

GAO

Report to the Chairman, Committee on  
Homeland Security, House of  
Representatives

August 2008

# TRANSPORTATION SECURITY

TSA Has Developed a  
Risk-Based Covert  
Testing Program, but  
Could Better Mitigate  
Aviation Security  
Vulnerabilities  
Identified Through  
Covert Tests





Highlights of [GAO-08-958](#), a report to the Chairman, Committee on Homeland Security, House of Representatives

## Why GAO Did This Study

The Transportation Security Administration (TSA) uses undercover, or covert, testing to approximate techniques that terrorists may use to identify vulnerabilities in and measure the performance of airport security systems. During these tests, undercover inspectors attempt to pass threat objects through passenger and baggage screening systems, and access secure airport areas. In response to a congressional request, GAO examined (1) TSA's strategy for conducting covert testing of the transportation system and the extent to which the agency has designed and implemented its covert tests to achieve identified goals; and (2) the results of TSA's national aviation covert tests conducted from September 2002 to June 2007, and the extent to which TSA uses the results of these tests to mitigate security vulnerabilities. To conduct this work, GAO analyzed covert testing documents and data and interviewed TSA and transportation industry officials.

## What GAO Recommends

To ensure that TSA is more fully using the results of covert tests, GAO recommends that TSA document causes of test failures; as TSA explores the use of covert testing in non-aviation modes of transportation, coordinate with transportation organizations that conduct covert tests; and develop a systematic process to evaluate covert testing recommendations. DHS and TSA reviewed a draft of this report and concurred with GAO's recommendations.

To view the full product, including the scope and methodology, click on [GAO-08-958](#). For more information, contact Cathleen A. Berrick at (202) 512-3404 or [berrickc@gao.gov](mailto:berrickc@gao.gov).

## TRANSPORTATION SECURITY

### TSA Has Developed a Risk-Based Covert Testing Program, but Could Better Mitigate Aviation Security Vulnerabilities Identified Through Covert Tests

#### What GAO Found

TSA has designed and implemented risk-based national and local covert testing programs to achieve its goals of identifying vulnerabilities in and measuring the performance the aviation security system, and has begun to determine the extent to which covert testing will be used in non-aviation modes of transportation. TSA's Office of Inspection (OI) used information on terrorist threats to design and implement its national covert tests and determine at which airports to conduct tests based on the likelihood of a terrorist attack. However, OI did not systematically record the causes of test failures or practices that resulted in higher pass rates for tests. Without systematically recording reasons for test failures, such as failures caused by screening equipment not working properly, as well as reasons for test passes, TSA is limited in its ability to mitigate identified vulnerabilities. OI officials stated that identifying a single cause for a test failure is difficult since failures can be caused by multiple factors. TSA recently redesigned its local covert testing program to more effectively measure the performance of passenger and baggage screening systems and identify vulnerabilities. However, it is too early to determine whether the program will meet its goals since it was only recently implemented and TSA is still analyzing the results of initial tests. While TSA has a well established covert testing program in commercial aviation, the agency does not regularly conduct covert tests in non-aviation modes of transportation. Furthermore, select domestic and foreign transportation organizations and DHS components use covert testing to identify security vulnerabilities in non-aviation settings. However, TSA lacks a systematic process for coordinating with these organizations.

TSA covert tests conducted from September 2002 to June 2007 have identified vulnerabilities in the commercial aviation system at airports of all sizes, and the agency could more fully use the results of tests to mitigate identified vulnerabilities. While the specific results of these tests and the vulnerabilities they identified are classified, covert test failures can be caused by multiple factors, including screening equipment that does not detect a threat item, Transportation Security Officers (TSOs), formerly known as screeners, not properly following TSA procedures when screening passengers, or TSA screening procedures that do not provide sufficient detail to enable TSOs to identify the threat item. TSA's Administrator and senior officials are routinely briefed on covert test results and are provided with test reports that contain recommendations to address identified vulnerabilities. However, TSA lacks a systematic process to ensure that OI's recommendations are considered and that the rationale for implementing or not implementing OI's recommendations is documented. Without such a process, TSA is limited in its ability to use covert test results to strengthen aviation security. TSA officials stated that opportunities exist to improve the agency's processes in this area.

In May 2008, GAO issued a classified report on TSA's covert testing program. That report contained information that was deemed either classified or sensitive. This version of the report summarizes our overall findings and recommendations while omitting classified or sensitive security information.

---

# Contents

---

<b>Letter</b>		1
	Results in Brief	6
	Background	9
	TSA Has a Risk-Based Covert Testing Strategy to Identify Vulnerabilities and Measure the Performance of Selected Aviation Security Systems, but Could Strengthen Its Testing Efforts	20
	TSA Could More Fully Use the Results of Covert Tests to Mitigate Security Vulnerabilities Identified in the Commercial Aviation System	30
	Conclusions	33
	Recommendations for Executive Action	34
	Agency Comments and Our Evaluation	35
<b>Appendix I</b>	<b>Objectives, Scope, and Methodology</b>	39
<b>Appendix II</b>	<b>Comments from the Department of Homeland Security</b>	42
<b>Figures</b>		
	Figure 1: TSA's Passenger Checkpoint and Checked Baggage Screening Operations and Equipment	12
	Figure 2: Diagram of Security Areas at a Typical Commercial Airport	15

---

---

## Abbreviations

AAR	American Association of Railroads
AOA	Air Operations Area
APTA	American Public Transportation Association
ASAP	Aviation Screening Assessment Program
ATSA	Aviation and Transportation Security Act
CBP	Customs and Border Patrol
DHS	Department of Homeland Security
DNDO	Domestic Nuclear Detection Office
DOT	Department of Transportation
EDS	Explosive Detection System
ETD	Explosive Trace Detection
FAA	Federal Aviation Administration
FSD	Federal Security Director
IED	Improvised Explosive Device
IRD	Internal Reviews Division
OI	Office of Inspection
OSO	Office of Security Operations
SIDA	Security Identification Display Area
STEA	Screener Training Exercises and Assessments
TSA	Transportation Security Administration
TSNM	Transportation Sector Network Management
TSO	Transportation Security Officer
TS-SSPP	Transportation System Sector Specific Plan

This is a work of the U.S. government and is not subject to copyright protection in the United States. It may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



United States Government Accountability Office  
Washington, DC 20548

August 8, 2008

The Honorable Bennie G. Thompson  
Chairman  
Committee on Homeland Security  
House of Representatives

Dear Mr. Chairman:

The March 2004 bombings of the rail system in Spain, July 2005 bombings of London's subway system, and August 2006 alleged terror plot to bring liquid explosives through airport security checkpoints in the United Kingdom and detonate them on board aircraft bound for the United States, are striking reminders that transportation systems have continued to be a target for terrorist attack. After the September 11, 2001 terrorist attacks, the Aviation and Transportation Security Act (ATSA) was enacted, creating the Transportation Security Administration (TSA) and mandating that it assume responsibility for security in all modes of transportation.<sup>1</sup> For the last 5 years, TSA has spent billions of dollars to screen airline passengers and checked baggage and to implement regulations and initiatives designed to strengthen the security of commercial aviation. TSA has also taken action to strengthen the security of surface modes of transportation, which includes mass transit and passenger rail, freight rail, and highways. Despite varying levels of progress in these respective areas, questions remain about the effectiveness of TSA's security programs and procedures.

One method that can be used to identify and mitigate vulnerabilities, measure the effectiveness of security programs, and identify needed changes to training procedures and technologies is undercover, or covert testing—also known as red team testing—which was advocated by the President's July 2002 National Strategy for Homeland Security to identify security vulnerabilities in the nation's critical infrastructure and to help prepare for terrorist attacks.<sup>2</sup> Regarding aviation security, and in

---

<sup>1</sup>See Pub. L. No. 107-71, 115 Stat. 597 (2001).

<sup>2</sup>TSA defines a covert test at domestic airports as any test of security systems, personnel, equipment, and procedures to obtain a snapshot of the effectiveness of airport passenger security checkpoint screening, checked baggage screening, and airport access controls to improve airport performance, safety, and security.

---

accordance with requirements established in law, TSA conducts covert testing of passenger and checked baggage screening operations, as well as airport perimeter security and access controls, and requires that Transportation Security Officers (TSO) who fail tests to undergo remedial training.<sup>3</sup> Prior to the creation of TSA, the Federal Aviation Administration (FAA) was responsible for ensuring compliance with aviation screening regulations and testing the performance of passenger and checked baggage systems in detecting threat objects. TSA began conducting covert testing in commercial aviation in September 2002. Covert testing is conducted at the national level by TSA's Office of Inspection (OI) and at the local, or individual airport level by the Office of Security Operations (OSO)—the division within TSA responsible for overseeing passenger and checked baggage screening at airports. During these tests, undercover inspectors attempt to pass threat objects, such as simulated explosive devices, through airport passenger screening checkpoints and checked baggage screening systems. Inspectors also attempt to access secure areas of the airport undetected, such as through doorways leading to aircraft and the airport's perimeter. The tests are designed to approximate techniques that terrorists may use in order to identify vulnerabilities in the people, processes, and technologies that comprise the aviation security system. With respect to some non-aviation modes of transportation, specifically mass transit, passenger rail, and maritime ferries, TSA has initiated pilot programs designed to test the feasibility of implementing screening of passengers at a centralized checkpoint, similar to the aviation system. According to OI officials, during these pilot programs, OI conducted covert testing to determine if they could pass threat objects through the passenger screening procedures and equipment that was being tested in these systems. In addition, TSA's May 2007 Transportation System Sector Specific Plan (TS-SSP) for mass transit describes TSA's strategy for securing mass transit and passenger rail, and encourages that transit and rail agencies should develop covert testing exercises.

We have previously reported on the results of TSA's national and local aviation covert tests, both of which have identified vulnerabilities in the aviation security system, and the results of our investigators' tests of TSA's passenger checkpoint and checked baggage security systems, which have also identified vulnerabilities. The Department of Homeland Security

---

<sup>3</sup>As used in this report and unless otherwise specifically stated, the term TSO, which ordinarily refers only to the federal screening workforce, includes the private screening workforce at airports in the screening partnership program.

---

(DHS) Office of Inspector General has also conducted its own covert testing of airport passenger and checked baggage screening, as well as perimeters and access controls, and has also identified vulnerabilities in these areas, most recently in March 2007.<sup>4</sup>

In light of the security vulnerabilities that covert testing has identified and concerns regarding the effectiveness of existing security procedures, you asked that we review TSA's national and local covert testing programs. In response, on May 13, 2008, we issued a classified report addressing the following key questions: (1) What is TSA's strategy for conducting covert testing of the transportation system, and to what extent has the agency designed and implemented its covert tests to achieve identified goals? and (2) What have been the results of TSA's national aviation covert tests conducted from September 2002 to June 2007, and to what extent does TSA use the results of these tests to mitigate security vulnerabilities in the commercial aviation system?

As our May 2008 report contained information that was deemed to be either classified or sensitive, this version of the report is intended to generally summarize our overall findings and recommendations while omitting classified or sensitive security information about TSA's covert testing processes and the results of TSA's covert tests conducted from September 2002 to June 2007. As our intent in preparing this report is to convey, in a publicly available format, the non-classified, non sensitive results of the classified May 2008 report, we did not attempt to update the information here to reflect changes that may have occurred since the publication of the May 2008 report.

To identify TSA's strategy for conducting covert testing of the transportation system and the extent to which the agency has designed and implemented tests to achieve its goals, we reviewed applicable laws, regulations, policies, and procedures for national and local covert testing. We interviewed TSA OI officials responsible for conducting national aviation covert tests, and OSO officials responsible for local aviation covert tests, regarding TSA's strategy for designing and implementing these tests, including the extent to which they used threat information to guide their efforts. We also observed OI inspectors during covert tests at

---

<sup>4</sup>Department of Homeland Security Office of Inspector General, *Audit of Access to Airport Secured Areas*, OIG-07-35 (March 2007). The results of TSA's, GAO's, and the DHS Office of Inspector General's covert tests are all classified and cannot be presented in this report.

---

seven airports, including airports with heavy passenger traffic and those with just a few flights per day, as well as airports with both TSOs and contract screeners.<sup>5</sup> During these covert tests, we accompanied OI inspectors during all phases of the test including planning, testing, and post-test reviews with TSOs and their supervisors. We interviewed TSOs and their supervisors that were involved in covert tests at each airport where we observed tests to discuss their experience with the national and local covert testing programs. We also interviewed the Federal Security Director (FSD) at each airport where we observed covert tests to obtain their views of the testing program and the results of tests at their airports.<sup>6</sup> While these seven airports represent reasonable variations in size and geographic locations, our observations of OI's covert tests and the perspectives provided by TSA officials at these airports cannot be generalized across all commercial airports. However, our observations at the seven airports provided us with an overall understanding of how OI conducts covert tests and useful insights provided by TSOs, their supervisors, and FSDs at these airports. We also reviewed TSA's procedures for screening passenger and checked baggage to determine how these procedures are used in designing and implementing national aviation covert tests. We interviewed OI officials and officials from TSA's Office of Transportation Sector Network Management (TSNM), which is responsible for developing security policies for non-aviation modes of transportation, regarding the extent to which covert testing has been conducted in non-aviation modes, the applicability of covert testing in other modes, and future plans for conducting covert testing in other modes. To understand how entities outside of TSA have used covert testing in non-aviation modes of transportation, we interviewed officials from DHS components and organizations that conduct covert testing,

---

<sup>5</sup>In accordance with ATSA, TSA began allowing all commercial airports to apply to TSA to transition from a federal to a private screening workforce in November 2004. See 49 U.S.C. § 44920. To support this effort, TSA created the Screening Partnership Program to allow all commercial airports an opportunity to apply to TSA for permission to use qualified private screening contractors and private sector screeners. Currently, private screening companies provide passenger and checked baggage screening at 11 airports.

<sup>6</sup>Federal Security Directors (FSD) are the ranking TSA authorities responsible for leading and coordinating TSA security activities at the nation's commercial airports. TSA had 122 FSD positions at commercial airports nationwide, as of January 2008. Although FSDs are responsible for security at all commercial airports, not every airport has an FSD dedicated solely to that airport. Most large airports have an FSD responsible for that airport alone. Other smaller airports are arranged in a "hub and spoke" configuration, in which an FSD is located at or near a hub airport but also has responsibility over one or more spoke airports of the same or smaller size.



---

including U.S. Customs and Border Protection, DHS Domestic Nuclear Detection Office (DNDO), Amtrak, the United Kingdom's Department for Transport Security (TRANSEC), and transportation industry associations, such as the American Association of Railroads and the American Public Transportation Association.

To determine the results of TSA's national covert tests and the extent to which TSA used the results of these tests to mitigate security vulnerabilities in the aviation system, we obtained and analyzed a database of the results of TSA's national covert tests conducted from September 2002 to June 2007. To determine how TSA gathers covert testing data, we reviewed the data collection instruments used at the airports where we observed covert tests, as well as other methods OI uses to gather covert testing data and observations. We also reviewed TSA's internal controls for collecting and maintaining the results of covert tests. We assessed the reliability of TSA's covert testing data and the systems used to produce the data by interviewing agency officials responsible for maintaining the database. We determined that the data were sufficiently reliable for our analysis and the purposes of this report. We also interviewed OI officials regarding how the results of covert tests are used in developing their recommendations to TSA management. We reviewed OI reports on the results of covert tests completed between March 2003 and June 2007 that were submitted to TSA's Administrator and OSO to identify OI's recommendations for mitigating the vulnerabilities identified during covert tests. We further obtained and analyzed a summary of the actions that OSO had taken to address OI's recommendations for mitigating vulnerabilities made from March 2003 to June 2007. More detailed information on our scope and methodology is contained in appendix I.

We conducted this performance audit from October 2006 to May 2008, in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

---

## Results in Brief

TSA has designed and implemented risk-based national and local covert testing programs to achieve its goals of identifying vulnerabilities in and measuring the performance of passenger checkpoint and checked baggage screening systems and airport perimeters and access controls, and has begun to determine the extent to which covert testing will be used to identify vulnerabilities and measure the effectiveness of security practices related to non-aviation modes of transportation. OI used information on terrorist threats to design and implement its national covert tests and determine at which airports to conduct tests based on analyses of risks. However, OI inspectors did not systematically record specific causes for test failures related to TSOs, procedures, or screening equipment that did not work properly. Standards for Internal Control in the Federal Government identify that information should be recorded and communicated to management and others in a form and within a time frame that enables them to carry out their internal control and other responsibilities.<sup>7</sup> OI officials stated that they do not record information on equipment failures because there is a possibility that the threat item was not designed properly and therefore should not have set off the alarm, and identifying a single cause for a test failure is difficult since covert testing failures can be caused by multiple factors. OI also did not systematically collect and analyze information on effective screening practices that may contribute to TSOs' ability to detect threat items, which could allow TSA to identify actions that may help improve screening performance across the system. Without systematically recording reasons for test failures, such as failures caused by screening equipment not working properly, as well as reasons for test passes, TSA is limited in its ability to mitigate identified vulnerabilities. Regarding TSA's local aviation covert testing program, the agency recently redesigned the program to address the limitations of the previous program, such as inconsistent structure and frequency of tests across airports. The new program, called the Aviation Screening Assessment Program (ASAP), is also risk-based, with tests being designed to mirror threat items that may be used by terrorists based on threat information. If implemented effectively, ASAP should provide TSA with a measure of the performance of passenger and checked baggage screening systems and help to identify security vulnerabilities. However, since the program was only recently implemented, it is too soon to determine whether ASAP will meet its goals of identifying vulnerabilities and measuring the performance of passenger and checked baggage

---

<sup>7</sup>GAO, *Internal Control: Standards for Internal Control in the Federal Government*, GAO/AIMD-00-21.3.1 (Washington, D.C.: November 1999).

---

screening systems. Furthermore, TSA has just begun to determine the extent to which covert testing will be used to identify vulnerabilities and measure the effectiveness of security practices in non-aviation modes of transportation. While TSA coordinates with domestic and foreign transportation organizations and DHS component agencies regarding its security efforts, they do not have a systematic process in place to coordinate with these organizations regarding covert testing in non-aviation settings, and opportunities for TSA to learn from these organizations' covert testing efforts exist.

TSA's national aviation covert testing program has identified vulnerabilities in selected aspects of the commercial aviation security system at airports of all sizes, however, the agency is not fully using the results of these tests to mitigate identified vulnerabilities. The specific results of these tests are classified and are presented in our classified May 2008 report. Although national covert test results provide only a snapshot of the effectiveness of airport security screening and cannot be considered a measurement of performance because the tests were not conducted using the principles of probability sampling, tests can be used to identify vulnerabilities in the commercial aviation security system. Covert test failures have been caused by various factors, including TSOs not properly following TSA procedures when screening passengers, screening equipment that does not detect a threat item, and TSA screening procedures that do not provide sufficient detail to enable TSOs to identify a threat item. Senior TSA officials, including TSA's Administrator, are routinely briefed on the results of covert tests and provided with OI reports that describe the vulnerabilities identified by these tests and recommendations to correct identified vulnerabilities. However, OSO, the office within TSA responsible for passenger and checked baggage screening, lacks a systematic process to ensure that OI's recommendations are considered, and does not systematically document its rationale for why it did or did not implement OI's recommendations. From March 2003 through June 2007, OI made 43 recommendations to OSO designed to mitigate vulnerabilities identified through covert tests. These recommendations related to providing additional training to TSOs and revising or clarifying existing TSA screening procedures, such as procedures for screening passengers and checked baggage. To date, OSO has taken actions to implement 25 of OI's 43 recommendations. OSO and OI also do not have a process in place to assess whether the corrective action implemented mitigated the identified vulnerabilities through follow-up national or local covert tests. According to OSO officials, TSA has other methods in place to identify whether corrective actions or other changes are effective; however, officials did not provide specific information

---

regarding these methods. For the remaining 18 of OI's 43 recommendations, OSO either took no action to address the recommendation or it is unclear how the action they took addressed the recommendation. Moreover, in those cases where OSO took no action to address OI's recommendation, they did not systematically document their rationale for why they took no action. Standards for Internal Control in the Federal Government identify that managers are to promptly evaluate findings, determine proper actions, and complete these actions to correct matters. In the absence of a systematic process for considering OI's recommendations, documenting their decision-making process, and evaluating whether corrective actions mitigated identified vulnerabilities, TSA is limited in its ability to use covert testing results to improve the security of the commercial aviation system. According to OSO officials, opportunities exist to improve OSO's internal processes for considering OI's recommendations and documenting its rationale for implementing or not implementing these recommendations. OSO senior leadership stated that they were committed to enhancing its partnership with OI and improving its processes for ensuring that OI recommendations are communicated to and considered by the appropriate TSA officials.

To better ensure that TSA is fully using the results of covert tests to identify and mitigate vulnerabilities in the transportation security system, we recommended in our May 2008 classified report that the Secretary of Homeland Security direct the Assistant Secretary for TSA to develop a systematic process for gathering and analyzing specific causes of all national aviation covert testing failures, record information on screening equipment that may not be working properly during covert tests, and identify effective practices used at airports that perform well on covert tests. In addition, as TSA explores the use of covert testing in non-aviation modes of transportation, we recommended that the agency coordinate with organizations that already conduct these tests to learn from their experiences. Further, we recommended that the Secretary of Homeland Security direct the Assistant Secretary for TSA to develop a systematic process to ensure that OSO considers all recommendations made by OI as a result of covert tests and systematically documents their rationale for either implementing or not implementing these recommendations. Finally, we recommended that when OSO implements OI's recommendations, they should evaluate whether the action taken has addressed the vulnerability identified through the covert tests, which could include the use of follow-up national or local covert tests or through other means determined by OSO.

---

We provided a draft of this report to DHS and TSA for review. DHS, in its written comments, concurred with the findings and recommendations in the report. The full text of DHS's comments is included in appendix II.

---

## Background

Congress and the Administration have advocated the use of covert or red team testing in all modes of transportation. Following the terrorist attacks on September 11, 2001, on November 19, 2001, the President signed ATSA into law, with the primary goal of strengthening the security of the nation's commercial aviation system.<sup>8</sup> ATSA created TSA within the Department of Transportation (DOT) as the agency responsible for securing all modes of transportation. Among other things, ATSA mandated that TSA assume responsibility for screening passengers and their property, which includes the hiring, training, and testing of the screening workforce.<sup>9</sup> ATSA also mandated that TSA conduct annual proficiency reviews and provide for the operational testing of screening personnel, and that TSA provide remedial training to any screener who fails such tests. In 2002, the President issued The National Strategy for Homeland Security that supports developing red team tactics in order to identify vulnerabilities in security measures at our Nation's critical infrastructure sectors, including the transportation sector. In 2007, TSA issued the TS-SSP that outlines its strategy and associated security programs to secure the transportation sector.<sup>10</sup> While the TS-SSP does not address covert testing in aviation, it

---

<sup>8</sup>Pub. L. No. 107-71, 115 Stat. 597 (2001).

<sup>9</sup>In accordance with ATSA, TSA assumed operational responsibility from air carriers for screening passengers and checked baggage for explosives at more than 450 commercial airports by November 19, 2002. Passenger screening is a process by which authorized personnel inspect individuals and property at designated screening locations to deter and prevent carriage of any unauthorized explosive, incendiary, weapon, or other dangerous item into a sterile area or aboard an aircraft. Sterile areas are located within the terminal and generally include areas past the screening checkpoint where passengers wait to board, or into which passengers deplane from, a departing or arriving aircraft. Checked baggage screening is a process by which authorized personnel inspect checked baggage to deter, detect, and prevent the carriage of any unauthorized object on board an aircraft. The Homeland Security Act of 2002, signed into law on November 25, 2002, transferred TSA from DOT to DHS. See Pub. L. No. 107-296, § 403, 116 Stat. 2135, 2178.

<sup>10</sup>TSA's TS-SSP describes the security framework that will enable TSA to prevent and deter acts of terrorism using or against the transportation system, enhance the resilience of the transportation system, and improve use of resources for transportation security, among other things. TS-SSP establishes TSA's strategic approach to securing the transportation sector in accordance with the National Infrastructure Protection Plan (NIPP) that obligates each critical infrastructure and key resources sector, such as transportation sector, to develop a sector specific plan.

---

does identify that mass transit and passenger rail operators should develop covert testing exercises. Moreover, the Implementing Recommendations of the 9/11 Commission Act of 2007 requires DHS to develop and implement the National Strategy for Railroad Transportation Security, which is to include prioritized goals, actions, objectives, policies, mechanisms, and schedules for assessing the usefulness of covert testing of railroad security systems.<sup>11</sup> Furthermore, the explanatory statement accompanying Division E of the Consolidated Appropriations Act, 2008 (the DHS Appropriations Act, 2008), directs TSA to be more proactive in red teaming for all modes of transportation.<sup>12</sup> Specifically, the statement directs approximately \$6 million of TSA's appropriated funds for red team activities to identify potential vulnerabilities and weaknesses in airports and air cargo facilities, as well as in transit, rail, and ferry systems.

Prior to the creation of TSA, the Department of Transportation's Federal Aviation Administration (FAA) monitored the performance of airport screeners. FAA created the "red team," as it came to be known, to assess the commercial aviation industry's compliance with FAA security requirements and to test whether U.S. aviation passenger and checked baggage screening systems were able to detect explosives and other threat items. TSA began its covert testing program in September 2002. TSA's covert testing program consists of a nationwide commercial aviation testing program conducted by OI, and a local commercial airport testing program implemented by OSO and FSDs at each airport.

---

## TSA's National Covert Testing Program for Commercial Aviation

OI conducts national covert tests of three aspects of aviation security at a commercial airport: (1) passenger checkpoint; (2) checked baggage; and (3) access controls to secure areas and airport perimeters. OI conducts covert tests by having undercover inspectors attempt to pass threat objects, such as guns, knives, and simulated improvised explosive devices (IED), through passenger screening checkpoints and in checked baggage, and to attempt to access secure areas of the airport undetected. OI officials stated that they derived their covert testing protocols and test scenarios from prior FAA red team protocols, but updated the threat items used and increased the difficulty of the tests. According to OI officials, they also began conducting tests at airports on a more frequent basis than

---

<sup>11</sup>See Pub. L. No. 110-53, § 1511(b), 121 Stat. 266, 426-29. See also 49 U.S.C. § 114(t).

<sup>12</sup> See explanatory statement accompanying Division E of the Consolidated Appropriations Act, 2008, Pub. L. No. 110-161, Div. E, 121 Stat. 1844 (2007), at 1054.

---

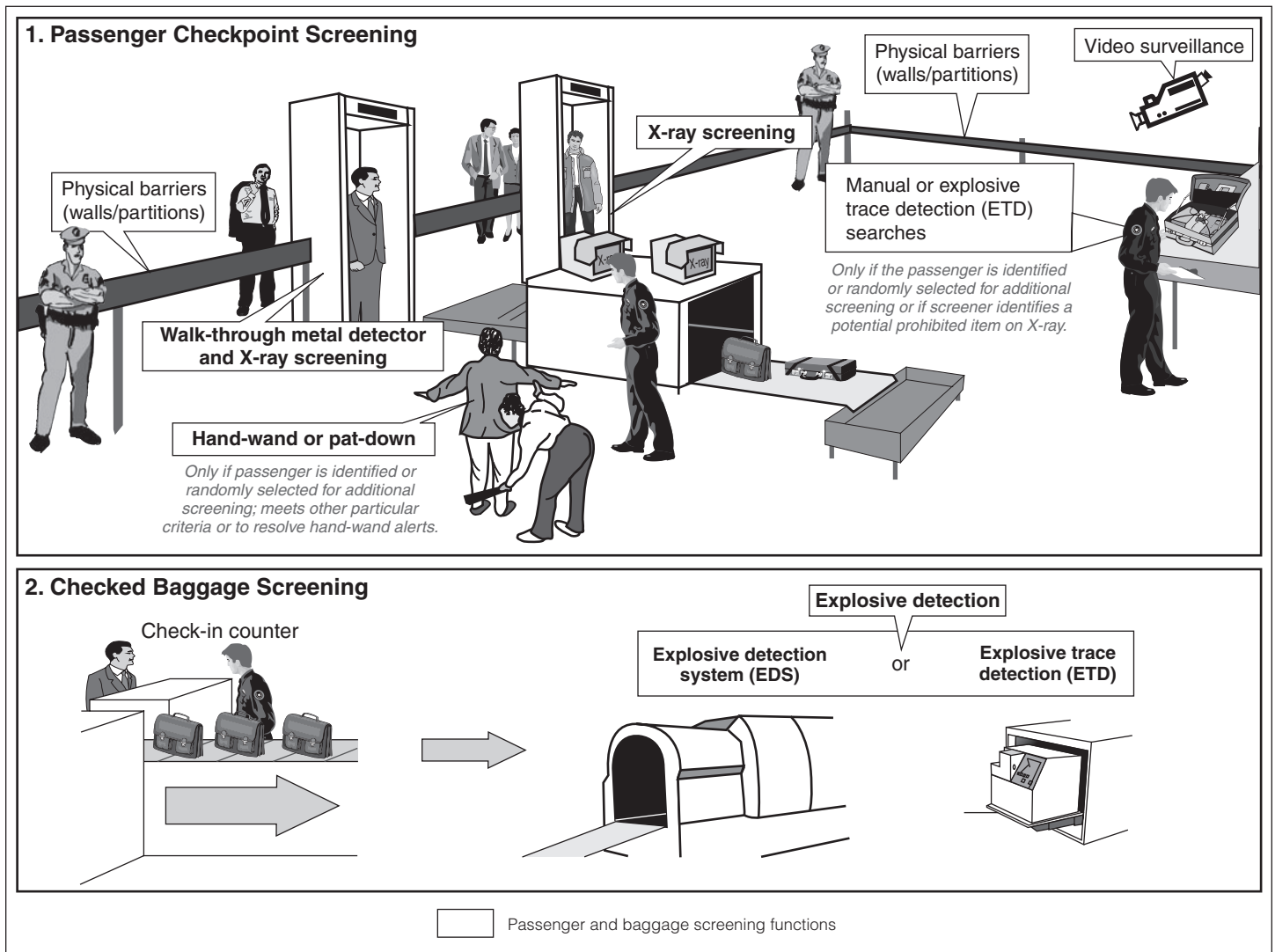
FAA. Initially, OI conducted tests at all of the estimated 450 commercial airports nationwide on a 3-year schedule, with the largest and busiest airports being tested each year.<sup>13</sup> TSA also began using threat information to make tests more closely replicate tactics that may be used by terrorists.

The number of covert tests that OI conducts during testing at a specific airport varies by the size of the airport. The size of the OI testing teams also varies depending upon the size of the airport being tested, the number of tests that OI plans to conduct, and the number of passenger checkpoints and access points to secure areas at a particular airport. OI testing teams consist of a team leader who observes the tests and leads post-test reviews with TSOs, and inspectors who transport threat items through passenger checkpoints and secure airport areas and record test results. Team leaders usually have previous federal law enforcement experience, while inspectors often include program analysts, administrative personnel, and other TSA personnel. Prior to testing, each team leader briefs their team to ensure that everyone understands their role, the type of test to be conducted, and the threat item they will be using. For tests at passenger checkpoints and in checked baggage, OI uses different IED configurations and places these IEDs in various areas of each inspector's body and checked baggage to create different test scenarios. Figure 1 provides an overview of TSA's passenger checkpoint and checked baggage screening operations and equipment.

---

<sup>13</sup>TSA classifies the commercial airports in the United States into one of five categories (X, I, II, III, and IV) based on various factors, such as the total number of takeoffs and landings annually and other special security considerations. In general, Category X airports have the largest number of passenger boardings, and category IV airports have the smallest. TSA periodically reviews airports in each category and, if appropriate, updates airport categorizations to reflect current operations. Until August 2005, OI conducted covert testing at category X airports once per year, category I and II airports once every 2 years, and category III and IV airports at least once every 3 years.

**Figure 1: TSA's Passenger Checkpoint and Checked Baggage Screening Operations and Equipment**



Sources: GAO and Nova Development Corporation.

According to OI officials, on the day of testing, OI typically notifies the airport police about one half hour, and the local FSD 5 minutes, before testing begins and instructs them not to notify the TSOs that testing is being conducted. OI officials stated that they provide this notification for security and safety reasons.



---

## Covert Testing Procedures at Passenger Checkpoints

During passenger checkpoint testing, each team of inspectors carries threat items through the passenger checkpoint. If the TSO identifies the threat item during screening, the inspector identifies him or herself to the TSO and the test is considered a pass. If the TSO does not identify the threat item, the inspector proceeds to the sterile area of the airport and the test is considered a failure. For each test, inspectors record the steps taken by the TSO during the screening process and test results, and the team leader assigns any requirements for remedial training as a consequence of a failed test. The specific types of covert tests conducted by TSA at the passenger checkpoint is sensitive security information and cannot be described in this report.

## Covert Testing Procedures for Checked Baggage

Covert tests of checked baggage are designed to measure the effectiveness of the TSOs' ability to utilize existing checked baggage screening equipment, not to test the effectiveness of the screening equipment. In covert tests of checked baggage screening, an inspector poses as a passenger and checks their baggage containing a simulated threat item at the airline ticket counter. The bag is then screened by TSOs using one of two checked baggage screening methods. At airports that have explosive detection systems (EDS), the TSO uses these machines to screen each bag.<sup>14</sup> At airports that do not have EDS and at airports where certain screening stations do not have EDS, such as curbside check-in stations, the TSOs use an Explosive Trace Detection (ETD) machine to screen checked baggage. During the ETD screening process of both carry-on and checked baggage, TSOs attempt to detect explosives on passengers' baggage by swabbing the target area and submitting the swab into the ETD machine for chemical analysis.<sup>15</sup> If the machine detects an explosive substance, it alarms, and produces a readout indicating the specific type of explosive detected. The TSO is then required to resolve the alarm by performing additional screening steps such as conducting a physical search of the bag or conducting further ETD testing on and X-raying of footwear. When testing EDS and ETD screening procedures, OI uses fully assembled objects such as laptop computers, books, or packages.

Whether using EDS or ETD, if the TSO fails to identify the threat item, the inspectors immediately identify themselves to stop the checked baggage

---

<sup>14</sup>EDS machines use specialized X-rays to detect characteristics of explosives that may be contained in passengers' checked baggage as it moves along a conveyor belt.

<sup>15</sup>ETD machines can detect chemical residues that may indicate the presence of explosives on a passenger or within passengers' baggage.

---

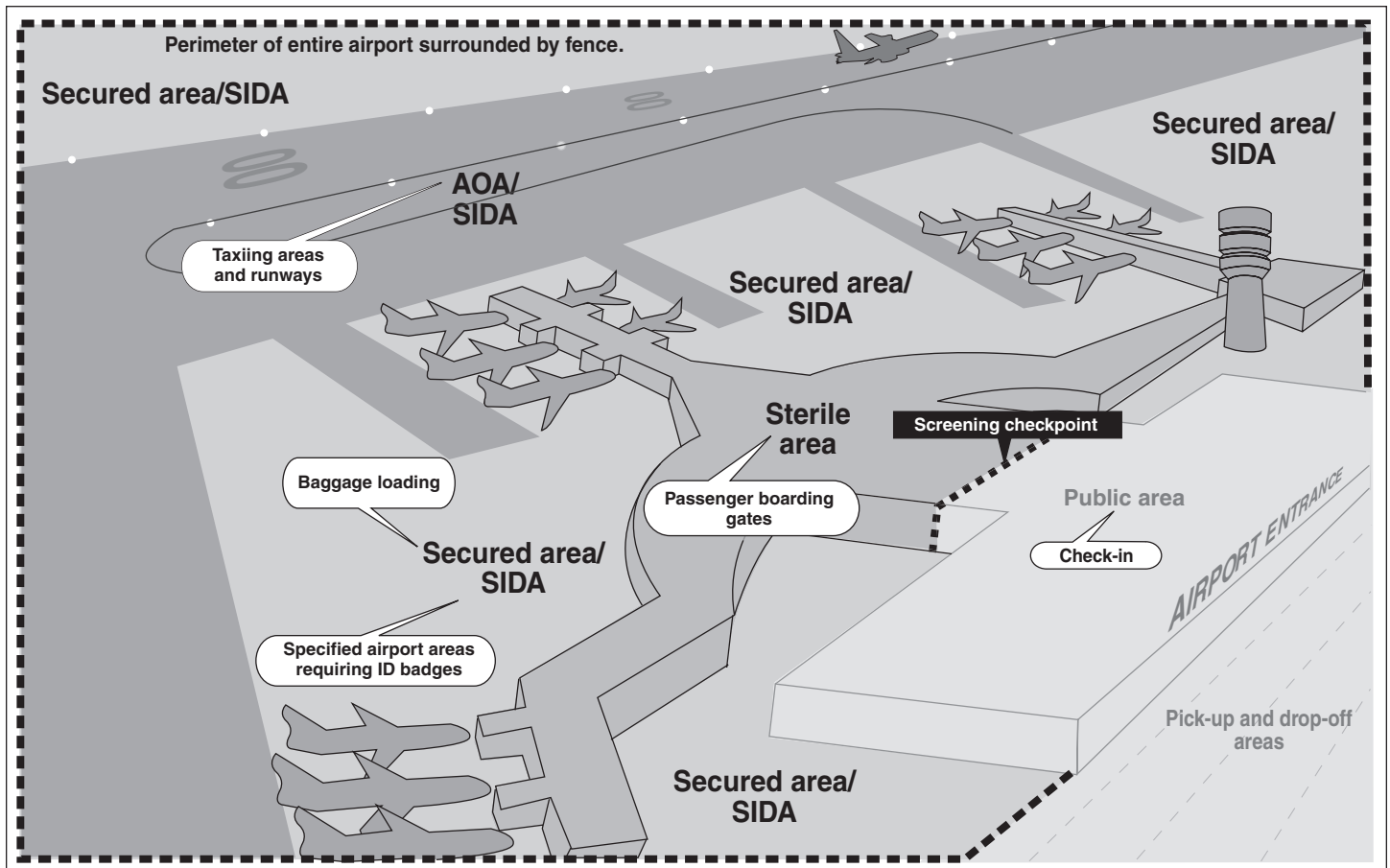
from being sent for loading onto the aircraft, and the test is considered a failure. If the TSO identifies the threat item, the inspectors also identify themselves and the test is considered a pass. If the OI inspector determines that the test failure was due to the screening equipment not working correctly, the test is considered invalid. OI conducts two types of checked baggage covert tests:

- **Opaque object:** This test is designed to determine if a TSO will identify opaque objects on the X-ray screen and conduct a physical search of the checked bag. During these tests, OI inspectors conceal a threat item that cannot be penetrated by the X-ray and appears on the EDS screen as an opaque object among normal travel objects within checked baggage.
- **IED in bag:** This test is designed to determine if a TSO will identify an IED during a search of the bag and use proper ETD procedures to identify it as a threat. During these tests, OI inspectors conceal a simulated IED within checked baggage. In addition, the IED may be contained within other objects inside of the bag.

#### Covert Testing Procedures for Access Controls

OI inspectors conduct covert tests to determine if they can infiltrate secure areas of the airport, such as jet ways or boarding doors to aircraft. Each U.S. commercial airport is divided into different areas with varying levels of security. Secure areas, security identification display areas (SIDA), and air operations areas (AOA) are not to be accessed by passengers, and typically encompass areas near terminal buildings, baggage loading areas, and other areas that are close to parked aircraft and airport facilities, including air traffic control towers and runways used for landing, taking off, or surface maneuvering. Figure 2 is a diagram of the security areas at a typical commercial airport.

Figure 2: Diagram of Security Areas at a Typical Commercial Airport



Source: GAO.

If inspectors are able to access secure areas of the airport or are not challenged by airport or airline employees, then the test is considered a failure. OI conducts four types of covert tests for airport access controls.

- **Access to SIDA:** During these tests, OI inspectors who are not wearing appropriate identification attempt to penetrate the SIDA through access points, such as boarding gates, employee doors, and other entrances leading to secure areas to determine if they are challenged by airport or airline personnel.
- **Access to AOA:** During these tests, OI inspectors who are not wearing appropriate identification attempt to penetrate access points leading from public areas to secured areas of the AOA, including vehicle and

---

pedestrian gates through the perimeter fence, cargo areas, and general aviation facilities that provide a direct path to passenger aircraft in secure areas to determine if they are challenged by airport or airline personnel.

- **Access to Aircraft:** During these tests, OI inspectors who are not wearing appropriate identification or who do not have a valid boarding pass attempt to penetrate access points past the passenger screening checkpoint which lead directly to aircraft, including boarding gates, employee doors, and jet ways to determine if they are challenged by airport or airline personnel.
- **SIDA Challenges:** During these tests, OI inspectors attempt to walk through secure areas of the airport, such as the tarmac and baggage loading areas, without appropriate identification to determine if they are challenged by airport personnel. If not challenged, then the test is considered a failure.

## Post-Test Reviews and Analysis

After testing at the airport is complete, team leaders conduct post-test reviews with the TSOs, supervisors, and screening managers involved in the testing. These post-test reviews include a hands-on demonstration of the threat items used during each test and provide an opportunity for TSOs to ask questions about the test. According to OI officials, the purpose of these post-test reviews is to serve as a training tool for TSOs. Following the post-test review, OI officials meet with the airport FSD to discuss the test results and any vulnerabilities identified at the airport. OI also provides the FSD with the names of each TSO required to undergo remedial training.<sup>16</sup> OI usually completes all aspects of its covert tests at an airport within several days. After completing tests at each airport, OI staff document test results on standardized data collection instruments and meet to discuss the results and identify the actions that they will recommend to TSA management to address the vulnerabilities identified by the tests. The airport testing data collected are then inputted into a database by OI headquarters staff, who develop reports that summarize the tests results and the vulnerabilities identified. These reports are then presented to TSA management, such as the Administrator. OI staff also regularly brief TSA's Administrator and management, such as the Assistant

---

<sup>16</sup>ATSA requires that each TSO who failed a covert test has to undergo remedial training for the function, such as X-ray screening, that he or she failed before returning to that function. The TSO can perform a different function, such as manual or ETD searches, while undergoing the remedial training.

---

Administrator of OSO, on the results of covert tests. Since 2003, when OI completed its first covert testing report, most of OI's reports contained specific recommendations aimed at addressing the vulnerabilities identified during covert testing.

---

## TSA's Local Covert Testing Program for Commercial Aviation

In February 2004, OSO authorized FSDs to conduct their own testing of local passenger and checked baggage screening operations at their airports to serve as a training tool for the TSOs and to measure their performance. Referred to as Screener Training Exercises and Assessments (STEA), FSDs conducted these local covert tests using federal employees, such as TSOs from other local airports and other federal law enforcement officers, and were given discretion to determine the number of tests conducted at their airports, the manner with which the tests were conducted, and the type of tests conducted. OSO considered STEA a tool for training TSOs in detecting threat items, and issued modular bomb kits (MBS II kits) containing simulated IEDs to be used during local testing.<sup>17</sup> During STEA tests, staff placed simulated IEDs in passenger and checked baggage to determine if they would be detected by TSOs.<sup>18</sup> Unlike OI's national covert tests, STEA tests did not include tests of airport access controls. TSOs that failed STEA tests were required to undergo remedial training. In May 2005, we reported that TSA officials stated that they had not yet begun to use data from STEA testing to identify training and performance needs for TSOs because of difficulties in ensuring that local covert testing was implemented consistently nationwide.<sup>19</sup> For example, because FSDs had discretion regarding the number of tests conducted, some airports conducted STEA tests regularly, while others rarely conducted tests. In addition, we previously reported that FSDs had difficulty in finding enough staff to help conduct STEA tests on a consistent basis. OSO officials recognized the limitations of the STEA

---

<sup>17</sup>The MBS II weapons training kits were provided to airports to address the identified training gap by allowing TSOs to see and feel the threat objects they were looking for. According to OSO officials, these kits contained some of the test objects used by OI to conduct covert testing.

<sup>18</sup>STEA included seven types of tests conducted at the passenger checkpoint: IED in property disassembled, IED in property assembled, stimulant on torso, weapon at an angle in property, weapon in property, weapon on individual, and weapon on inner thigh. There is one type of STEA test conducted at the checked baggage screening system—assembled IED in baggage.

<sup>19</sup>GAO, *Aviation Security: Enhancements Made in Passenger and Checked Baggage Screening, but Challenges Remain*, [GAO-06-371T](#) (Washington, D.C.: Apr. 4, 2006).

---

program and, as a result, began to re-structure the program in September 2006. This local covert testing program was renamed the Aviation Screening Assessment Program (ASAP). ASAP is designed to test the performance of passenger and checked baggage screening systems and identify security vulnerabilities at each airport.

In April 2007, OSO began its initial 6-month cycle of ASAP, in which 1,600 tests were conducted in each grouping of airports—Category X (27 airports), category I (55 airports), and Category II through IV (369 airports). OSO compliance inspectors at each airport conduct the tests. Specific test requirements are distributed to FSDs before the start of each 6-month cycle. These test requirements stipulate the percentage of tests to conduct during peak and non-peak passenger screening periods; the percentage of basic, intermediate, or advanced tests to be conducted; and specific types of threat items that should be used during each type of test, such as IEDs or weapons. Following each test, inspectors are to brief the TSOs, supervisors, and screening managers involved in the tests on the results and notify the FSD of the results. With the first cycle of tests initiated in April 2007, TSA officials plan that any recommendations resulting from ASAP tests will be submitted to OSO management and other offices within TSA that need to know the test results. Although the testing requirements, including the level of frequency and types of tests, will not change during the initial 6-month cycle to preserve the validity of the test results, TSA officials plan to analyze the results of the tests and evaluate the need to revise the structure of the tests or the type of threat items used after testing is complete. According to OSO officials, the first cycle of ASAP tests are complete, but the results are still being analyzed by TSA to determine the overall findings from the tests.

---

## Covert Testing as a Key Component of TSA's Broader Risk Management Approach

TSA's national and local aviation covert testing programs contribute to TSA's broader risk management approach for securing the transportation sector by applying principles of risk assessment to identify vulnerabilities in commercial aviation. Risk management is a systematic and analytical process to consider the likelihood that a threat will endanger an asset, individual, or function, and to identify actions to reduce the risk and mitigate the consequences of an attack. Risk management, as applied in the homeland security context, can help federal decision-makers determine where and how to invest limited resources within and among the various modes of transportation. In recent years, the President, through Homeland Security Presidential Directives (HSPD), and laws such as the Intelligence Reform and Terrorism Prevention Act of 2004, have provided that federal agencies with homeland security responsibilities

---

should apply risk-based principles to inform their decision making regarding allocating limited resources and prioritizing security activities.<sup>20</sup> The 9/11 Commission recommended that the U.S. government should identify and evaluate the transportation assets that need to be protected, set risk-based priorities for defending them, select the most practical and cost-effective ways of doing so, and then develop a plan, budget, and funding to implement the effort.<sup>21</sup> In 2002, the President issued The National Strategy for Homeland Security that instructs the federal government to allocate resources in a balanced way to manage risk in our border and transportation security systems while ensuring the expedient flow of goods, services, and people. Further, the Secretary of DHS has made risk-based decision-making a cornerstone of departmental policy. In May 2007, TSA issued the TS-SSP and supporting plans for each mode of transportation that establish a system based risk management approach for securing the transportation sector.

We have previously reported that a risk management approach can help to prioritize and focus the programs designed to combat terrorism.<sup>22</sup> A risk assessment, one component of a risk management approach, consists of three primary elements: a vulnerability assessment, a threat assessment, and a criticality assessment. A vulnerability assessment is a process that identifies weaknesses in physical structures, personnel protection systems, processes, or other areas that may be exploited by terrorists, and may suggest options to eliminate or mitigate those weaknesses. TSA uses both national and local aviation covert testing as a method to identify and mitigate security vulnerabilities in the aviation sector. A threat assessment identifies and evaluates threats based on various factors, including capability and intentions as well as the lethality of an attack. Criticality assessment evaluates and prioritizes assets and functions in terms of

---

<sup>20</sup>See, e.g., 49 U.S.C. § 114(t).

<sup>21</sup>National Commission on Terrorist Attacks upon the United States, the 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks upon the United States (Washington, D.C.: 2004). The 9/11 Commission was an independent, bipartisan commission created in late 2002, to prepare a complete account of the circumstances surrounding the September 11, 2001 terrorist attacks, including preparedness for and the immediate response to the attacks. The Commission was also mandated to provide recommendations designed to guard against future attacks.

<sup>22</sup>GAO, *Passenger Rail Security: Enhanced Federal Leadership Needed to Prioritize and Guide Security Efforts*, [GAO-05-851](#) (Washington, D.C.: Sept. 9, 2005); and GAO, *Aviation Security: Federal Efforts to Secure U.S.-Bound Air Cargo Are in the Early Stages and Could Be Strengthened*, [GAO-07-660](#) (Washington, D.C.: Apr. 30, 2007).

---

specific criteria, such as their importance to public safety and the economy, as a basis for identifying which structures or processes require higher or special protection from attack.

---

## TSA Has a Risk-Based Covert Testing Strategy to Identify Vulnerabilities and Measure the Performance of Selected Aviation Security Systems, but Could Strengthen Its Testing Efforts

TSA has designed and implemented risk-based national and local covert testing programs to achieve its goals of identifying vulnerabilities in and measuring the performance of passenger checkpoint and checked baggage screening systems and airport perimeters and access controls, and has begun to determine the extent to which covert testing will be used to identify vulnerabilities and measure the effectiveness of security practices related to non-aviation modes of transportation. OI used information on terrorist threats to design and implement its national covert tests and determine at which airports to conduct tests based on analyses of risks. However, OI inspectors did not systematically record specific causes for test failures related to TSOs, procedures, or screening equipment that did not work properly. OI also did not systematically collect and analyze information on effective screening practices that may contribute to TSOs' ability to detect threat items. Without systematically recording reasons for test failures, such as failures caused by screening equipment not working properly, as well as reasons for test passes, TSA is limited in its ability to mitigate identified vulnerabilities. TSA recently redesigned its local covert testing program to address limitations in its previous program. The new program, ASAP, should provide TSA with a measure of the performance of passenger and checked baggage screening systems and help to identify security vulnerabilities. Furthermore, TSA has begun to determine the extent to which covert testing will be used to identify vulnerabilities and measure the effectiveness of security practices in non-aviation modes of transportation. While TSA coordinates with domestic and foreign organizations regarding transportation security efforts, they do not have a systematic process in place to coordinate with these organizations regarding covert testing in non-aviation settings, and opportunities for TSA to learn from these organizations' covert testing efforts exist.

---

## TSA Uses a Risk-Based Covert Testing Strategy

OI uses threat assessments and intelligence information to design and implement national covert tests that meet its goals of identifying vulnerabilities in passenger checkpoint and checked baggage screening systems, and airport perimeters and access controls. While OI currently focuses its covert tests on these three areas of aviation security, it has recently begun to establish procedures for the testing of air cargo facilities. According to OI officials, as of March 2008, OI has not yet conducted any tests of air cargo. In designing its covert tests, OI works



---

with DHS's Transportation Security Laboratory to create threat items to be used during covert tests. OI also uses threat information to replicate tactics that may be used by terrorists. The tactics that OI uses are all designed to test the capabilities of passenger checkpoint and checked baggage screening systems to identify where vulnerabilities exist. The process OI uses to select which airports to test has evolved since covert testing began in September 2002 to focus more on those airports determined to be at greater risk of a terrorist attack. Initially, OI's goals were to conduct covert tests at all commercial airports, with tests being conducted more frequently at those airports with the largest number of passenger boardings than smaller airports with fewer flights. In August 2005, when TSA began focusing on the most catastrophic threats, OI changed its testing strategy to utilize a risk-based approach to mitigate those threats.

---

### OI Could Better Identify Vulnerabilities by Recording and Analyzing Specific Causes of Covert Testing Failures and Passes of National Covert Tests in Its Testing Database

OI inspectors record information on the results of national covert tests on data collection instruments after each test is conducted, including the extent to which TSOs properly followed TSA screening procedures and whether the test was passed or failed. After airport testing is complete, OI headquarters analysts input the covert test results into a centralized database. While analysts input whether the test was a pass or a fail and inspectors observations regarding some tests, they do not systematically capture OI's assessment of the cause of the test failure and include that information in the database.<sup>23</sup> Test failures could be caused by (1) TSOs not properly following existing TSA screening procedures, (2) screening procedures that are not clear to TSOs, (3) screening procedures that lack sufficient guidance to enable TSOs to identify threat items, and (4) screening equipment that does not work properly. Moreover, when inspectors determine the cause of a covert test failure to be due to screening equipment, such as the walk through metal detector, the hand-held metal detector, or ETD not alarming in response to a threat item, OI considers these tests to be invalid. While OI officials stated that they report instances when equipment may not be working properly to the airport FSD and officials from the Transportation Security Laboratory, they do not input that equipment caused the failure in the covert testing database. TSA management may find this information useful in identifying

---

<sup>23</sup>While some test entries recorded in the database include observations made by inspectors, OI officials told us that these observations are not intended to identify the reason for a test failure. Moreover, these observations are not consistently recorded in the database to allow OI to analyze trends in test outcomes.

---

vulnerabilities in the aviation system that relate to screening equipment not working properly. OI officials stated that they do not record information on equipment failures because there is always a possibility that the simulated threat item was not designed properly and therefore should not have set off the alarm. Further, they stated that DHS's Transportation Security Laboratory is responsible for ensuring that screening equipment is working properly. However, the Laboratory does not test screening equipment at airports in an operational environment. Furthermore, according to OI officials, identifying a single cause for a test failure may be difficult since covert testing failures can be caused by multiple factors. However, in discussions with OI officials about selected individual test results, inspectors were able in their view, in most of these cases, to identify the cause they believed contributed most to the test failure. According to the Standards for Internal Control in the Federal Government, information should be recorded and communicated to management and others in a form and within a time frame that enables them to carry out their internal control and other responsibilities. The Standards further call for pertinent information to be identified, captured, and distributed in a form and time frame that permits people to perform their duties efficiently. By not systematically inputting the specific causes for test failures in its database, including failures due to equipment, OI may be limiting its ability to identify trends that impact screening performance across the aviation security systems tested.

In addition to not identifying reasons the inspectors believed caused the test failures, OI officials do not systematically record information on screening practices that may contribute to covert test passes. However, OI inspectors occasionally captured information of effective practices used by TSOs to detect threat items during covert tests in the data collection instruments used during these tests. Further, during covert tests that we observed, OI inspectors routinely discussed with us those practices used during certain tests that they viewed as effective, such as effective communication between TSOs and supervisors in identifying threat items. In 2006, OSO officials requested a TSA internal review of differences in checkpoint screening operations at three airports to identify whether the airports employed certain practices that contributed to their ability to detect threat items during covert tests, among other things. Between June and October 2006, OI's Internal Reviews Division (IRD) reviewed passenger checkpoint covert test results for each airport, observed airport operations, interviewed TSA personnel, and reviewed documents and information relevant to checkpoint operations. IRD's review identified a number of key factors that may contribute to an airport's ability to detect threat items. While IRD conducted this one time review of effective

---

screening practices that may have led to higher test pass rates, OI does not systematically collect information on those practices that may lead to test passes. As discussed earlier in this report, Standards for Internal Control in the Federal Government stated the need for pertinent information to be identified and captured to permit managers to perform their duties efficiently. Without collecting information on effective screening practices that, based on the inspectors' views, may lead to test passes, TSA managers are limited in their ability to identify measures that could help to improve screening performance across the aviation security system.

---

### TSA Redesigned Its Local Covert Testing Program to Address Limitations of Its Previous Program and to Measure the Performance of Passenger Checkpoint and Checked Baggage Screening

In April 2007, TSA initiated its local covert testing program, the Aviation Screening Assessment Program (ASAP). TSA is planning to use the results of ASAP as a statistical measure of the performance of passenger checkpoint and checked baggage screening systems, in addition as a tool to identify security vulnerabilities. TSA ASAP guidance applies a standardized methodology for the types and frequency of covert tests to be conducted in order to provide a national statistical sample. If implemented as planned, ASAP should provide TSA with a measure of the performance of passenger and checked baggage screening systems and help identify security vulnerabilities. According to OSO officials, the first cycle of ASAP tests were completed, but the results are still being internally analyzed by TSA to determine the overall findings from the tests. As a result, it is too soon to determine whether ASAP will meet its goals of measuring the performance of passenger and checked baggage screening systems and identifying vulnerabilities.

Similar to OI's national covert testing program, OSO applies elements of risk in designing and implementing ASAP tests. Unlike national covert tests, the ASAP program does not use elements of a risk-based approach to determine the location and frequency of the tests because, according to TSA officials, in order to establish a national baseline against which TSA can measure performance, all airports must be tested consistently and with the same types of tests. OSO officials plan to analyze the results of the tests and evaluate the need to revise the tests or the type of threat items used after the first and second testing cycle and annually thereafter. Furthermore, OSO officials stated that they plan to assess the data, including the types of vulnerabilities identified and the performance of the TSOs in detecting threat items, and develop recommendations for mitigating vulnerabilities and improving screening performance. Officials stated that OSO also plans to conduct follow-up testing to determine whether vulnerabilities that were previously identified have been addressed or if recommendations made were effective.

---

According to TSA's ASAP guidance, individuals conducting the ASAP tests will be required to identify specific causes for all test failures. In addition to identifying test failures attributed to TSOs, such as the TSO not being attentive to their duties or not following TSA screening procedures, individuals conducting ASAP tests are also required to identify and record causes for failures related to TSOs, screening procedures that TSOs said were not clear or lack sufficient detail to enable them to detect threat items, and screening equipment.

OSO officials further stated that they plan to develop performance measures for the ASAP tests after the results of the first 6 month cycle of tests are evaluated. However, officials stated that performance measures for the more difficult category of tests will not be developed because these tests are designed to challenge the aviation security system and the pass rates are expected to be low. Furthermore, TSA officials stated that the results of ASAP tests will not be used to measure the performance of individual TSOs, FSDs, or airports, but rather to measure the performance of the passenger checkpoint and checked baggage screening system. TSA officials stated that there will not be a sufficient number of ASAP tests to measure individual TSO, FSD, or airport performance. We previously reported that TSA had not established performance measures for its national covert testing program and that doing so would enable TSA to focus its improvement efforts on areas determined to be most critical, as 100 percent detection during tests may not be attainable.<sup>24</sup> While TSA has chosen not to establish performance measures for the national covert testing program, as stated above, they plan to develop such measures for only the less difficult ASAP tests.

---

## Covert Testing in Non-Aviation Modes of Transportation

Since the initiation of TSA's covert testing program in 2002, the agency has focused on testing commercial aviation passenger checkpoints, checked baggage, and airport perimeters and access controls. However, TSA is in the early stages of determining the extent to which covert testing will be used to identify vulnerabilities and measure the effectiveness of security practices in non-aviation modes of transportation. In addition, TSA officials stated that it would be difficult to conduct covert tests in non-aviation modes because these modes typically do not have established security screening procedures to test, such as those in place at airports.

---

<sup>24</sup>GAO, *Aviation Security: Screener Training and Performance Measurement Strengthened, but More Work Remains*, [GAO-05-457](#) (Washington, D.C.: May 2, 2005).

---

Specifically, passengers and their baggage are not generally physically screened through metal detectors and X-rays prior to boarding trains or ferries as they are prior to boarding a commercial aircraft, making it difficult to conduct tests. OI officials also stated that they do not currently have the resources necessary to conduct covert tests in both aviation and non-aviation modes of transportation.

Although OI does not regularly conduct covert tests in non-aviation modes of transportation, it has conducted tests during three TSA pilot programs designed to test the feasibility of implementing airport style screening in non-aviation modes of transportation to include mass transit, passenger rail, and maritime ferry facilities. In 2004, TSA conducted a Transit and Rail Inspection pilot program in which passenger and baggage screening procedures were tested on select railways.<sup>25</sup> TSA also tested similar screening procedures at several bus stations during the Bus Explosives Screening Technology pilot in 2005.<sup>26</sup> In addition, TSA has also been testing screening equipment on ferries in the maritime mode through the Secure Automated Inspection Lanes program.<sup>27</sup> According to OI officials, during these three pilot programs, OI conducted covert testing to determine if they could pass threat objects through the piloted passenger screening procedures and equipment. However, these tests were only conducted on a trial basis during these pilot programs. While OI has not developed plans or procedures for testing in non-aviation modes of transportation, the office has begun to explore the types of covert tests that it might conduct if it receives additional resources to test in these modes.

---

<sup>25</sup>The goal of TSA's Transit and Rail Inspection Pilot program was to evaluate the use of existing and emerging technologies in the rail environment to screen passengers' carry-on items, checked baggage, cargo, and parcels for explosives. The pilot was conducted in three phases. Phase I evaluated the use of screening technologies to screen passengers and baggage prior to boarding trains at the New Carrollton, Maryland, train station. Phase II tested screening of checked and unclaimed baggage and cargo prior to loading on board Amtrak trains at Union Station in Washington, D.C. Phase III evaluated the use of screening technologies installed on a rail car to screen passengers and their baggage while the rail car was in transit on a Shoreline East commuter rail car.

<sup>26</sup>The Bus Explosives Screening Technology pilot tested emerging and existing technologies to screen passengers, baggage, and cargo for explosives prior to boarding buses. The pilot was conducted at the Greyhound Bus terminal in Washington, D.C.

<sup>27</sup>TSA's Secure Automated Inspection Lanes pilot program tested portable screening equipment and explosive detection technologies on maritime ferry passengers to identify traces of explosive residue on papers and documents carried by passengers.

---

In addition to OI, TSA's Office of Transportation Sector Network Management (TSNM) may have a role in any covert tests that are conducted in non-aviation modes of transportation. TSNM is responsible for securing the nation's intermodal transportation system and has specific divisions responsible for each mode of transportation—mass transit, maritime, highway and motor carriers, freight rail, pipelines, commercial airports, and commercial airlines. TSNM is also responsible for TSA's efforts to coordinate with operators in all modes of transportation. A TSNM official stated that TSNM has only begun to consider using covert testing in mass transit. In April 2007, TSA coordinated with the Los Angeles County Metropolitan Transportation Authority, Amtrak, and Los Angeles Sheriff's Department during a covert test of the effectiveness of security measures at Los Angeles' Union Station. During the test, several individuals carried threat items, such as simulated IEDs, into the rail system to determine if K-9 patrols, random bag checks, and other random procedures could detect these items. The official from TSNM's mass transit office stated that the agency is incorporating the use of covert testing as a component of the mass transit and passenger rail national exercise program being developed pursuant to the Implementing Recommendations of the 9/11 Commission Act of 2007. However, TSNM has not developed a strategy or plan for how covert testing will be incorporated into these various programs. The TSNM official further stated that he was not aware of other mass transit or passenger rail operators that are currently conducting or planning covert testing of their systems. Furthermore, TSNM does not have a systematic process in place to coordinate with domestic or foreign transportation organizations to learn from their covert testing experiences.

The use of covert or red team testing in non-aviation modes of transportation has been supported in law. The Implementing Recommendations of the 9/11 Commission Act of 2007 directs DHS to develop and implement the National Strategy for Railroad Transportation Security, which is to include prioritized goals, actions, objectives, policies, mechanisms, and schedules for assessing, among other things, the usefulness of covert testing of railroad security systems. Furthermore, the explanatory statement accompanying the Homeland Security Appropriations Act, 2008, directed TSA to be more proactive in red teaming for airports and air cargo facilities, as well as in transit, rail, and ferry systems. Specifically, the statement directed approximately \$6 million of TSA's appropriated amount for red team activities to identify vulnerabilities in airports and air cargo facilities, as well as in transit, rail, and ferry systems. Regarding covert testing of non-aviation modes of transportation, the report of the House of Representatives Appropriations

---

Committee, which accompanies its fiscal year 2008 proposal for DHS appropriations, directed TSA to randomly conduct red team operations at rail, transit, bus, and ferry facilities that receive federal grant funds to ensure that vulnerabilities are identified and corrected.<sup>28</sup>

DHS has also identified covert, or red team, testing as a priority for the Department. The President's July 2002 National Strategy for Homeland Security identified that DHS, working with the intelligence community, should use red team or covert tactics to help identify security vulnerabilities in the nation's critical infrastructure, which includes the transportation sector. The strategy further identifies that red team techniques will help decision makers view vulnerabilities from the terrorists' perspective and help to develop security measures to address these security gaps. In addition, TSA's May 2007 TS-SSP identified that transit agencies should develop meaningful exercises, including covert testing, that test the effectiveness of their response capabilities and coordination with first responders. However, the TS-SSP does not provide any details on the type of covert testing that transit agencies should conduct and does not identify that TSA itself should conduct covert testing in non-aviation modes of transportation.

---

### Select Domestic and Foreign Transportation Organizations and DHS Component Agencies Use Covert Testing to Identify Vulnerabilities and Measure System Effectiveness

Domestic and foreign transportation organizations and DHS component agencies that we interviewed conduct covert testing to identify and mitigate vulnerabilities in non-aviation settings that lack the standardized passenger screening procedures found in the commercial aviation sector and measure the effectiveness of security measures. Our previous work on passenger rail security identified foreign rail systems that use such covert testing to keep employees alert about their security responsibilities. One of these foreign organizations—the United Kingdom Department for Transport's Transport Security and Contingencies Directorate (TRANSEC)—conducts covert testing of passenger rail and seaports in addition to aviation facilities to identify vulnerabilities related to people, security processes, and technologies. According to a TRANSEC official, TRANSEC's non-aviation covert testing includes testing of the nation's passenger rail system and the United Kingdom's side of the channel tunnel between the United Kingdom and France. TRANSEC conducts a number of covert tests to determine whether employees are

---

<sup>28</sup>H.R. Rpt. No. 110-181, at 62-63 (2007), accompanying H.R. 2638, 110th Cong. (as passed by House of Representatives. June 15, 2007).

---

following security procedures established by TRANSEC or the rail operator, whether processes in place assist employees in identifying threat items, and whether screening equipment works properly. A TRANSEC official responsible for the agency's covert testing program stated that these tests are carried out on a regular basis and are beneficial because, as well as providing objective data on the effectiveness of people and processes, they encourage staff to be vigilant with respect to security.

In our September 2005 report on passenger rail security, we recommended that TSA evaluate the potential benefits and applicability—as risk analyses warrant and as opportunities permit—of implementing covert testing processes to evaluate the effectiveness of rail system security personnel.<sup>29</sup> Like TRANSEC in the United Kingdom, TSA has existing security directives that must be followed by passenger rail operators that could be tested. TSA generally agreed with this recommendation. In responding to the recommendation, TSA officials stated that the agency regularly interacts and communicates with its security counterparts in foreign countries to share best practices regarding passenger rail and transit security and will continue to do so in the future. TSA officials further stated that the agency has representatives stationed overseas at U.S. embassies that are knowledgeable about security issues across all modes of transportation. While TSA coordinates with domestic and foreign organizations regarding transportation security efforts, they do not have a systematic process in place to coordinate with these organizations regarding covert testing in non-aviation modes of transportation, and opportunities for TSA to learn from these organizations' covert testing efforts exist.

In the United States, Amtrak has conducted covert tests to identify and mitigate vulnerabilities in their passenger rail system. Amtrak's Office of Inspector General has conducted covert tests of intercity passenger rail systems to identify vulnerabilities in the system related to security personnel and Amtrak infrastructure. The results from these tests were used to develop security priorities that are currently being implemented by Amtrak. According to an Amtrak official, as the security posture of the organization matures, the covert testing program will shift from identifying vulnerabilities to assessing the performance of existing rail security measures.

---

<sup>29</sup>See [GAO-05-851](#).



---

Transportation industry associations with whom we spoke, who represented various non-aviation modes of transportation, supported the use of covert testing as a means to identify security vulnerabilities and to test existing security measures. Officials from the American Association of Railroads (AAR), which represents U.S. passenger and freight railroads, and the American Public Transportation Association (APTA), which represents the U.S. transit industry, stated that covert testing in the passenger rail and transit industries would help to identify and mitigate security vulnerabilities and increase employee awareness of established security procedures. AAR and APTA officials stated that covert testing might include placing bags and unattended items throughout a rail station or system to see if employees or law enforcement personnel respond appropriately and in accordance with security procedures. AAR and APTA officials further stated that any testing conducted by TSA would require close coordination with rail operators to determine what should be tested, the testing procedures to be used, and the practicality of such testing.

Within DHS, the U.S. Customs and Border Protection (CBP) also conducts covert testing at land, sea, and air ports of entry in the United States to test and evaluate CBP's capabilities to detect and prevent terrorists and illicit radioactive material from entering the United States. According to CBP officials, the purpose of CBP's covert testing program is to identify potential technological vulnerabilities and procedural weaknesses related to the screening and detection of passengers and containers entering the United States with illicit radioactive material, and to assess CBP officers' ability to identify potential threats. As of June 2008, CBP tested and evaluated two land border crossings on their capabilities to detect and prevent terrorists and illicit radioactive material from entering the United States. In addition, CBP covertly and overtly evaluated the nation's 22 busiest seaports for radiation detection and the effectiveness of the non-intrusive imaging radiation equipment deployed at the seaports. CBP officials also stated that the agency is planning to expand testing to address overseas ports that process cargo bound for the United States.

In addition to CBP, the DHS Domestic Nuclear Detection Office (DNDO) conducts red team testing to measure the performance of and identify vulnerabilities in equipment and procedures used to detect nuclear and radiological threats in the United States and around the world. According to DNDO officials, the agency uses the results of red team tests to help mitigate security vulnerabilities, such as identifying nuclear detection equipment that is not working correctly. DNDO also uses red team testing to determine if unclassified information exists in open sources, such as on the internet, which could potentially be used by terrorists to exploit

---

vulnerabilities in nuclear detections systems. DNDO's program, according to its officials, provides a means to assess vulnerabilities that an adversary is likely to exploit, and to make recommendations to either implement or improve security procedures.

---

## TSA Could More Fully Use the Results of Covert Tests to Mitigate Security Vulnerabilities Identified in the Commercial Aviation System

TSA's national aviation covert testing program has identified vulnerabilities in select aspects of the commercial aviation security system at airports of all sizes; however, the agency is not fully using the results of these tests to mitigate identified vulnerabilities. The specific results of these tests are classified and are presented in our classified May 2008 report. Covert test failures can be caused by various factors, including TSOs not properly following TSA procedures when screening passengers, screening equipment that does not detect a threat item, or TSA screening procedures that do not provide sufficient detail to enable TSOs to identify the threat item. Senior TSA officials, including TSA's Administrator, are routinely briefed on the results of covert tests and provided with OI reports that describe the vulnerabilities identified by these tests and recommendations to correct identified vulnerabilities. However, OSO lacks a systematic process to ensure that OI's recommendations are considered, and does not systematically document its rationale for why it did or did not implement OI's recommendations. OSO and OI also do not have a process in place to assess whether the corrective action implemented mitigated the identified vulnerabilities through follow-up national or local covert tests, and if covert test results improved. According to OSO officials, TSA has other methods in place to identify whether corrective actions or other changes to the system are effective; however, officials did not provide specific information regarding these methods. Moreover, in those cases where OSO took no action to address OI's recommendation, they did not systematically document their rationale for why they took no action. In the absence of a systematic process for considering OI's recommendations, documenting their decision-making process, and evaluating whether corrective actions mitigated identified vulnerabilities, TSA is limited in its ability to use covert testing results to improve the security of the commercial aviation system. OSO senior leadership stated that opportunities exist to improve the agency's processes in this area.

---

## TSA Covert Test Results Identified Vulnerabilities in the Aviation Security System

Between September 2002 and June 2007, OI conducted more than 20,000 covert tests of passenger checkpoints, checked baggage screening systems, and airport perimeters and access control points collectively at every commercial airport in the United States regulated by TSA. The results of these tests identified vulnerabilities in select aspects of the commercial aviation security system at airports of all sizes. While the specific results of these tests and the vulnerabilities they identified are classified, covert test failures can be caused by multiple factors, including TSOs not properly following TSA procedures when screening passengers, screening equipment that does not detect a threat item, or TSA screening procedures that do not provide sufficient detail to enable TSOs to identify the threat item. TSA cannot generalize covert test results either to the airports where the tests were conducted or to airports nationwide because the tests were not conducted using the principles of probability sampling.<sup>30</sup> For example, TSA did not randomly select times at which tests were conducted, nor did they randomly select passenger screening checkpoints within the airports. Therefore, each airport's test results represent a snapshot of the effectiveness of passenger checkpoint screening, checked baggage screening, and airport access control systems, and should not be considered a measurement of any one airport's performance or any individual TSO's performance in detecting threat objects. Although the results of the covert tests cannot be generalized to all airports, they can be used to identify vulnerabilities in the aviation security system. TSA officials stated that they do not want airports to achieve a 100 percent pass rate during covert tests because they believe that high pass rates would indicate that covert tests were too easy and therefore were not an effective tool to identify vulnerabilities in the system.

---

<sup>30</sup>A well-designed probability sample would enable failure rates to be generalized to the airports in which the tests were conducted and to all airports. In a probability sample, each item in the population being studied has a known, non-zero probability of being selected.

---

**TSA Lacks a Systematic Process to Ensure that Covert Testing Recommendations Are Considered and Actions Are Taken to Address Them If Determined Necessary**

After completing its covert tests, OI provides written reports and briefings on the test results to senior TSA management, including TSA's Administrator, Assistant Administrator of OSO, and area FSDs. In these reports and briefings, OI officials provide TSA management with the results of covert tests, describe the security vulnerabilities identified during the tests, and present recommendations to OSO that OI believes will mitigate the identified vulnerabilities. TSA's Administrator and senior OSO officials stated that they consider the aviation security system vulnerabilities that OI presents in its reports and briefings as well as the recommendations made. However, OSO officials we spoke with stated that they do not have a systematic process in place to ensure that all of OI's recommendations are considered or to document their rationale for implementing or not implementing these recommendations.<sup>31</sup> Furthermore, TSA does not have a process in place to assess whether corrective actions taken in response to OI's recommendations have mitigated identified vulnerabilities. Specifically, in those cases where corrective actions were taken to address OI's recommendation, neither OSO nor OI conducted follow-up national or local covert tests to determine if the actions taken were effective. For example, in cases where OI determined that additional TSO training was needed and OSO implemented such training, OSO or OI did not conduct follow-up national or local covert testing to determine if the additional training that was implemented to address the recommendation helped to mitigate the identified vulnerability. According to OSO officials, TSA has other methods in place to identify whether corrective actions or other changes are effective; however, officials did not provide specific information regarding these methods.

Standards for Internal Control in the Federal Government require that internal controls be designed to ensure that ongoing monitoring occurs during the course of normal operations. Specifically, internal controls direct managers to (1) promptly evaluate and resolve findings from audits and other reviews, including those showing deficiencies and recommendations reported by auditors and others who evaluate agencies' operations, (2) determine proper actions in response to findings and recommendations from audits and reviews, and (3) complete, within established time frames, all actions that correct or otherwise resolve the matters brought to management's attention. The standards further identify

---

<sup>31</sup>According to TSA officials, recommendations made as a result of ASAP tests are provided in a report to senior TSA leadership, who make the decision whether or not to implement the recommendation, and the status of each recommendation is tracked in the ASAP database.

---

that the resolution process begins when audit or other review results are reported to management, and is completed only after action has been taken that (1) corrects identified deficiencies, (2) produces improvements, or (3) demonstrates the findings and recommendations do not warrant management action. In the absence of a systematic process for considering and resolving the findings and recommendations from OI's covert tests and ensuring that the effectiveness of actions taken to address these recommendations are evaluated, TSA management is limited in its ability to mitigate identified vulnerabilities to strengthen the aviation security system.

---

### OI Made 43 Recommendations to OSO to Mitigate Vulnerabilities Identified by Covert Tests from March 2003 to June 2007

While neither OSO nor OI have a systematic process for tracking the status of OI covert testing recommendations, at our request, OSO officials provided information indicating what actions, if any, were taken to address OI's recommendations.<sup>32</sup> From March 2003 to June 2007, OI made 43 recommendations to OSO designed to mitigate vulnerabilities identified by national covert tests. To date, OSO has taken actions to implement 25 of these recommendations. For the remaining 18 of OI's 43 recommendations, OSO either took no action to address the recommendation, or it is unclear how the action they took addressed the recommendation. OI did not make any recommendations to OSO related to screening equipment. The specific vulnerabilities identified by OI during covert tests and the specific recommendations made, as well as corrective actions taken by OSO, are classified.

---

## Conclusions

TSA has developed a risk-based covert testing strategy to identify vulnerabilities and measure the performance of select aspects of the aviation security system. OI's national covert testing program is designed and implemented using elements of a risk-based approach, including using information on terrorist threats to design simulated threat items and tactics. However, this program could be strengthened by ensuring that all of the information from the tests conducted is used to help identify and mitigate security vulnerabilities. For example, without a process for recording and analyzing the specific causes of all national covert test failures, including TSOs not properly following TSA's existing screening procedures, procedures that are unclear to TSOs, or screening equipment

---

<sup>32</sup>OSO officials told us that they did not systematically monitor the status of OI's recommendations prior to our request.

---

that is not working properly, TSA is limited in its ability to identify specific areas for improvement, such as screening equipment that may be in need of repair or is not working correctly. Moreover, without collecting and analyzing information on effective practices used at airports that performed particularly well on national covert tests, TSA may be missing opportunities to improve TSO performance across the commercial aviation security system. TSA has only recently begun to determine the extent to which covert testing may be used to identify vulnerabilities and measure the effectiveness of security practices in non-aviation modes of transportation if it receives additional resources to test in these modes. Nevertheless, several transportation industry stakeholders can provide useful information on how they currently conduct covert tests in non-aviation settings, and systematically coordinating with these organizations could prove useful for TSA.

National aviation covert tests have identified vulnerabilities in the commercial aviation security system. However, TSA could better use the covert testing program to mitigate these vulnerabilities by promptly evaluating and responding to OI's findings and recommendations. We recognize that TSA must balance a number of competing interests when considering whether to make changes to TSO training, screening procedures, and screening equipment within the commercial aviation security system, including cost and customer service, in addition to security concerns. We further recognize that, in some cases, it may not be feasible or appropriate to implement all of OI's recommendations. However, without a systematic process in place to consider OI's recommendations, evaluate whether corrective action is needed to mitigate identified vulnerabilities, and evaluate whether the corrective action effectively addressed the vulnerability, OSO is limited in the extent to which it can use the results of covert tests to improve the security of the commercial aviation system.

---

## Recommendations for Executive Action

To help ensure that the results of covert tests are more fully used to mitigate vulnerabilities identified in the transportation security system, we recommended in our May 2008 classified report that the Assistant Secretary of Homeland Security for TSA take the following five actions:

- Require OI inspectors to document the specific causes of all national covert testing failures—including documenting failures related to TSOs, screening procedures, and equipment—in the covert testing database to help TSA better identify areas for improvement, such as additional TSO training or revisions to screening procedures.

- 
- Develop a process for collecting, analyzing, and disseminating information on practices in place at those airports that perform well during national and local covert tests in order to assist TSA managers in improving the effectiveness of checkpoint screening operations.
  - As TSA explores the use of covert testing in non-aviation modes of transportation, develop a process to systematically coordinate with domestic and foreign transportation organizations that already conduct these tests to learn from their experiences.
  - Develop a systematic process to ensure that OSO considers all recommendations made by OI in a timely manner as a result of covert tests, and document its rationale for either taking or not taking action to address these recommendations.
  - Require OSO to develop a process for evaluating whether the action taken to implement OI's recommendations mitigated the vulnerability identified during covert tests, such as using follow-up national or local covert tests to determine if these actions were effective.

---

## Agency Comments and Our Evaluation

We provided a draft of this report to DHS for review and comment. On April 24, 2008, we received written comments on the draft report, which are reproduced in full in appendix II. DHS and TSA concurred with the findings and recommendations, and stated that the report will be useful in strengthening TSA's covert testing programs. In addition, TSA provided technical comments, which we incorporated as appropriate.

Regarding our recommendation that OI document the specific causes of all national covert testing failures related to TSOs, screening procedures, and equipment in the covert testing database, DHS stated that TSA's Office of Inspection (OI) plans to expand the covert testing database to all causes of test failures. DHS further stated that the specific causes of all OI covert testing failures are documented in data collection instruments used during covert tests and within a comment field in the covert testing database when the cause can be determined. However, TSA acknowledged that covert test failures caused by screening equipment not working properly are not recorded in the database in a systematic manner. Documenting test failures caused by equipment should help OI better analyze the specific causes of all national covert testing failures and assist TSA management in identifying corrective actions to mitigate identified vulnerabilities.

---

Concerning our recommendation that OI develop a process for collecting, analyzing, and disseminating information on practices in place at those airports that perform well during national and local covert tests in order to assist TSA managers in improving the effectiveness of checkpoint screening operations, DHS stated that it recognizes the value in identifying factors that may lead to improved screening performance. TSA officials stated that, while OI or ASAP test results can be used to establish a national baseline for screening performance at individual airports, the results are not statistically significant. As a result, additional assessments would be required to provide a statistical measure for individual airports. According to DHS, OI plans to develop a more formal process for collecting and analyzing test results to identify best practices that may lead to test passes. Officials stated that when specific screening practices indicate a positive effect on screening performance, TSA plans to share and institutionalize best practices in the form of management advisories to appropriate TSA managers. Developing a more formal process for collecting and analyzing test results to identify best practices that may lead to test passes should address the intent of this recommendation.

In response to our recommendation that TSA develop a process to systematically coordinate with domestic and foreign transportation organizations as the agency explores the use of covert testing in non-aviation modes of transportation to learn from their experiences, DHS stated that it is taking a number of actions. Specifically, according to DHS, TSNM has worked closely with transit agencies and internal TSA covert testing experts during red team testing exercises and is currently exploring programs in which covert testing may be used to evaluate the effectiveness of security measures. For example, TSNM is considering incorporating covert testing as a part of its Intermodal Security Training and Exercise Program. While considering the use of covert testing in its programs should help TSA evaluate the effectiveness of security measures, it is also important that TSA establish a systematic process for coordinating with domestic and foreign organizations that already conduct testing in non-aviation modes of transportation to learn from their experiences.

DHS further stated that it plans to take action to address our recommendation that the agency develop a systematic process to ensure that OSO considers all recommendations made by OI as a result of covert tests in a timely manner, and documents its rationale for either taking or not taking action to address these recommendations. Specifically, DHS stated that OSO is coordinating with OI to develop a directive requiring that OI's covert testing recommendations be formally reviewed and



---

approved by TSA management, and OSO is establishing a database to track all OI recommendations and determine what action, if any, has been taken to address the recommendation. Taking these steps should address the intent of this recommendation and help TSA to more systematically record whether OI's covert testing recommendations have been addressed.

Concerning our recommendation that OSO develop a process to evaluate whether the action taken to implement OI's recommendations mitigated the vulnerability identified during covert tests, such as using follow-up national or local covert tests or information collected through other methods, to determine if these actions were effective, DHS stated that OSO established a new program to study various aspects of TSO and screening performance in 2007 that considers recommendations originating from OI national covert tests and ASAP tests. According to DHS, after completing each study, recommendations resulting from this analysis will be provided to TSA leadership for consideration. DHS further stated that the results of ASAP tests will also likely be a focus of these future studies. While these actions should help to address the intent of this recommendation, it is also important that OSO assess whether the actions taken to mitigate the vulnerabilities identified by OI's national covert tests are effective.


---

As agreed with your office, unless you publicly announce its contents earlier, we plan no further distribution of this report until 30 days after its issue date. At that time, we will send copies of this report to the Secretary of Homeland Security, Assistant Secretary of DHS for the Transportation Security Administration, and the Ranking Member of the Committee on Homeland Security, House of Representatives, and other interested congressional committees as appropriate. We will also make this report available at no charge on GAO's Web site at <http://www.gao.gov>.

---

If you or your staff have any questions about this report, please contact me at (202) 512-3404 or at [berrickc@gao.gov](mailto:berrickc@gao.gov). Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. Other key contributors to this report were John Hansen, Assistant Director; Chris Currie; Yanina Golburt; Samantha Goodman; Art James; Wendy Johnson; Thomas Lombardi; and Linda Miller.

Sincerely yours,

A handwritten signature in black ink that reads "Cathleen A. Berrick". The signature is written in a cursive style with a long, sweeping tail on the final letter.

Cathleen A. Berrick  
Director, Homeland Security and Justice Issues

---

# Appendix I: Objectives, Scope, and Methodology

---

This report addresses the following questions: (1) what is the Transportation Security Administration's (TSA) strategy for conducting covert testing of the transportation system, and to what extent has the agency designed and implemented its covert tests to achieve identified goals? and (2) what have been the results of TSA's national aviation covert tests conducted from September 2002 to June 2007, and to what extent does TSA use the results of these tests to mitigate security vulnerabilities in the commercial aviation system?

To identify TSA's strategy for conducting covert testing of the transportation system and the extent to which the agency has designed and implemented its covert tests to achieve identified goals, we reviewed applicable laws, regulations, policies, and procedures to determine the requirements for conducting covert testing in the transportation sector. To assess TSA's strategy specifically in the aviation covert testing program, we interviewed TSA Office of Inspection (OI) officials responsible for conducting national covert tests and Office of Security Operations (OSO) officials responsible for local covert tests regarding the extent to which information on risks is included in the design and implementation of tests. We also interviewed the Transportation Security Officers (TSO), supervisors, screening managers, and Federal Security Directors (FSD) who participated in covert tests at each airport where we observed tests to discuss their experience with the national and local covert testing programs. We observed OI inspectors during covert tests at seven airports including airports with heavy passenger traffic and those with just a few flights per day, as well as airports with both federal and contract TSOs. During these observations, we accompanied OI inspectors during all phases of the covert test including planning and observations, testing, and post test reviews with TSOs, supervisors, and screening managers. While these seven airports represent reasonable variations in size and geographic locations, our observations of OI's covert tests and the perspectives provided by TSA officials at these airports cannot be generalized across all commercial airports. However, our observations at the seven airports provided us an overall understanding of how OI conduct covert tests and useful insights provided by TSOs, their supervisors, and FSDs at these airports. We analyzed TSA documents including established protocols for national and local covert testing, procedures for screening passengers and checked baggage, and OI covert testing reports issued from 2002 to 2007 to identify procedures for designing and implementing TSA's covert testing program. Furthermore, to determine the extent to which TSA met the goals of the program, we conducted a detailed analysis of the data collection instrument and methods that OI used to collect covert testing data for the seven airports where we observed covert tests.

We also assessed the adequacy of TSA's internal controls for collecting and maintaining the results of covert tests by evaluating TSA's processes for collecting covert testing data and inputting this data into its database. In assessing the adequacy of internal controls, we used the criteria in GAO's Standards for Internal Control in the Federal Government, GAO/AIMD 00-21.3.1, dated November 1999. These standards, issued pursuant to the requirements of the Federal Managers' Financial Integrity Act of 1982 (FMFIA), provide the overall framework for establishing and maintaining internal control in the federal government. Also pursuant to FMFIA, the Office of Management and Budget issued Circular A-123, revised December 21, 2004, to provide the specific requirements for assessing the reporting on internal controls. To assess TSA's strategy for conducting covert tests in non-aviation modes of transportation, we interviewed officials from TSA's Office of Transportation Sector Network Management (TSNM) regarding the extent to which TSA has conducted covert testing in non-aviation modes of transportation, the applicability and potential use of covert testing in other modes, and their future plans for conducting covert testing in other modes. To understand how other organizations and federal agencies have used covert testing in the non-aviation arena, we interviewed officials from selected federal agencies and organizations that conduct covert testing including Amtrak, the United Kingdom Department for Transport Security (TRANSEC), U.S. Customs and Border Protection (CBP), DHS Domestic Nuclear Detection Office (DNDO), and select transportation industry associations. We reviewed the president's National Strategy for Homeland Security and TSA's Transportation Systems Sector Specific plan, including individual plans for each mode of transportation, to determine the role and use of covert testing across the transportation system. We also reviewed the fiscal year 2008 DHS appropriations legislation, enacted as Division E of the Consolidated Appropriations Act, 2008, and associated committee reports and statements to identify any funding allocated to TSA to conduct covert testing in non-aviation modes.

To determine the results of TSA's national covert tests and the extent to which TSA used the results of these tests to mitigate security vulnerabilities in the aviation system, we obtained and analyzed a database of the results of TSA's national covert tests conducted from September 2002 to June 2007. We analyzed the test data according to airport category, threat item, and type of test conducted between September 2002 and June 2007. We also examined trends in pass and failure rates when required screening steps were or were not followed and examined differences in covert test results between private and federal airports. We assessed the reliability of TSA's covert testing data by reviewing existing information

about the data and the systems used to produce them, and by interviewing agency officials responsible for maintaining the database. We determined that the data were sufficiently reliable for our analysis and the purposes of this report. TSA provided us with a copy of their covert testing database which contained a table with one record, or entry, per test for all of the tests conducted between 2002 and 2007. In order to accurately interpret the data, we reviewed information provided by OI officials regarding each of the fields recorded in the database and information about how they enter test results into the database. We also conducted manual testing of the data, conducting searches for missing data and outliers. To further assess the reliability of the data, we reviewed the source documents used to initially collect the data as well as OI's published reports. We also interviewed OI officials regarding how the results of covert tests are used in developing their recommendations to TSA management. We reviewed OI reports on the results of covert tests issued between March 2003 and June 2007 that were submitted to TSA's Administrator and OSO to identify OI's recommendations for mitigating the vulnerabilities identified during covert tests. We obtained and analyzed a summary of the actions that OSO had taken to address OI's recommendations for mitigating vulnerabilities made from March 2003 to June 2007. We also asked officials to discuss the extent to which OSO has addressed and implemented recommendations made by OI based on covert test results, and we analyzed information provided by TSA regarding the status of each covert testing recommendation made by OI from 2003 to 2007.

We conducted this performance audit from October 2006 to May 2008, in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

# Appendix II: Comments from the Department of Homeland Security

U.S. Department of Homeland Security  
Washington, DC 20528



Homeland  
Security

April 24, 2008

Ms. Cathleen A. Berrick  
Director, Homeland Security and Justice Issues  
Government Accountability Office  
441 G Street, NW  
Washington, DC 20548

Dear Ms. Berrick:

Thank you for the opportunity to comment on GAO-08-958 draft report entitled, *TSA Has Developed a Risk-Based Covert Testing Program, but Could Better Mitigate Aviation Security Vulnerabilities Identified Through Covert Testing*. The findings in the report will be useful in strengthening our testing programs. This letter replicates the Department of Homeland Security's response to the classified version of this report.

As you are well aware through your previous work with us on this topic, covert testing results are only one of many inputs into our security evaluation strategy. Given the nonlinear nature of the risks to transportation security, and our strategy to manage them, the Transportation Security Administration (TSA) relies upon a wide variety of input, including intelligence and workforce feedback, to evaluate and respond to security vulnerabilities.

As your report demonstrates, TSA's Office of Inspection (OI) has a very robust testing program designed to identify systemic vulnerabilities in transportation security systems. Through the OI program, subject matter experts develop and test specific hypotheses regarding potential system vulnerabilities. These tests are not designed to be performance measures. Rather, they are evaluations of system vulnerabilities that can be used to design countermeasures. When viewed in this light, the qualitative results from these experiments are highly valuable in analyzing vulnerabilities, with the conclusions from these experiments informing decisions at the strategic level.

TSA has also recognized the need for a more systematic framework to accurately assess the effectiveness of our screening process and identify areas for improvement. In April 2007, TSA therefore established the Aviation Screening Assessment Program (ASAP) to greatly expand our internal covert testing, provide statistically sound data to support operational decisions, and create a framework to systematically review and report the data. As a result, ASAP runs thousands of covert tests at hundreds of airports nationwide every six months. The information produced from these tests provides data necessary to make more informed decisions on improving the screening process. ASAP analysis also focuses on particular areas of screening for

[www.dhs.gov](http://www.dhs.gov)

targeted improvement including: operations; procedures; technology; training; and management. Through ASAP, we have established a formal process to thoroughly assess the screening process and implement appropriate courses of action addressing concerns revealed through expansive covert testing. The increased pace of ASAP's testing and review cycles also permits TSA to evaluate the effectiveness of actions taken in response to both ASAP and OI recommendations.

While TSA has utilized the results of OI covert testing to make adjustments to the screening process, we do not rely solely on them to make screening process changes. Rather, it is one of many data points used in evaluating system vulnerabilities. Nevertheless, we understand the need for accurate data collection, reporting, and follow-through from all of our testing programs, and will continue program improvements in those areas as recommended.

With respect to testing in other modes, TSA values the relationships that have been fostered with industry stakeholders in non-aviation modes. These relationships have proven valuable over the years in establishing necessary protocols for national security. We will continue to work with our industry partners to obtain lessons learned and best practices to enable the development of testing protocols in various modes of transportation to improve security.

The following are TSA's responses to the specific recommendations:

**Recommendation 1: Require OI to document the specific causes of all national covert testing failures—including documenting failures related to TSOs, screening procedures, and equipment—in the covert testing database to help TSA better identify areas for improvement.**

**Concur:** Specific causes of all Office of Inspection (OI) covert testing failures are documented (on testing documents and in the covert testing database as a comment field) when that cause can be determined. To date, specific causes of equipment failure have not been recorded in the database in a uniform manner. OI will expand the covert testing database to document test failures related to screening equipment.

**Recommendation 2: Develop a process for collecting, analyzing, and disseminating information on practices in place at those airports that perform well during national and local covert tests in order to assist TSA managers in improving the effectiveness of checkpoint screening operations.**

**Concur:** TSA recognizes the value in identifying the elements that lead to improved performance at the checkpoint. Through the Aviation Screening Assessment Program (ASAP) TSA has already begun to establish a national baseline, by detection point, for screening performance. Using OI covert testing data or ASAP data alone to draw conclusions as to which airports are high- and low-performing is not statistically sound, as additional assessments would be required to provide a statistical "scorecard" for individual airport locations, but covert testing data will be used as a data point in these evaluations. OI will develop a more formal process to uniformly collect and analyze test results to identify best practices attributed to test passes. When specific screening programs or national practices indicate a positive effect on screening performance, TSA will take steps to share and institutionalize best practices in the form of

management advisories to appropriate TSA managers. TSA will explore additional means to effectively disseminate this information.

**Recommendation 3:** As TSA explores the use of covert testing in non-aviation modes of transportation, the agency should work closely with transportation industry stakeholders to develop protocols for conducting tests and coordinate with domestic and foreign organizations that already conduct these tests to learn from their experiences.

**Concur:** TSA has established an excellent rapport with industry stakeholders to allow a free flow of information sharing. Specifically in Mass Transit, TSA follows these practices in approaching systems for “red team” exercises and developing exercise protocols to execute covert testing. In the covert testing exercises TSA Mass Transit has participated in, we have worked closely with the transit agencies being tested, TSA’s internal covert testing program experts, as well as those who have established programs and lessons learned from their own activities. TSA Mass Transit is currently exploring several programs where covert testing may be used as a means for evaluating the effectiveness of various security measures. One example is the I-STEP (Intermodal Security Training and Exercise Program), where TSA is working to develop a model table top exercise to foster the use of similar type exercises around the country. This current effort is centered on the National Capital Region. In subsequent phases of I-STEP, when the program culminates in full scale exercises, opportunities to include covert testing may present themselves.

Another area TSA is exploring is the Visible Intermodal Prevention and Response (VIPR) team activity in support of mass transit and passenger rail security. TSA mass transit is looking at ways to enhance the training and effectiveness of this random, unpredictable deterrence measure. Covert testing could play an appropriate role in this effort. As we move forward in working with our transit security partners, covert testing of VIPR activities, to demonstrative how they integrate into overall security activities, could be important to our collective efforts.

These efforts outlined for Mass Transit will be explored in other non-aviation modes to glean all valuable information concerning lessons learned and other experiences.

**Recommendation 4:** Develop a systematic process to ensure that OSO considers all recommendations made by OI in a timely manner as a result of covert tests, and document its rationale for either taking or not taking action to address these recommendations.

**Concur:** TSA’s Procedures Division is currently coordinating an Operations Directive with the Office of Inspection to formally approve the recommendation and review process for handling recommendations made by OI. In addition the Procedures Division is establishing an access database to track all incoming recommendations and final outcome for incorporating the recommendations into the Standard Operating Procedures.

**Recommendation 5:** Require OSO to develop a process or use information collected through other methods, for evaluating whether the action taken to implement OI’s recommendations mitigated the vulnerability identified during covert tests, such as using follow-up national or local covert tests to determine if these actions were effective.



---

**Appendix II: Comments from the Department  
of Homeland Security**

---

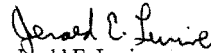
**Concur:** In 2007, the Office of Security Operations established a new program to study various aspects of Transportation Security Officer (TSO) and program performance. Topics of the study include, but are not limited to, recommendations originating from ASAP as well as OI recommendations. Upon the conclusion of each study, recommendations resulting from analysis are provided to TSA leadership for adoption. The program includes:

- Identifying and correlating TSA TSO and/or program data elements, (i.e., scores for Threat Image Projection, Image Interpretation Test, Image Mastery Test; Performance Accountability and Standards System, etc.)
- Conducting related studies, interviews, evaluations, observations, and surveys;
- Convening best practice focus groups; and,
- Providing recommendations for improved performance, effectiveness and/or efficiency.

The first study focused on evaluating the effectiveness of the IED Checkpoint Drills, a nationally implemented program originating from an OI recommendation. Results of this study are expected FY08 Q3. Depending upon the nature of the recommendation, ASAP may also be used to evaluate the national effectiveness of recommendations targeting a specific detection point.

Once again, DHS and TSA appreciate the work GAO has done in the review of TSA's covert testing program.

Sincerely,

  
Gerald E. Levine  
Director  
Departmental GAO/OIG Liaison Office

---

## GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

---

## Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's Web site ([www.gao.gov](http://www.gao.gov)). Each weekday, GAO posts newly released reports, testimony, and correspondence on its Web site. To have GAO e-mail you a list of newly posted products every afternoon, go to [www.gao.gov](http://www.gao.gov) and select "E-mail Updates."

---

## Order by Mail or Phone

The first copy of each printed report is free. Additional copies are \$2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. Government Accountability Office  
441 G Street NW, Room LM  
Washington, DC 20548

To order by Phone: Voice: (202) 512-6000  
TDD: (202) 512-2537  
Fax: (202) 512-6061

---

## To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: [www.gao.gov/fraudnet/fraudnet.htm](http://www.gao.gov/fraudnet/fraudnet.htm)

E-mail: [fraudnet@gao.gov](mailto:fraudnet@gao.gov)

Automated answering system: (800) 424-5454 or (202) 512-7470

---

## Congressional Relations

Ralph Dawn, Managing Director, [dawnr@gao.gov](mailto:dawnr@gao.gov), (202) 512-4400  
U.S. Government Accountability Office, 441 G Street NW, Room 7125  
Washington, DC 20548

---

## Public Affairs

Chuck Young, Managing Director, [youngc1@gao.gov](mailto:youngc1@gao.gov), (202) 512-4800  
U.S. Government Accountability Office, 441 G Street NW, Room 7149  
Washington, DC 20548