that provides additional meaning and insight to the test results.

Critical to the analysis of QT results is the documentation of observations, even if not directly associated to a requirement, that may have a profound impact in an operational environment. Furthermore, the QT analysts leverage photographs and screen captures taken during QT to provide additional information for users during adjudication sessions. Analysts generally leverage hypothesis testing when analyzing data. The confidence level used is typically 90 percent when documenting results. The System Evaluator is a critical reviewer of the final test results to ensure the test data accurately reflect the capabilities and deficiencies of the system.

The TSIF Test Director prepares a QT QLR on a case-by-case basis as documented in the QTP, usually when information is required prior to a formal QT Report (QTR). The QT QLR provides preliminary findings based on QT data.

The TSIF Test Director develops the QLR upon completion of the QT with a delivery date of approximately 15 calendar days after the final DAG meeting. The TSIF Test Director gains concurrence from the System Evaluator, and final approval from the TSIF Test Branch Manager prior to final delivery of the QLR.

The TSIF Test Director also prepares a QTR (or DCAR for TSIF Data Collection Activities) describing system performance against each requirement. The TSIF Test Director provides these to the System Evaluator for use as an input during SER development and to the Operational Test Director to assist with OT planning. The System Evaluator concurs with the QTR.

The TSIF Test Branch Manager is the final approver of the QTR or DCAR and in general delivers the QTR or DCAR 45 calendar days after the final DAG unless otherwise documented in the test plan. The QT QLR, QTR, and/or DCAR will indicate the severity rating for any failures encountered during testing. These reports are provided to the PM as input to the SRR. Generally, satisfactory QT results are required for a successful SRR outcome.

5.0 ACCEPTANCE TESTING

The TSA Acceptance Test Team performs AT on TSE after production installation and, when applicable, integration. This ensures consistency of the manufacturing process and a verification of system configuration, performance, Baggage Handling System (BHS)/Checked Baggage Inspection System (CBIS) performance and operability. The Acceptance Test Team generally conducts testing prior to the start of OT and performs Factory Acceptance Test (FAT) followed by additional AT for each subsequent unit during system deployment. This verification is typically performed in support of the test events listed in Table 5-1.

Table 5-1: Types of Acceptance Tests

Test Type / Purpose	Additional Detail / Decision Type	Sponsor	Planning Level	Test Length
First Article Test and Evaluation (FAT&E). FAT&E involves testing and evaluating the first article for conformance with specified contract requirements before or in the initial stage of production under a contract. FAT&E results in a product baseline. First articles include preproduction models, initial production samples, test samples, first lots, pilot models, and pilot lots.	TSA production baseline establishment. Proceed to Factory Acceptance Test (FAT)/ production decision	PMO DLD	High	1 week - 2 months
FAT Production testing that is planned, conducted, and monitored by the materiel developer. FAT includes preproduction and initial production testing conducted to ensure that the contractor can furnish a product that meets the established technical criteria. Validates applicable requirements on units prior to shipment from the manufacturing facility	TSA deployment / purchase / ownership decision	PMO DLD	Medium/ Low	1 day
Site Acceptance Test (SAT). The SAT entails inspection and dynamic testing of systems or major system components to support the qualification of equipment. The SAT verifies that the installed functionality of the equipment meets or exceeds the operational requirements. The SAT is executed on completion of all commissioning tasks and validates applicable requirements on units installed for the first time at their operational location	TSE installation operational decision/user acceptance	PMO DLD	Medium/ Low	1 day
Operational Readiness Test (ORT) Similar scope than the SAT, validates applicable requirements on units relocated	TSE relocation operational decision	PMO DLD	Medium/ Low	1 day

Test Type / Purpose	Additional Detail / Decision Type	Sponsor	Planning Level	Test Length
and/or significantly modified. Can be performed any time after a SAT.	/ TSE modifications operational decision			
Network Acceptance Test (NAT) Validates the functionality of vendor-installed network environments in support of EDS installation and communication with components.	TSE installation operational decision	PMO DLD	Medium/ Low	1 Day
Integrated Site Acceptance Test (ISAT) Validates inline CBIS performance, consisting of an Integrated EDS and BHS. The ISAT test procedures are written by ATSA and verifies that the equipment functionality meets or exceeds the site specific operational requirements.	TSE installation and integration operational decision	PMO DLD	High	Varies by site

The following sections further describe the test events in the above table.

5.1 First Article Test and Evaluation

First Article Test and Evaluation (FAT&E) ensures that the very first manufactured unit by the vendor after acquisition approval conforms to all contract requirements for Government acceptance. FAT&E is also used to establish the configuration baseline upon which all units manufactured in the future are compared against. The PMO oversees FAT&E in conjunction with the vendor with the Acceptance Test Team serving as a witness. The TSA APM CM group performs a physical configuration audit to verify configuration control of the unit under test. TSA and the vendor must successfully complete FAT&E prior to proceeding with any subsequent unit production acceptance unless otherwise directed by the Government.

5.1.1 Factory Acceptance Test, Site Acceptance Test, and Operational Readiness Test

The Acceptance Test Team performs Factory Acceptance Tests (FATs), Site Acceptance Tests (SATs), and Operational Readiness Tests (ORTs) to ensure proper functionality of a system. The Acceptance Test Team performs this testing with a Government representative witnesses.

The FAT/SAT/ORT process begins with the creation and delivery of test plans for FATs and SAT/ORTs from the vendor. TSA reviews the supplied procedures and provides any comments to the vendor, to include any additional required tests. The vendor owns the FAT procedure and will ensure a TSA approved copy is utilized for testing. TSA will maintain the SAT/ORT procedure and update it as system changes dictate. As part of the review process, the Acceptance Test Team and the vendor perform a dry run on an available unit at the TSIF or vendor facility using the vendor-submitted test plan, procedure, and report. During this dry run, the test procedures will be red lined for any changes or omissions to the procedure.

Once the units are ready for FAT, the vendor executes the test with a TSA witness in accordance with the approved test plan and procedures. For SATs/ORTs the Acceptance Test Team will perform the test and the vendor will support/witness the test. Upon completion of testing the Acceptance Test Team obtains test reports from the vendor and develops a QLR for posting/uploading in the approved TSA storage location.

The Acceptance Test Team assigns the system one of the following statuses in the test summary QLR:

- *Pass* – The unit under test met all criteria specified within the test procedures and, therefore, is ready for operations.
- *Fail* – The unit under test did not meet the criteria specified within the procedures, or is not ready to be installed for operations.
- *Defects Found* – A unique rating that is assigned to units that may have an anomaly that is outside of the test requirements, or other issues of concern that can potentially impact operational usage.

TSA gives the vendor 24 hours to diagnose and correct failures upon notification of a failed QLR. A failed QLR will be stored/uploaded into the TSA approved storage location. The Acceptance Test Team will remain on site for 24-hours if the vendor indicates that the system fix will be completed during that time. If the fix cannot be completed in 24-hours the team will depart the site and a new TRN will need to be submitted at least 5 business days prior to the requested test date. Once a system passes a FAT, the Government takes ownership of the TSE and affixes a TSA barcode on the TSE. At this point the vendor payment is made according to contract terms. The system is then eligible to be shipped to its intended operational location for a SAT. The Acceptance Test Team will perform an ORT for relocated units or systems that have a significant hardware or software change (e.g. upgrades, additional functionality).

5.1.2 TSE Network Acceptance Test

The Network Acceptance Test (NAT) verifies the functionality of the system network as installed in the field by the vendor. The NAT confirms the hardware and software configuration, network setting options, and reporting. The NAT also confirms functionality such as connectivity validation, ability to control operation of network elements, image routing, data recording, and network fail overs. The Acceptance Test Team assesses each system against criteria within the approved NAT procedures and generates a test report documenting unit pass or fail. The same failure process described under FAT is applicable to a NAT, SAT, and ORT. Once a system passes a NAT and SAT/ORT, the system can be operated by the end-user to screen live passengers.

5.1.3 Integrated Site Acceptance Test

The Integrated Site Acceptance Test (ISAT) is currently applicable only to in-line CBISs which consist of the airport BHS and EDS. The testing portion of an ISAT consists of two parts. The first part is a series of controlled tests (line tests) that check the specific functionality. The second part is a set of controlled stream of commerce system tests. These tests are completed

with a team consisting of personnel from the OTA, Airport Project Sponsor, EDS vendor, BHS Programmer, and local (airport) TSA personnel.

The ISAT process is a three-part process, with test planning activities occurring early in the planning and construction phases to assist the project team in developing the schedule, recommending test periods, and scoping test phases. The following process flow outlines the ISAT process.

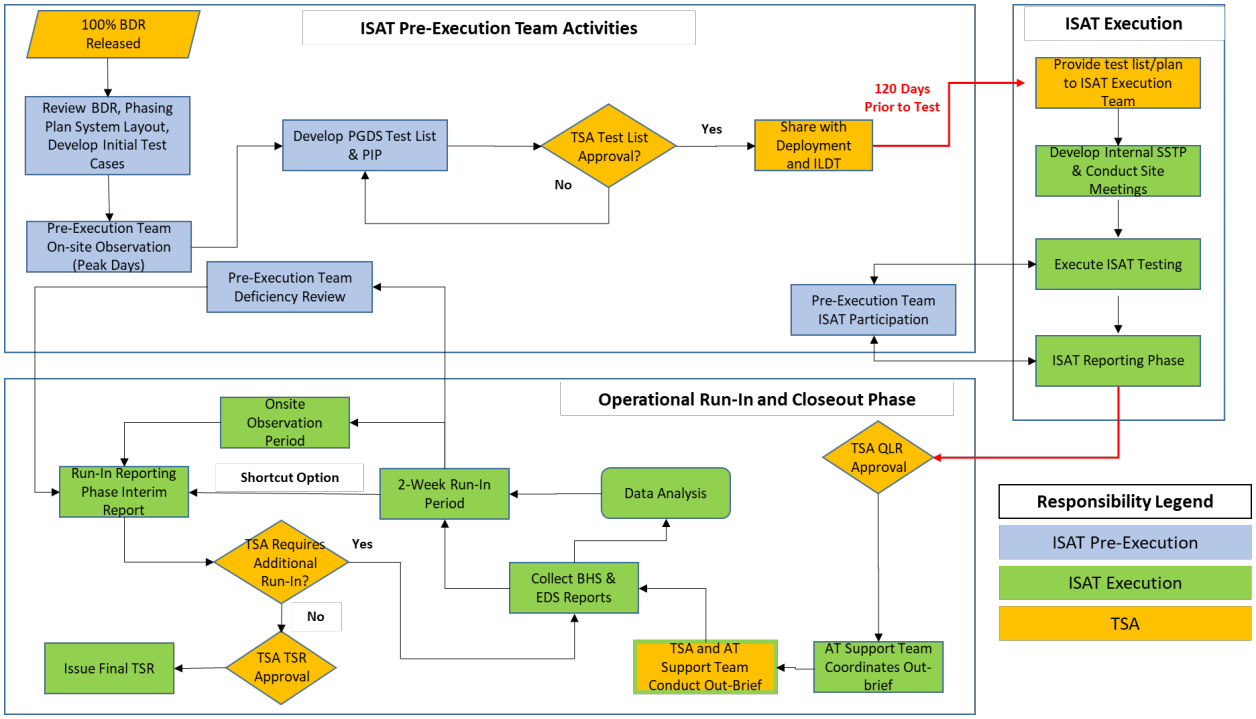

Figure 5-1: ISAT Process Flow

Upon completion of the ISAT an operational run-in period is performed. During this period the system is monitored during live operations using stream of commerce luggage to assess system performance. The Acceptance Test Team will require access to both EDS and BHS data to analyze the system. The duration of the operational run-in depends on the performance of the system and typically runs between 14 and 30 days.

Upon completion of testing two reports will be drafted. The first report is a QLR which documents the CBIS performance based on the ISAT testing. The second report is the Test Summary Report (TSR) which is an encompassing document that includes the results of the QLR and also observations collected during the run-in period. The issuance of the reports and recommendations will occur as follows:

- Executive Review Board (TSA Go Live Authorization): Available from TSA two (2) days after testing. The Regional Deployment Coordinator will provide this to the airport.
- Draft QLR: Due to TSA four (4) days after the completion of ISAT testing.
- Airport ISAT Outbrief: Due five (5) days after completion of ISAT testing.

- Final QLR: Due seven (7) days after completion of ISAT testing.
- TSR: Due five (5) days after the completion of the run-in period.

The QLR and TSR results will be depicted in a standard risk cube. The risk cube is divided into three colors (red, yellow, and green). If there are any test categories that fall into the red zone of the cube the system will not proceed to live operations until the deficiencies are corrected. If the overall system classification falls into the green shaded boxes, then the system will be recommended to proceed to live operations. A classification in the yellow shaded boxes will require some mitigation prior to system operations. This mitigation could be in the form of a procedure modification, software/hardware modification, or staffing modification. Figure 5-2 provides an example of a risk cube.

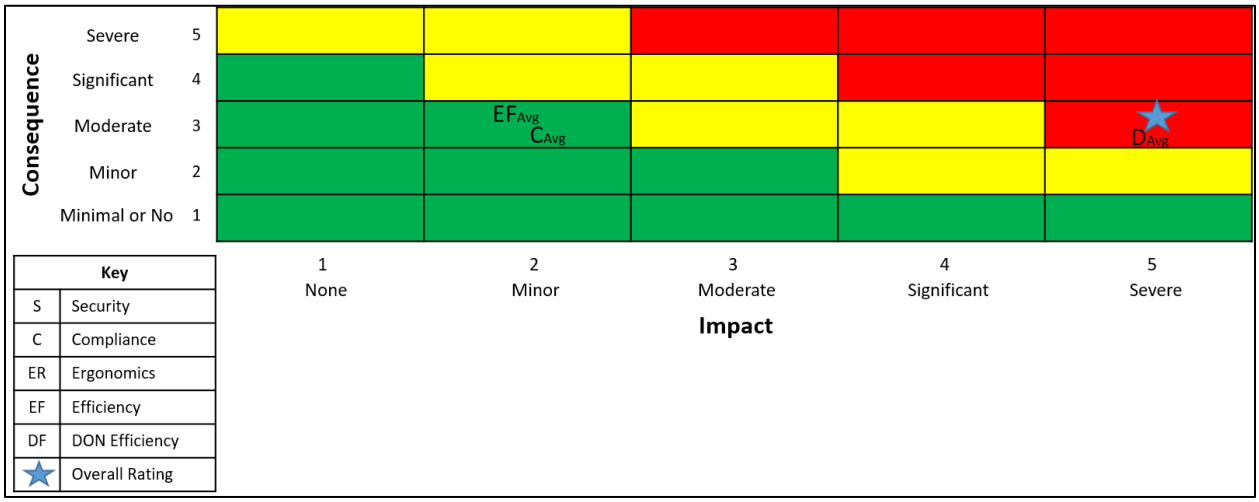

Figure 5-2: Example of a Risk Cube

If a system receives an overall risk rating in any of the red boxes, the system will not be retested for a period of at least thirty (30) days to allow for system or procedure modification.

5.1.4 Site Acceptance Test in Support of Qualification Testing

The TSIF Test Branch conducts a SAT upon request prior to QT at the TSIF. The SAT ensures the system as delivered by the vendor functions as expected and documented in the vendor's QDP. The SAT normally consists of a review of vendor system documentation and vendor demonstration of core system functionality.

5.1.5 Site Acceptance Test in support of Operational Testing

The Acceptance Test Team will conduct either a SAT or ORT based on the scope of the Operational Test. For example, if a system has an algorithm change, the Acceptance Test Team performs an ORT. If a host computer is replaced, then the Acceptance Test Team should perform a SAT. The SAT and ORT procedures are essentially identical in scope.

5.1.6 Integrated Site Acceptance Test in Support of Operational Testing

The Acceptance Test Team conducts an OT ISAT upon completion of the installation and integration of the integrated SUT. ISATs generally apply to integrated EDS. As opposed to a regular ISAT, an ISAT in support of OT does not require a formal TRR to kick off testing activities. Instead, the Acceptance Test Team holds a combined TRR/ISAT. This allows for a seamless transition from BHS pre-testing to record ISAT testing. In some cases, results from the TRR are used for ISAT.

Upon ISAT completion, the Acceptance Test Team gathers data and reports from the EDS and BHS for data reconciliation. The Acceptance Test Team generates a QLR within four business days of ISAT completion and provides it to the System Evaluator and the PMO.

5.2 Acceptance Test Articles

Test articles are required to support a wide range of AT activities across the various security technology platforms. These test articles can include image quality phantoms and stream-of-commerce representative test luggage. The following sections provide descriptions of two types of test articles and their intended purpose.

5.2.1 Image Quality Test Kits

Technology-specific image quality phantoms ensure the proper setup and operation of image-producing security technologies. These specialty phantoms are designed to characterize the imaging subsystem of candidate platforms and provide a quantitative assessment of image quality. These test articles focus on a measurement of specific image quality parameters predetermined to be critical to the operation of the system. For further reference on the new phantom kits, please reference the American Society for Testing and Materials (ASTM) F792-08, Standard Practice for Evaluating the Imaging Performance of Security X-Ray Systems or American National Standards Institute (ANSI) 42.45-2011(or most recent), American National Standard for Evaluating the Image Quality of X-Ray Computed Tomography Security-Screening Systems.

5.2.2 Stream-of-Commerce Representative Test Luggage

Stream-of-commerce representative test luggage is used to evaluate performance on platforms during SAT or ORT events. This luggage represents the types of bags used by the traveling public.

6.0 OPERATIONAL TESTING

OT independently validates the extent to which candidate systems are Operationally Effective, Suitable, and Cyber Resilient in the operational environment. This phase of testing focuses on evaluating operational requirements from the ORD as defined by the user organization and serves as a realistic demonstration of a systems expected field capability.

6.1 Operational Test Events Overview

OTB performs Field Tests and OTs in support of an acquisition decision. Due to threat intelligence, risk-based initiatives, procedural development, and emerging technology exploration, needs for operational data collection arise outside of a planned acquisition activity. These events may lead to the need to perform a FDCA. Table 6-1 lists the various types of operational evaluation/assessment activities and their purpose.

Table 6-1: Operational Testing Activities

Test Type	Additional Detail/Decision Type	Plan / Report	Sponsor	Typical Length
Initial Operational Test and Evaluation (IOT&E)	OTB conducts an IOT&E to inform an ADE 3 and evaluate a system's Operational Effectiveness, Suitability, and Cyber Resilience using typical, trained personnel in a realistic operational environment. The IOT&E is the initial OT for a system.	SEP OTP QLR ERB SER	PMO	28-45 days
FOT&E	An FOT&E may be necessary to verify the resolution of issues discovered during an IOT&E or previous FOT&E and/or to reevaluate the system when a vendor proposes changes to ensure that it continues to meet operational needs.	OTP Update QLR ERB SER or SER Redress	PMO	28-45 days
OA	An evaluation of specific Operational Effectiveness, Suitability, and/or Cyber Resilience measures. OAs focus on developmental efforts, programmatic voids, risk areas, adequacy of requirements, and the ability of the program to support adequate OT. OAs may be conducted at any time in a program lifecycle using technology demonstrators, mock-ups, engineering development models or simulators, but are typically done early in the concept or development phases. OAs will not substitute for an OT to support full rate production and deployment decisions.	OAP QLR ERB OAR	PMO	21-45 days

Test Type	Additional Detail/Decision Type	Plan / Report	Sponsor	Typical Length
Field Test	OTB conducts a Field Test to provide the program office an indication of a system's readiness, and risk, to proceed to an OT. The scope of a Field Test generally consists of the measures, or subset of measures, listed in the OTP. A Field Test is an optional step in a system's formal qualification process, and is typically performed at the request of the PMO.	DCAP QLR DCAR	PMO	5-45 days
FDCA	OTB conducts a FDCA to provide data on system and/or procedural capabilities and limitations in an operational environment.	DCAP QLR DCAR	PMO or requesting organization	5-45 days

6.2 Operational Test Planning

The Operational Test Director, supported by the OT team, performs a number of planning activities to prepare for tests in the operational environment. Central to these planning activities is the development of an OTP, OAP, or DCAP. In general, the OT team may need to perform many of the same activities described in the following sections when planning for OTs, OAs, Field Tests, and FDCAs with the exception that in some cases, the planning activity may not apply (or be less stringent) for Field Tests and FDCAs. For example, when assessing a procedural change, a vendor TRN (Section 6.2.6) is not required.

Figure 6-1 describes the various activities taking place during test planning by the OTB, EQAB, and other stakeholders in support of an acquisition program OT.

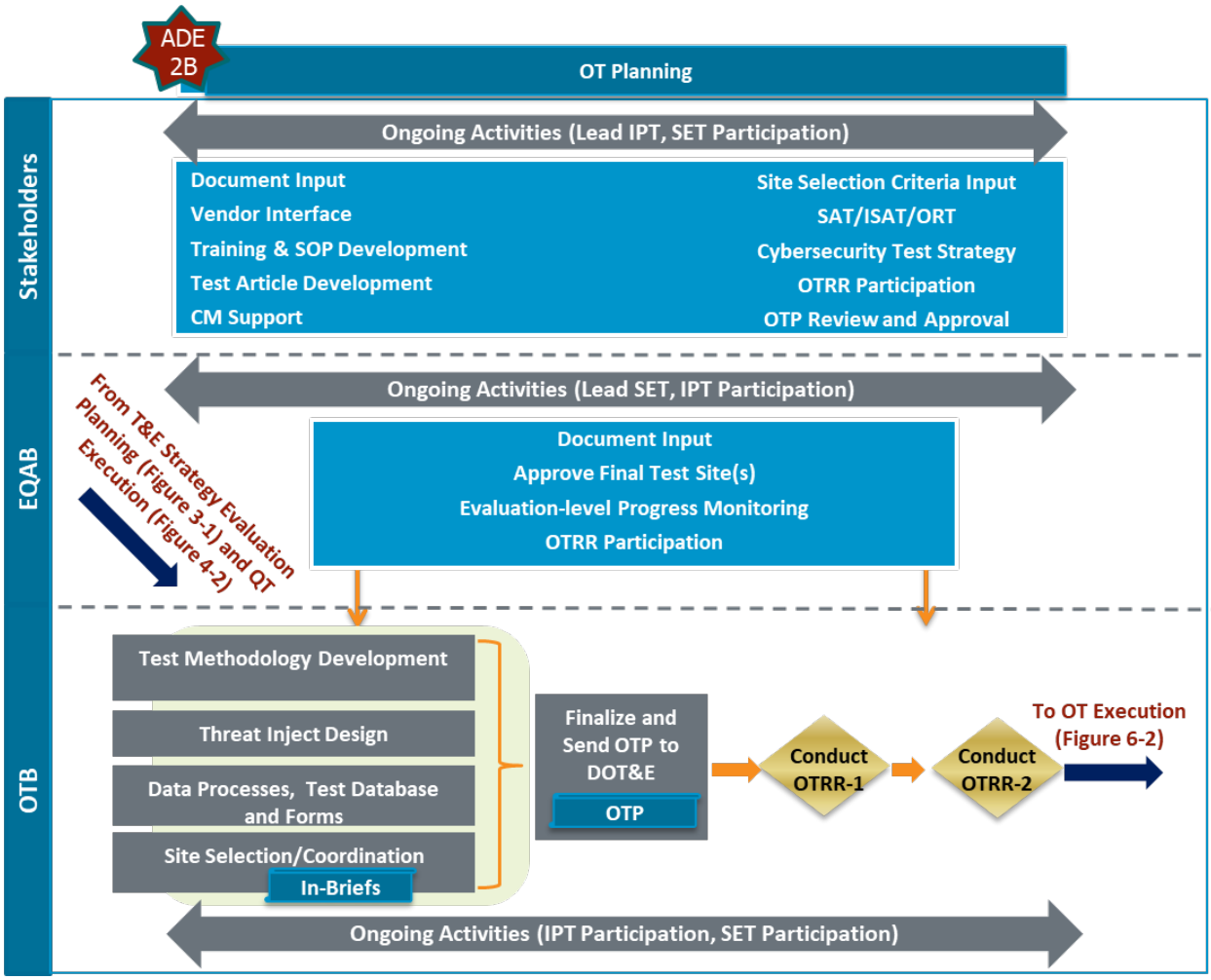

Figure 6-1: Operational Test Planning Process Flow

6.2.1 Planning Documentation

The Operational Test Director creates an OTP for test events in support of an acquisition program based on the evaluation strategy and requirements allocated to OT as documented in the IEF. The OTP describes the details of the test to include logistical considerations, resources, test/data collection/data analysis methodologies, data collection schedules, and Cyber Resilience assessment methodologies. The Operational Test Director gains concurrence from the System Evaluator that the OTP will adequately support the evaluation prior to initial approval from the OTB Manager and EQAB Manager and initial approval by the T&E Director, OTA. Upon T&E Director, OTA approval, the Operational Test Director routes the OTP for final approval by DOT&E.

The DOT&E approves the OTP within 60 days of submission, therefore, the Operational Test Director should ensure it is submitted with adequate time to not impact the start of OT.

The Operational Test Director creates a DCAP in support of a Field Test and for FDCAs. The DCAP is similar to the OTP in that it contains all the necessary planning information for the test event. The PMO (or other sponsor of the test event) may determine a change in test scope/requirements is necessary after the start of a Field Test or FDCA. In these cases, the Operational Test Director will follow the procedure described in Appendix D. The Operational Test Director gains approval from the System Evaluator that the DCAP will adequately support the sponsoring organizations requested assessment. The Operational Test Director gains final DCAP approval from the OTB Manager.

6.2.2 Site Selection

The Site Selection Process for each SUT may vary based on the system and testing requirements. Each site selection will be tailored to the specific conditions required for the SUT, and the type of testing required (e.g., Field Test, OT). The Site Selection Authority (SSA) for Field Tests is the Program Manager. The SSA for OT events is the System Evaluator. Typically, a representative from the sponsoring agency (e.g., RCA, Program Office] will be the SSA for FDCAs. The following illustrates the typical site selection process.

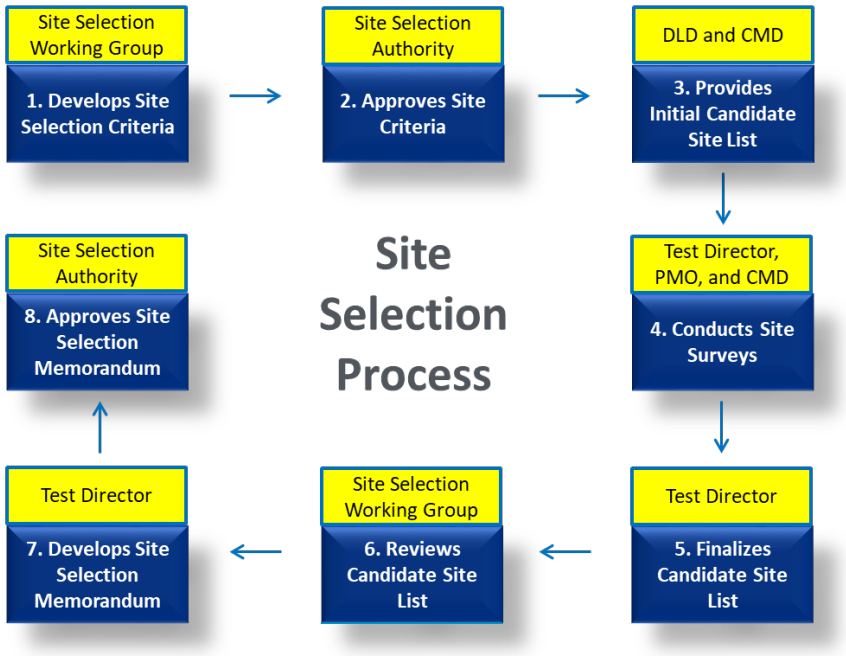

Figure 6-2: Site Selection Process

6.2.2.1 Site Selection Working Group Develops Site Selection Criteria

Initial site selection criteria are identified in the TEMP. Additional site selection criteria are developed by the SET, or designated site selection working group, based on the operational requirements for the SUT. Site selection criteria may include such attributes as:

- Passenger mix, bag population, or other operational factors representative of a typical deployment site if the system is procured.
- Ability to support test objectives to include required system configurations (if needed, e.g., baseline systems), RMA thresholds, volume, counts, and sample sizes.
- Geographically diverse sites and checkpoints to make sure the new system can operate the same way with different operator and passenger populations.
- Representative utilization rates and ability to support required operating hours.
- Local ability to support testing (e.g., sufficient Transportation Security Officer manpower to support training and operations throughout the test duration).
- Compatible existing infrastructure (same vendor, dedicated power circuit, layout, etc.) and infrastructure change needs and impact.
- Ability to integrate the system into the operational environment.
- Airport badging process support test schedules and planned costs.
- Other airport activities do not negatively impact operations or user availability (e.g., training, blackout dates, construction, or other testing).

6.2.2.2 SSA Approves Site Criteria

The SSA as described in Section 6.2.2 approves the site selection criteria.

6.2.2.3 DLD and CMD Provides an Initial Candidate Site List

DLD and the CMD user representative provide representative sites that potentially meet the site selection criteria. These sites serve as initial locations for consideration. Once the candidate site list is generated, the Operational Test Director will gather additional data as needed from sources such as, but not limited to the Staffing Allocation Matrix (SAM) or through request to the user organization. The Operational Test Director will use the compiled data to analyze/score the sites against the identified criteria in a Site Selection Worksheet resulting in a best qualified list. A Site Selection Worksheet template is available in the OT template library.

6.2.2.4 Operational Test Director, PMO, and CMD Representatives Conduct Site Surveys

CMD will make initial contact with the candidate sites. After initial contact is made, the Operational Test Director will coordinate with the best qualified sites, as needed, to schedule field surveys with appropriate stakeholders such as PMO and CMD representatives. The Operational Test Director will manage communications with airport leadership to keep them informed and set the expectations and ground rules for the test events.

The field survey will verify site capabilities and identify site limitations. In parallel, the site selection working group will begin test schedule coordination with the site to determine site availability, site access requirements, and site logistics capabilities to support the planned testing.

6.2.2.5 Operational Test Director Finalizes Candidate Site List

The Operational Test Director updates the Site Selection Worksheet based on additional information gained during the field survey, prioritizes the sites, and provides it to the site selection working group for feedback.

6.2.2.6 Site Selection Working Group Reviews Candidate Site List

The site selection working group reviews the candidate site list and provides feedback to the Operational Test Director to inform the site ratings and selections.

6.2.2.7 Operational Test Director Develops the Site Selection Memorandum

After the site selection working group review of the final site ratings and selections, the Operational Test Director develops a Site Selection Memorandum documenting the findings of the site selection analysis, includes a description of each site's capabilities and limitations, and recommends primary and backup sites if necessary.

6.2.2.8 Site Selection Results and Approval

The Operational Test Director then initiates routing of the Site Selection Memorandum to the appropriate SSA. The Site Selection Memorandum serves as the final record for site selection and as a reference document in the corresponding test plan (e.g., OTP, DCAP).

6.2.2.9 Test Site Coordination and In-Briefs

The OT team coordinates regularly with the test sites from site selection through the completion of the test event. Coordination can consist of informal communication by email or phone over matters such as tester badging, test dates, and the current status of the test. Site coordination also includes formal, in-person meetings to inform the site of upcoming test details. Prior to the start of OT at an airport, the Operational Test Director coordinates and provides the airport Federal Security Director (FSD) management team with an in-brief explaining the roles and responsibilities throughout the test, data being collected, and anything else that should be expected. The session also serves as a forum for resolving any remaining questions and issues prior to the start of testing.

The Operational Test Director is also responsible for coordinating a Law Enforcement Officer (LEO) in-brief prior to the start of any test activity involving TI testing. The LEO in-brief should discuss the frequency of planned threat inject testing as well as actions to be taken to terminate the trial without issue.

6.2.3 Establishing Baseline Capability

As determined by the System Evaluator, through coordination with the PMO, the OT team may travel to each of the test sites and perform baseline data collection prior to the start of OT. The objective of this activity is to determine the performance capability of current fielded systems. The baseline FDCA should be performed at the same sites and locations (e.g., terminal,

checkpoint) of the system that will be tested. Baseline data collection planning is typically documented in a DCAP or included in the OTP. The System Evaluator may use the data from baseline data collection as a reference for possible inclusion in the SER.

6.2.4 Operational Testing Test Readiness Notification

The OT TRN describes the ground rules for vendor presence and participation during an OT event. The Operational Test Director will coordinate the OT TRN through the PMO COR and will transmit the TRN to the vendor at the beginning of burn-in.

The OT TRN includes the procedures for a vendor to gain access to the TSE prior-to and during test. There are several activities at the test site where a vendor's presence may be required. The vendor may be authorized to support one or all of the following activities:

- *Training*: Vendor onsite with OTD teams to provide TSE training.
- *Configuration Audit Support*: Vendor onsite with TIOB, Acceptance Test, and/or OTB teams to support the configuration audit.
- *TSE Installation*: Vendor onsite to install the system in coordination with the DLD.
- *SAT/ISAT*: Vendor onsite to support the Acceptance Test Team-led SAT/ISAT.

Once the test begins, the primary need for vendor onsite participation is preventive maintenance and corrective maintenance. Vendors must coordinate preventive maintenance activities with the Operational Test Director prior to the test. Prior to each scheduled preventive maintenance activity, the vendor should coordinate with the OT team to gain access to the system. TSA will notify the vendor through the established maintenance concept when TSA requires corrective maintenance. The OT team must monitor all preventive maintenance and corrective maintenance during OT. The OT TRN explains to the vendor the following test protocols and restrictions.

- A vendor is not authorized to access the TSE without prior coordination with the Operational Test Director or the OT team.
- A vendor must not change their typical maintenance protocol (e.g., increasing the number of maintenance personnel at a test site) to appear more responsive during the test period.
- During maintenance activities, the OT team monitors to ensure only the required maintenance is performed. The OT team documents any attempt to review system logs, change system configuration, or update system files.
- The System Evaluator must approve any release of test results and observations outside of DAG meetings. OT team members are not authorized to give test status reports, system performance results, FDRS data, etc., to vendors or others.
- At no time will vendor conversations with the OT team be construed as direction from the Government unless coordination and concurrence is obtained from the COR.
- Vendors are not authorized to contact or interrogate TSOs prior to, during, or after shifts.
- No vendor photography/videos may be taken of the SUT.

All vendors supporting an OT event are required to formally sign the OT TRN.

6.2.5 Operational Test Readiness Reviews

The OTB utilizes a four-level OTRR progression for systems going to OT or an OA:

- *OTRR #1*: OTB Manager approval to enter burn-in.
- *OTRR #2*: CAE approval to enter burn-in.
- *OTRR #3*: OTB Manager approval to enter OT.
- *OTRR #4*: CAE approval to enter OT.

OTRR #1 and OTRR #3 are internal, T&E team final reviews prior to OTRR #2 and OTRR #4, respectively. The goal of OTRR #1 and OTRR #3 is to ensure that documentation, programmatic and system readiness, and test resources are in place to proceed to the next test phase. OTB will not coordinate scheduling for OTRR #2 and OTRR #4 without satisfactorily completing OTRR #1 and OTRR #3, respectively. Voting members of OTRR #1 and OTRR #3include the System Evaluator and Operational Test Director with the final approval belonging to the OTB Manager. Additional stakeholders may be invited to OTRR #1 and OTRR #3 at the discretion of the Operational Test Director.

The T&E team is represented at OTRR #2 and OTRR #4 by the T&E Director (OTA), the OTB and EQAB Managers, the Operational Test Director, the System Evaluator, and other test entity stakeholders as required. Voting members of OTRR #2 and OTRR #4 include the user representative; Program Office representative; Operational Test Director; System Evaluator; OTB and EQAB Managers; T&E Director, OTA; and the DOT&E with final approval belonging to the CAE. OTRR voting members may delegate their vote when necessary.

6.2.5.1 Operational Test Readiness Review Checklist

Prior to any OTRR, the Operational Test Director should ensure the SET has reviewed the OTRR Checklist and all open items can be discussed with sufficient information to allow the OTRR Chair to make a determination. The OTRR Chair weighs the inputs as presented by the Operational Test Director to make a determination to proceed to the next level of testing. Each checklist item will have a color rating defined as follows:

- *Green*: Action completed.
- *Yellow*: Not started or incomplete; expected completion by the start of testing.
- *Red*: Not started or incomplete; a significant risk exists to an on-time completion which could delay the burn-in or OT start date.

The Operational Test Director must address any item that is in a yellow or red status for OTRR-1 and OTRR-3 to ensure green status for all items in OTRR #2 and OTRR #4. If green status is not possible due to event timing or another reason, the OTB Manager may approve that the Operational Test Director continue forward with scheduling OTRR #2 or OTRR #4.

The end product of any OTRR may also include required changes to ensure successful test execution, or a delay or cancellation of the OT.

6.2.5.2 Operational Test Readiness Review #1 (OTRR #1)

The Operational Test Director presents OTRR #1 to determine readiness to enter burn-in. OTRR #1 entrance criteria may include the following; however, criteria are subject to change:

- Test planning (TES Briefing, TEMP, TECB, SEP, and OTP) documents completed, approved, and signed.
- VV&A completed of any necessary test articles and documentation produced.
- Test resources, including equipment, personnel, and instrumentation, are in place.
- CERT and QT completed with favorable results as adjudicated by the user representative.
- SUT configuration documented and frozen.
- Integration and installation completed and validated (SAT and ISAT, as applicable).
- Site coordination complete to include site visits and FSD management team in-brief.
- System operator and test personnel training. Training is required for only a sufficient number of TSOs to conduct burn-in at OTRR #1 and OTRR #2.
- Training documentation, operator and maintenance manuals, and any required TSE login and password information for system access provided to test team.
- CONOPs and SOP for all modes of operation approved.
- Privacy Impact Assessment completed.
- Maintenance concept defined and communicated to all relevant parties.
- PMO-led SRR conducted.

6.2.5.3 Operational Test Readiness Review #2 (OTRR #2)

The Operational Test Director schedules the OTRR #2 with the TSA CAE following a successful OTRR #1. The checklist items should not vary from OTRR #1 to OTRR #2. The Operational Test Director requires OTB Manager concurrence prior to scheduling the OTRR #2.

6.2.5.4 Operational Test Readiness Review #3 (OTRR #3)

OTRR #3 is conducted at the end of burn-in once system operations are stable. The focus of OTRR #3 is to assess the system's readiness to enter the OT. During OTRR #3, OTB, EQAB, and other stakeholders (as needed) confirm readiness to support the OT. The Operational Test Director presents OTRR #3 to the OTB Manager. The OTRR #3 chair confirms the readiness to proceed to OTRR 4. In addition to the OTRR #1 and OTRR 2 entrance criteria, OTRR #3 entrance criteria include:

- All TSO training complete and user operations stable.

- System performance is stable.
- OT team data collection processes are viable, stable, and do not interfere with operations.
- Completed TI planning, to include LEO in-brief and TI site visit.
- All TI test articles received and stored at the site.

The outcome of the OTRR #3 is a decision that the system is ready for record test and to proceed to OTRR #4, to direct required changes to ensure successful record test execution, or to recommend (to the PMO) a delay or cancellation of record testing. Upon a decision that the system is ready for record test, the Operational Test Director schedules OTRR #4.

6.2.5.5 Operational Test Readiness Review #4 (OTRR #4)

The CAE chairs OTRR #4 with attendance by the DOT&E; the T&E Director, OTA; and representatives from key SET stakeholder organizations. Generally, OTRR #4 should be held no later than 1-week prior to the start of OT. If schedules or the status of events do not allow for this to occur, the OTRR #4 may be held anytime up to the day prior to the start of OT. OTRR #4 provides the forum for the CAE to approve the start of the OT.

Prior to scheduling OTRR #4, all of the items from OTRR #3 must be in green status and the Operational Test Director must be prepared to present any issues encountered during the burn-in stages for consideration prior to the CAE making an OTRR decision.

6.2.5.6 Test Readiness Review for Field Test and Field Data Collection Activities

The Operational Test Director schedules a Test Readiness Review (TRR) prior to burn-in, and one prior to record test, in support of a Field Test or FDCA. The Operational Test Director may also schedule a single TRR for shorter tests, tests with minimal measures, and when conditions are not expected to change between burn-in and record test. The TRR participants should include, at a minimum, the Operational Test Director, System Evaluator, and sponsoring organization (e.g., PMO). The OTB Manager is the TRR Chair for FDCAs.

6.3 Burn-In and Test Execution

During burn-in and test execution, the OT team collects data in accordance with the OTP, OAP, or DCAP while its intended users, in their operational environment, use the system in accordance with the approved SOP. Figure 6-2 describes the various activities taking place during test execution in support of an acquisition program OT.

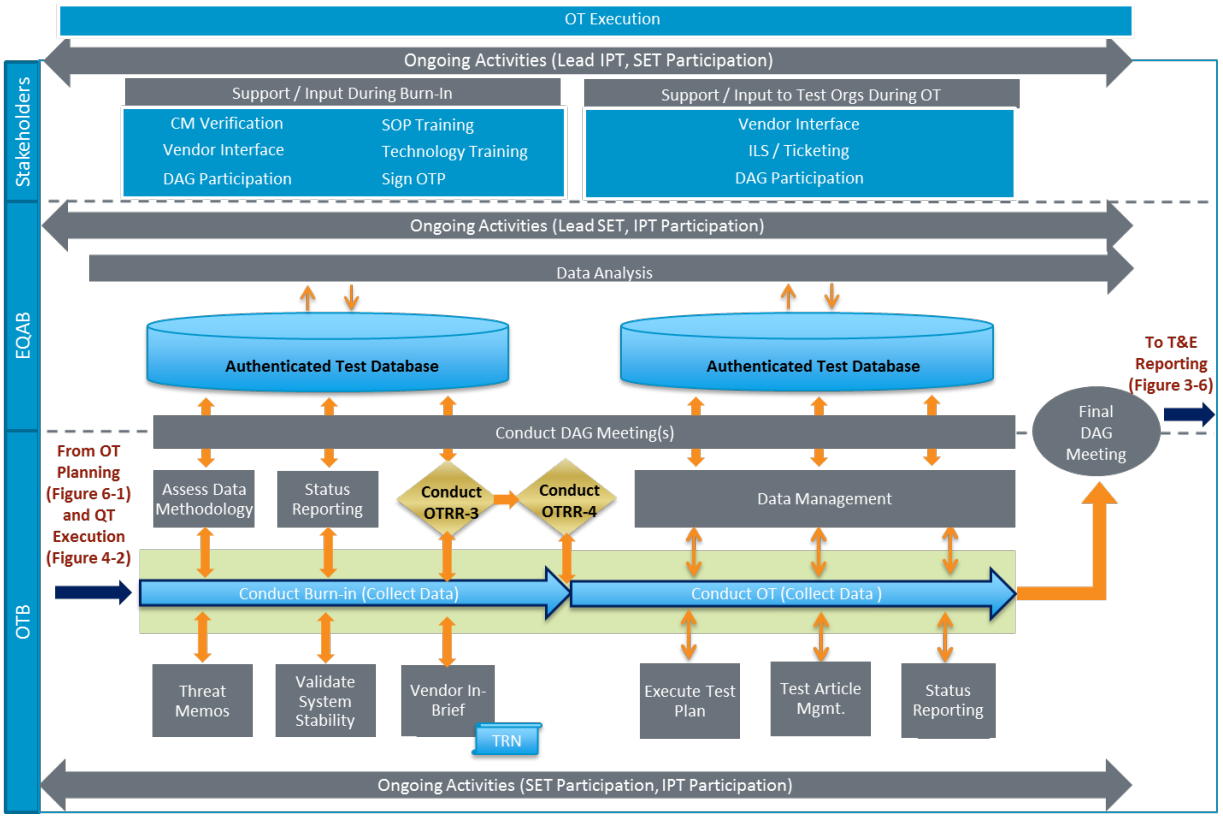

Figure 6-3: Operational Test Execution Process Flow

6.3.1 Burn-In

Immediately following successful installation, SAT, and CAE approval at OTRR-2, a system enters burn-in. During this period, operators use the system with the approved SOP to gain familiarity with the SUT and associated procedures. Burn-in also provides an opportunity for the test team to validate data collection methodologies and forms. As users become familiar with the system, test team data collection processes stabilize, enabling data integrity when record testing begins. Finally, burn-in provides the system an opportunity to demonstrate stable operational performance. The burn-in period duration is normally scheduled from seven to 30 days depending on the complexity of the system and integration. Burn-in duration may be shorter for FDCAs and Field Tests. Burn-in is event driven, meaning it will last as long as required to meet its purpose. If a system (technology or procedures) experiences considerable issues in the operational environment, a decision not to proceed to record testing may occur.

6.3.2 Test Execution and Data Collection

Once record testing begins, the OT team executes test procedures for throughput, TI testing, RMA, HSI, and other areas as described in the OTP, OAP, or DCAP. The test team collects data in a non-interference manner, however, they should intervene (notify the TSO) if observing a procedural or system deficiency that introduces a possible breach. The test team must document

any intervention as a test incident and report the intervention in the next daily status report, ensuring the PMO and user representative are copied.

OTB uses a TIR to capture essential data on test incidents as they occur. All system failures should have an associated TIR. It contains at a minimum, a description of the test incident, and any associated downtime and corrective action. The System Evaluator uses these data to resolve RMA measures in the SER. The OT team stores all TIRs in the test database and reviews them regularly during the DAG for authentication and use in the system evaluation. The OT team may also record TIRs not associated with system failures (e.g., improper vendor access of TSE, checkpoint closure or breach, TI incident, etc.).

6.3.3 Configuration Control During Test Execution

Configuration control throughout all phases of test ensures the SUT represents the vendor's intended configuration for a deployed system as established in the Program Baseline (PBL). Generally, the SUT configuration during OT should not change from the final configuration that passed testing at the TSIF. After the system has been installed and integrated in the operational environment, the TIOB CM team performs a pre-OT configuration audit. The TIOB CM team confirms the results of the audit against the PBL to ensure the system configuration has remained unchanged. The TIOB CM team may engage the Acceptance Test Team and/or onsite OT team in support of this audit. The TIOB CM team provides audit results to the Operational Test Director and continues to manage the PBL through completion of OT. This includes receipt of DEPs and dispositioning of the DEP in coordination with stakeholders, to include the CCB.

6.3.4 Early Termination of Operational Testing

In addition to the restrictions listed in the TRN (Section 6.2.5), other events could also cause the suspension of the OT event. These include the following:

- A critical safety deficiency arises.
- TSE disrupts, impedes, or degrades operations.
- Operator or technical procedures are not adequate, disrupt, or delay testing.
- Vendor inappropriately interferes with test or system configuration.
- The OT team discovers an unapproved system configuration discrepancy.

6.3.5 Test Closeout and Site Restoration

Upon completion of the OT, OTB works with the system integrator, the user organization, DLD, and airport authorities to remove test instrumentation and systems, to restore the normal screening TSE layout in a timely fashion, and to ensure minimum impact to local operations. In some cases, a joint decision is made to leave the test TSE at the site, in which case these site restoration activities will not take place. Upon request, airport officials may request an OT closeout briefing. The Operational Test Director should include in this briefing, a high level overview of test activities during the period and any pertinent site-specific topics. Detailed test results will not be released until formally reported in a SER, OAR, or SER.

6.4 Threat Inject Testing

6.4.1 Threat Inject Test Planning

OTB conducts TI testing when required in the SEP and OTP to assess whether the system (inclusive of operator and procedures) is effective at detecting threats. Proper planning ensures test integrity, risks are minimized, and that the required numbers of samples are collected.

The TIOB develops the threat test articles in support of TI testing. Unless otherwise specified in the OTP, the test articles should represent the spectrum of possible system threats (e.g., various detection tiers, prohibited items). The System Evaluator follows the *TSA VV&A Policy* to accredit the threat articles when possible.

The OT team creates the TI test methodology, Design of Experiments (DOE), and supporting data collection forms in conjunction with the System Evaluator and documented in the OTP. The scenarios listed in the DOE are typically staggered to ensure variability in times and the frequency of scenarios per day. The OTP should discuss the analysis methodology to include how a pass/fail determination will be reached. This should include comparison to thresholds, baseline system performance, or user adjudication post-test to determine whether the results are acceptable.

Additional critical planning elements include developing test scenarios with the user, test team training materials, memorandums to be used for ending trials, and test termination protocols

6.4.2 Threat Inject Execution

The Threat Inject Carrier (TIC) executes the scenarios in accordance with the OTP and TI training. If no detection is made, the nature of the test is not revealed. If a detection is made, procedures defined in the "Terminating the Test" section of training are followed.

During test execution, it is critical that the OT team take precautions to not compromise the TIC and test articles. The following components should not be revealed to operators prior to and during the test period to maintain test integrity:

- Test article storage location
- TIC identity
- Scenario details and schedule

When executing checkpoint TI trials, anonymity of TICs is paramount to test integrity. If a TIC's identity is revealed, TSOs may modify their typical screening behavior. The TIC should be prepared to answer any common questions asked by TSOs. The TIC should always remain calm, not drawing attention to the threat object even when discovered.

The OT team must always be cognizant of the TIC's identity, whereabouts, and time of execution of the TI scenario. Data collectors should not change their normal activities in the checkpoint during a TI test and should ensure the appropriate number of data collectors are present for each TI trial per the approved test plan.

6.4.3 Working with Threat Inject Data

TI data should be safeguarded appropriately per assigned classification to prevent vulnerability data from being released.

During test execution, the OT team collects TI results and observations each day and transmits them as appropriate for analysis and reporting. TI data collected daily is considered SSI. In addition to SSI guidelines, the OT team should take precautions when collecting the results of the test runs, encoding results whenever possible. As data begin to accumulate, and results/trends can be compiled, analysts should transmit the data to the HSDN network and delete any original files. The compiled data is classified as SECRET on the HSDN and handled in accordance with the approved DHS TSA Security Classification Guide as signed by the TSA Administrator.

6.4.4 Test Article Security

Test bags, threat objects, fraudulent documents, and any other test articles used during TI testing must be securely stored on airport grounds (unless otherwise approved by the TIOB Manager) and under OT team control at all times.

When shipping test articles, only a traceable shipping method from a reputable carrier must be used. Test bags and threat-based test articles (e.g., simulants, prohibited items) must be shipped from the TSIF (or other test article manufacturing location as approved by the TIOB Manager) with a DD1149 form completed to document chain of custody of the test articles. A member of the TIOB team must witness the bags palletized, shrink wrapped, and placed on the shipping vehicle prior to shipment. They must also receive a signed bill of lading detailing the shipment.

6.5 Operational Test Data Authentication

The mission of the DAG is to periodically review and authenticate data collected during the test event. Data authentication consists of a comprehensive review to ensure the data are complete, accurate, consistent, and representative of system performance within the tested operational environment. The DAG also ascertains whether the test team properly collected and reduced the data. The output of the DAG is an authenticated database used for data analysis and reporting.

6.5.1 Operational Test Data Authentication Group Personnel

The System Evaluator for the SUT chairs the DAG and is the final decision maker when authenticating data. The DAG is composed of a multi-disciplined team of representatives from the EQAB, OTB, user representative, and PMO. The on-site test team may be called upon to provide additional information regarding the data collected. DAG participants may not release data, nor disseminate test results, without the prior permission of the System Evaluator. Table 6-2 lists the minimum required DAG personnel and their role.

Table 6-2: Operational Test Data Authentication Group Personnel

Test Role	DAG Role
System Evaluator	DAG Chair
Lead Analyst	Presents data
Operational Test Director and Support Team	Provides details on collected data
User Representative	Participant
PMO Representative	Participant

6.5.2 Data Authentication Group Scoring Criteria

The DAG assigns scores to the data based upon data validation and verification, investigation results, and system knowledge. The test team records the authentication scores in the test database after the DAG to ensure authenticated data are used during final analysis and reporting. There are four authentication scores that the DAG can assign to each record:

- *Authenticated* data are complete, accurately represent operations, and can be used by the System Evaluator during final analysis and reporting.
- *Investigation* data are those records that require further research by the OT team before the DAG can score the data. The OT team must resubmit these records to the DAG when the investigation is complete for scoring.
- *Limited Use* data are those records that contain partially authenticated data elements and can be used for certain analyses, but not others.
- *No Test* data are non-representative, incomplete, or erroneous and not used for analysis.

6.5.3 Data Authentication Group Procedures

Prior to the DAG meeting, the OT team sends all DAG participants a data file consisting of the data to review for the DAG session. Typically, only flagged data are distributed in order for the test team to maintain control of SUT data prior to final reporting. The OT team instructs DAG participants to review the data prior to the meeting to expedite the review.

At the DAG meeting, the lead analyst, presents the flagged data file to include all data collected since the previous DAG data cutoff and all records previously marked as *Investigation*. The DAG reviews the records and recommends a score to the System Evaluator. The System Evaluator assigns the final DAG score. Any inquiries about the data records are directed to the Operational Test Director, who may need to conduct an investigation and provide the answers during the next DAG meeting.

APPENDIX A. ACRONYM DEFINITIONS

AA	Assistant Administrator
AA	Adversarial Assessment
ADA	Acquisition Decision Authority
ADE	Acquisition Decision Event
AI	Additional Issue
ALF	Acquisition Life Cycle Framework
ANSI	American National Standards Institute
APB	Acquisition Program Baseline
APM	Acquisition Program Management
ARP	Acquisition Review Process
ASTM	American Society for Testing Materials
AT	Acceptance Test
AT-CB	As-Tested Configuration Baseline
ATO	Authority to Operate
BCM	Baseline Correlation Matrix
BHS	Baggage Handling System
CAE	Component Acquisition Executive
CAR	Capability Analysis Report
CASP	Capability Analysis Study Plan
CBIS	Checked Baggage Inspection System
CC	Chief Counsel
CCB	Change Control Board
CERT	Certification Testing
CI	Configuration Item
CM	Configuration Management
CMD	Capabilities Management Division
CMP	Configuration Management Plan
CO	Contracting Officer
COI	Critical Operational Issue
COIC	Critical Operational Issues and Criteria
CONOPs	Concept of Operations
COOP	Continuity of Operations
COR	Contracting Officer Representative

CTP	Critical Technical Parameter
CVPA	Cooperative Vulnerability and Penetration Assessment
DAG	Data Authentication Group
DBL	Developmental Baseline
DCAP	Data Collection Activity Plan
DCAR	Data Collection Activity Report
DCCB	Developmental Configuration Control Board
DEP	Developmental Engineering Proposal
DHS	Department of Homeland Security
DLD	Deployment and Logistics Division
DOE	Design of Experiments
DOT&E	Director, Office of Test and Evaluation
DPBCA	Developmental Product Baseline Configuration Audit
EBSP	Electronic Baggage Screening Program
EDS	Explosives Detection System
EQAB	Evaluation and Quality Assurance Branch
ERB	Emerging Results Briefing
ERF	Event Request Form
FAT	Factory Acceptance Test
FAT&E	First Article Test and Evaluation
FCA	Functional Configuration Audit
FDRS	Field Data Reporting System
FOT&E	Follow-on Operational Test and Evaluation
FRD	Functional Requirements Document
FSD	Federal Security Director
HSDN	Homeland Security Domain Network
HSI	Human-Systems Integration / Human-Systems Interfaces
I&A	Intelligence and Analysis
IA	Information Assurance
IAD	Information Assurance and Cybersecurity Division
ICD	Interface Control Document
ID	Identification Document / Insufficient Data
IEF	Integrated Evaluation Framework
ILSP	Integrated Logistics Support Plan
INS	Inspections
IOT&E	Initial Operational Test and Evaluation

IPT	Integrated Product Team
IRD	Interface Requirements Document
ISAT	Integrated Site Acceptance Test
IT	Information Technology
IT&E	Independent Test and Evaluation
KPP	Key Performance Parameter
LA	Legislative Affairs
LBA	Lead Business Authority
LCCE	Life Cycle Cost Estimate
LEO	Law Enforcement Officer
LOA	Letter of Assessment
M&S	Modeling and Simulation
MAD	Mission Analysis Division
MAOL	Master Acquisition Oversight List
MCIL	Master Configuration Items List
MD	Management Directive
MNS	Mission Needs Statement
NAT	Network Acceptance Test
NT	Not Tested
OA	Operational Assessment
OAP	Operational Assessment Plan
OAR	Operational Assessment Report
OEM	Original Equipment Manufacturer
ORD	Operational Requirements Document
ORT	Operational Readiness Test
OSHE	Occupational Safety, Health, and Environment
OT	Operational Test
OT&E	Operational Test and Evaluation
OTA	Operational Test Agent
OTB	Operational Test Branch
OTA	Operational Test Agent
OTP	Operational Test Plan
OTRR	Operational Test Readiness Review
PARM	Program Accountability and Risk Management
PBL	Program Baseline
PGDS	Planning Guidelines and Design Standards

PIA	Post-ISAT Audit
PLC	Programmable Logic Controller
PM	Program Manager
PMO	Program Management Office
POAM	Plan of Action and Milestones
PoC	Proof of Concept
POC	Point of Contact
PRA	Preliminary Risk Assessment
PRR	Product Readiness Review
PSP	Passenger Screening Program
QDP	Qualification Data Package
QLR	Quick Look Report
QPL	Qualified Products List
QT	Qualification Testing
QTP	Qualification Test Plan
QTR	Qualification Test Report
QTRR	Qualification Test Readiness Review
QVP	Qualification Verification Package
RA	Risk Assessment
RCA	Requirements and Capability Analysis
RCM	Requirements Crosswalk Matrix
RFD	Request for Deviation
RFP	Request for Proposal
RMA	Reliability, Maintainability, and Availability
RTS	Recommendation of Test Scope
S&T	Science and Technology
SAM	Staffing Allocation Model
SAT	Site Acceptance Test
SCPA	Strategy, Communications and Public Affairs
SELC	System Engineering Life Cycle
SEP	System Evaluation Plan
SER	System Evaluation Report
SET	System Evaluation Team
SME	Subject Matter Expert
SOP	Standard Operating Procedures
SOS	Systems Optimization and Support

SRR	System Readiness Review
SSA	Site Selection Authority
SSI	Sensitive Security Information
SSTP	Site-Specific Test Plan
SSW	Site Selection Worksheet
STIP	Security Technology Integrated Program
STSO	Supervisory Transportation Security Officer
SUT	System under Test
T&D	Training and Development
T&E	Test and Evaluation
TECB	Test and Evaluation Concept Briefing
TES	Test and Evaluation Strategy
TEMP	Test and Evaluation Master Plan
TI	Threat Inject
TIC	Threat Inject Carrier
TIOB	Test Infrastructure and Operations Branch
TIR	Test Incident Report
TRN	Test Readiness Notification
TRR	Test Readiness Review
TSA	Transportation Security Administration
TSE	Transportation Security Equipment
TSIF	TSA Systems Integration Facility
TSL	Transportation Security Laboratory
TSM	Transportation Security Manager
TSO	Transportation Security Officer
TSR	Test Summary Report
TSS-E	Transportation Security Specialist-Explosives
VRTM	Verification Requirements Traceability Matrix
VV&A	Verification, Validation, and Accreditation

APPENDIX B.TEST AND EVALUATION PLANS, REPORTS, BEST PRACTICES AND OTHER SUPPORTING DOCUMENTATION

Current Test and Evaluation and Quality Assurance Branch, Operational Test Branch, and Transportation Security Administration (TSA) Systems Integration Facility (TSIF) Test Branch templates and best practices are located in the TSA iShare repository. Please contact the test organization manager to request access to the repository.

APPENDIX C. REFERENCES

Department of Homeland Security

- Acquisition Management Directive 102-01, Revision 03, July 28, 2015
- Acquisition Management Instruction 102-01-001, Revision 01, Incorporating Change 1, May 03, 2019
- DHS Directive 026-06, Test and Evaluation, Revision 01, May 05, 2017
- Memorandum: Operational Test and Evaluation (OT&E) Planning, Execution, and Reporting, August 06, 2015
- Memorandum: Procedures for Operational Test and Evaluation of Cybersecurity, October 15, 2015
- DHS Supplemental Guidance, Threat Assessment in Support of T&E, Version 1.0, September 2018
- DHS Supplemental Guidance, Cyber Resilience T&E, Version 2.0, October 2019
- DHS Instruction Guide 02606-001-01, Revision 1, Test and Evaluation Master Plan (TEMP), October 3, 2018

Transportation Security Administration

- Test and Evaluation Policy, Revision 3.0, August 13, 2018
- Office of Security Capabilities Test and Evaluation Verification, Validation and Accreditation (VV&A) Policy, Version 1.0, May 27, 2014
- Office of Security Capabilities Test and Evaluation Statistical Methodology Policy Memorandum, April 04, 2016
- Configuration Management Plan for the Office of Acquisition Program Management, Revision E, August 31, 2018
- Acceptance Testing Process Guide, Version 1.0, January 30, 2014
- Event Request Form, https://apps2010.ishare.tsa.dhs.gov/sites/OSC/eReview/Pages/eReviewLandingPage.aspx

APPENDIX D. TEST SCOPE CHANGE PROCEDURES

The following steps demonstrate the process for changing the scope of a Field Test and Field Data Collection Activity (FDCA) when the Program Office (or other test sponsor) determines a change in test scope/requirements is needed.

1. The Test and Evaluation (T&E) Manager updates the Event Request Form (ERF) with change scope/focus.
2. The Operational Test (OT) Director provides inputs in support of cost impact and input on when the data collection team will be ready to collect the new data.
3. The System Evaluator provides technical impact (e.g., estimated sample sizes and confidence/range of results).
4. The T&E Manager authorizes request and updates the ERF. ERF is the official record of the change/added scope.
5. The Operational Test Director develops internal test approach/plan for meeting the change/added scope. The signed test plan will not be changed.
6. The Operational Test Director notifies the System Evaluation Team (SET) of change implementation date.
7. The Data Collection Activity Report (DCAR) will include a section discussing the deviation from the original test plan.

APPENDIX E. CYBER RESILIENCE EVALUATION GUIDANCE

The overall objective of the operational Cyber Resilience evaluation is to ensure the System Under Test (SUT) is able to withstand and respond to realistic threat representative cyber-attacks while continuing to support mission operations. The system comprises the hardware, software, operators, maintainers, and the training and Standard Operating Procedures (SOPs) used to carry out the Concept of Operations (CONOPs). The Cyber Resilience evaluation includes test activities to gather sufficient data to identify all significant vulnerabilities of the SUT. Results of the Cyber Resilience test activities will be used to determine the operational Cyber Resilience of the system.

The Cyber Resilience evaluation strategy should align with the following departmental guidance:

- Procedures for Operational Test and Evaluation of Cybersecurity, October 15, 2015
- DHS Supplemental Guidance. Threat Assessment in Support of T&E, Version 1.0, September 2018
- DHS Supplemental Guidance. Cyber Resilience T&E, Version 1.0, July 2018

Cyber Resilience Threat Assessment

In accordance with Department of Homeland Security (DHS) Supplemental Guidance, Threat Assessment in Support of T&E, the Requirements Sponsor works with TSA intelligence groups and the Information Technology (IT) Information Assurance and Cybersecurity Division (IAD) to develop the threat assessment. The threat assessment is developed prior to the Test and Evaluation Master Plan (TEMP). Unclassified portions of the threat assessment such as the cybersecurity threats and course of action table (Table E-1) are included in this TEMP. Any classified portions detailing threat intelligence gathering should be documented separately in a classified briefing/document. The requirements sponsor should schedule and lead a threat assessment briefing for the Director, Operational Test and Evaluation (DOT&E). The threat assessment includes the following information:

- Which threat actors may target the mission(s) that the system supports?
- What is the intent (e.g., denial, degradation, attack, exfiltration)?
- What are the most likely and dangerous attack vectors based upon intent and capabilities?

The threat to system mission accomplishment includes vulnerabilities which inhibit the system's ability to fulfill the screening mission and pose a threat to the aviation transportation system. The TEMP should include likely threats to the SUT in a table as listed below.

Table E-1: Cyber Security Threats to CPSS

Actor	Intent	Capability	Most Likely Course of Action	Most Dangerous Course of Action
Hacktivist / Insider / Terrorist / Nation State	Exfiltration / Degradation/ Attack / Denial	Basic / Moderate / Advanced	Likely course of action	Most dangerous course of action

The Operational Test Agent (OTA) develops adversary scenarios upon review of the threat assessment. The adversary scenarios should indicate the highest risk scenarios from the user representative's viewpoint and clearly denote the user representative's single highest risk scenario. Threat scenarios may be revised into the future as the threat landscape could change.

Table-Top Exercise

The OTA, working with IAD staff, convenes a table-top exercise to assist in identifying the threats to the system, further define SUT threat vectors, and to begin documenting the test scenarios to be used to evaluate the system. An element of the table-top exercise is to leverage system architecture diagrams and the threat assessment to document the attack surface and kill chain. These elements will become the foundation for the development of Cyber Resilience test activity plans and are further described below.

- Identification and analysis of attack surface: The different points in a system architecture where an attacker could gain entry to compromise a system.
- Identification and analysis of kill chain: The sequence of adversary activities that take advantage of reachable and exploitable vulnerabilities as shown in the attack surface.

Cyber Resilience T&E Activities

The IAD team will support the System Evaluator's operational Cyber Resilience determination by conducting a Risk Assessment (RA) of the SUT. The RA may include any of the following activities. Note: these test activities may apply to systems regardless of whether they are intended to be connected to a network. However, the impact rating of vulnerabilities may vary based on the SUT configuration (e.g., network-enabled). T&E stakeholders should avoid using the term "standalone" when describing the SUT's Cyber Resilience T&E strategy and evaluation planning.

1. Preliminary Risk Assessment (PRA) – The PRA is an initial analysis of an information system used to identify potential vulnerabilities, determine the extent of the potential threat, and assess the risk to the system throughout its life cycle. The PRA is used to determine if existing countermeasures and safeguards adequately reduce the probability of loss and risk to an acceptable level and help validate the need for additional cost-effective countermeasures. The DHS 4300A *Sensitive Systems Handbook* requires that a risk assessment be conducted to provide information to support the decision to formally accredit the systems under the Authorizing Official's responsibility. The PRA includes a limited security controls assessment. The System Evaluator anticipates this activity occurring prior-to, or during Qualification Testing (QT).
2. Physical Security Assessment – The IAD team conducts a physical security assessment in conjunction with the PRA, which will include assessing open ports. Test teams will try to use any exposed ports for exfiltration of scanned images, system configuration files stored on the system, system reports, any other system data, and/or to otherwise gain unauthorized access to, and exploit the system. The System Evaluator anticipates this activity occurring prior-to, or during QT.

3. Cooperative Vulnerability and Penetration Assessment (CVPA) – The IAD team will conduct vulnerability scans using software tools. The System Evaluator anticipates the scans to occur prior-to, or during QT. As part of the CVPA, the IAD team will also conduct a penetration test against the SUT. As part of the penetration test, IAD test teams will try to penetrate the system by accessing the system's critical settings, code, and data. The IAD team will exploit vulnerabilities in an attempt to gain access to the systems configuration files and critical settings. The System Evaluator anticipates this activity occurring prior-to, or in parallel with the Operational Test (OT).
4. Authority to Operate (ATO) - When network connected, test teams support all pre-requisite activities and documentation to ensure issuance of a TSA network ATO. The SUT undergoes security authorization activities which will encompass a full security controls assessment. The System Evaluator anticipates these activities to occur prior-to OT.

Cyber Resilience Evaluation Reporting

IT IAD will compile results from the Cyber Resilience test activities, assign severity ratings to any vulnerabilities, and provide the results to the System Evaluator. The System Evaluator convenes Cyber Resilience results sessions with Integrated Product Team (IPT) stakeholders to discuss the findings, false-positives, remediation, and discuss any resulting Plans of Action and Milestones (POAMs). IT IAD documents the full findings and recommendations.

The System Evaluator leverages IT IAD's findings when determining the SUT's level of operational Cyber Resilience. The basis for the final Cyber Resilience rating is documented in the System Evaluation Plan (SEP). In general, one or more open/un-waivered critical or high vulnerability may result in a "Not Cyber Resilient" rating. In all cases, the System Evaluator considers the mission impact of all vulnerabilities that may allow an attacker to penetrate, pivot, or otherwise exploit the SUT when determining the level of operational Cyber Resilience. To be operationally cyber resilient, the system must remain effective and suitable to satisfy the mission while under duress from offensive cyber-attacks.

APPENDIX F.INSTRUCTIONS FOR USING THE EXTERNAL DATA SOURCE EVALUATION CHECKLIST

The System Evaluator follows the guidelines below when using the worksheet.

1. On vendor QVP submission, place an "X" in Column B of the *Data Source Risk* worksheet for each applicable assessment criteria. This will auto populate “Applicable Criteria” in the *Scoring Summary*.
2. On vendor QVP submission, identify the data source risk level by providing a “0”, “1”, or “2” score in Column G of the *Data Source Risk* worksheet.
3. On vendor QVP submission, assess whether an adequate test plan was provided by indicating “yes” or “no” on the *Test Plan Scorecard (QVP)* worksheet, by requirement, for each of the following sub areas. Note: if an area is not applicable, mark “yes.”
    a. Acceptable test methodology
    b. Acceptable sample size
    c. Acceptable CM approach
    d. Acceptable test environment
    e. No significant limitations or constraints
    f. Acceptable test articles
4. On vendor QDP submission, assess the data provided for each requirement by indicating “yes” or “no” on the *Test Data Scorecard (QDP)* for each of the following sub areas. Note: if an area is not applicable, mark “yes.”
    a. Acceptable test methodology
    b. Acceptable sample size
    c. Tested on representative configuration?
    d. Acceptable test environment
    e. Acceptable test articles

The System Evaluator can review the overall data source risk ratings and requirements compliance on the *Scoring Summary* worksheet.

APPENDIX G.TEST AND EVALUATION DOCUMENT RESPONSIBILITY AND ASSIGNMENT MATRIX

The Test and Evaluation (T&E) Document Responsibility Matrix listed in Table G-1 indicates responsibilities by T&E document.

Key		
Developed By and Approved	D/A	The individual responsible for developing the document. Through its development, the individual acknowledges they have fully reviewed, understand, and approve the document content.
Review Only	R	The individual(s) responsible for reviewing documents and providing feedback to the developer.
Approve	A	The individual(s) responsible for reviewing and approving the document either by signing/initialing the document and/or routing sheet. These approvals are typically required prior to final approval.
Final Approver	F	The individual responsible for providing the final approval for a document.
Distributed for Information	I	The individual(s) who receive the document after final approval for informational purposes.
Letter of Assessment (LOA)	L	The LOA is DOT&E's assessment of the adequacy of T&E as documented in the SER.

Table G-1: Document Responsibility Matrix

	T&E Staff			PMO		User	T&E Branch Managers			OTA		Other	
Document	**Evaluator**	**Qualification Test Lead**	**Operational Test Director**	**T&E Manager**	**Program Manager**	**User Representative**	**TSIF Test Branch Manager**	**OTB Manager**	**EQAB Manager**	**T&E Deputy Director**	**OTA Director, T&E**	**DOT&E**	**CAE**
Evaluation Strategy/Planning													
Risk Assessment Level of Testing (RALOT)	D/A	R	R	I	I	I	A	A	A	A	F	I	I
Test and Evaluation Strategy (TES) Briefing	A	R	R	D/A	A	A	I	I	I	I	I	F	A
Test and Evaluation Master Plan (TEMP)	R	R	R	D/A	A	A	R	R	R	R	A	F	A
Test and Evaluation Concept Brief (TECB)	D/A	R	R	R	R	R	A	A	A	A	A	F	
System Evaluation Plan (SEP) / Requirements Crosswalk Matrix (RCM)	D/A	R	R	R	R	R	A	A	A	A	F	I	
Test Event Planning													
Qualification Test Plan (QTP)	A	D/A	I	R	R	R	A	I	I	A	F	I	

Document	T&E Staff			PMO		User	T&E Branch Managers			OTA		Other	
	Evaluator	Qualification Test Lead	Operational Test Director	T&E Manager	Program Manager	User Representative	TSIF Test Branch Manager	OTB Manager	EQAB Manager	T&E Deputy Director	OTA Director, T&E	DOT&E	CAE
TSIF Data Collection Activity Plan (DCAP)*		D/F		I	I		F						
QTRR Slides	R	D/A	I	R	I	I	F	I	I	I	I		
Site Selection Memorandum (OT/FOT&E)	F		D/A	R	A	R		A		I	I	I	
Site Selection Memorandum (Field Test)	A		D/A	R	F	R							
Test Readiness Notification (TRN)			D/A	A				F					
Test Article Validation Memorandum	D/A		R	R	R	F							
Threat Inject Authorization Letter			D/A					A		A	F		
DCAP (Field Data Collection Activities and Field Tests)	A		D/A	R	R	R		F	A				
Operational Assessment Plan (OAP)**	A		D/A	R	R	R		A	A	A	A	F	
Operational Test Plan (OTP)**	A		D/A	R	R	R		A	A	A	A	F	
OTRR 1 and 3 Slides	A		D/A	R	R	R		A	A		F		
OTRR 2 and 4 Slides	A		D/A	R	A	A		A	A	R	A	A	F
Reporting													
QT Quick Look Report (QLR)	R	D/A	I	I	I	I	F	I	I				
QT Report	R	D/A	I	R	R	R	F	I	I				
TSIF Data Collection Activity Report (DCAR)*		D/F		I	I		F						
Operational Test (OT) QLR	D/A		R	I	I	I		A	F				
OT Emerging Results Briefing (ERB)	D/A		R	R	R	R		A	A	A	A	F	
Operational Assessment Report (OAR)	D/A		R	R	R	R		A	A	A	A	F	
DCAR (Field Data Collection Activities and Field Tests)	D/A		R	R	R	R		A	F				
System Evaluation Report (SER)	D/A	I	R	R	R	R	I	A	A	A	F	L	I

Note: APM Communications routing is optional. It is acceptable for the document developer to send a copy once the document is final.
* Final signature may be required from the TSIF Branch Manager on a case-by-case basis.
**Assumes OT&E or OA is under DOT&E oversight. If not, then signature authority lies with the OTA Director, T&E.