

Mr. Steve Lord
Director, Homeland Security & Justice Issues
U.S. Government Accountability Office (GAO)
441 G Street, NW
Washington, DC 20548

Dear Mr. Lord:

Thank you for the opportunity to review and comment on GAO-10-157SU, the draft report titled: *Aviation Security: Efforts to Validate Aspects of TSA's Screening of Passengers by Observation Techniques (SPOT) Program Underway, But Opportunities Exist to Strengthen Validation and Address Operational Changes*. The Transportation Security Administration (TSA) appreciates the U.S. Government Accountability Office's work in planning and conducting its review and issuing this report.

TSA deployed the SPOT program in an effort to mitigate the threat of individuals with potentially hostile intent from boarding a commercial airplane and causing harm. Congress has encouraged the use of behavior recognition to enhance aviation security and has provided resources to support its implementation and expansion. The SPOT program fulfills the mandate of Section 1611 of the Implementing Recommendations of the 9/11 Commission Act, P.L. 110-53, that "TSA shall provide advanced training to the transportation security officers for the development of specialized security skills, including behavior observation and analysis ... in order to enhance the effectiveness of layered transportation security measures."

Intelligence continues to show there is no specific terrorist profile. In a March 10, 2010, hearing before the Senate Homeland Security and Governmental Affairs Committee, TSA Acting Administrator Gale Rossides highlighted the challenge faced by TSA leaders in "balancing the requirement to screen all passengers and to actually focus our officers' attention on the right passengers." TSA designed SPOT to increase its ability to focus on the "right passengers" by identifying persons exhibiting behaviors and appearances that may indicate stress, fear, and deception, and distinguish them from other travelers.

SPOT is Based on Scientific Research and Law Enforcement Practices

TSA's development and deployment of SPOT was a planned and deliberate process based on more than 3 years of operational test-bed assessment of SPOT at Boston's Logan International Airport from June 2003 until nationwide rollout began in fiscal year (FY) 2007. TSA carefully developed SPOT by using selective behaviors recognized within both the scientific and law enforcement communities as displaying stress, fear, and deception. A SPOT working group,

made up of various TSA and U.S. Department of Homeland Security (DHS) components,¹ was created in February 2004. Other organizations, such as the Massachusetts State Police, the Federal Bureau of Investigation (FBI) Behavioral Sciences Unit, and the Federal Law Enforcement Training Center, were also involved in SPOT development. Through these working groups, TSA has developed and finalized SPOT standard operating procedures (SOPs) for a common ability to assess behaviors indicating hostile intent for both aviation and mass transit modes of transportation. TSA continues to consult with its SPOT working group partners as it updates the procedures and science behind the program.

Decades of scientific research have shown the behaviors to be universal in their manifestation. In fact, the DHS Science and Technology Directorate (S&T) completed a study on suicide bomber indicators in July 2009 that illustrates a very high degree of overlap between operationally reported suicide bomber indicators and TSA SPOT behaviors. This result further bolsters TSA's contention that the SPOT program draws from the best practices of many defense, intelligence, and law enforcement organizations.

SPOT Scientific Validation is Ongoing

S&T began research in 2007 to examine the validity of the SPOT program. The series of studies involved in this research is designed to assess the validity of the SPOT scoring system, including the use of individual behavioral indicators to identify high-risk travelers. More specifically, S&T's research plan aims to examine the extent to which these behavioral indicators are appropriate for screening purposes and lead to appropriate and correct security decisions. When this study is complete, SPOT will be one of the most, if not the most, rigorously tested behavior-based security screening programs in existence.

Results of this work will establish a scientific basis of the **extent** to which the SPOT program, including its instrument and methods, such as the SPOT Referral Report and SOPs, are valid. Although it is challenging to establish the validity of a deterrent program in which the outcomes of interest are extremely rare, critical elements of reliability and validity will be rigorously assessed. Of particular importance is the evaluation of criterion-related validity, or the extent to which travelers are correctly selected for screening based on the SPOT scoring system. Establishing this degree of classification accuracy justifies the use of the SPOT program to discriminate high-risk travelers from low-risk travelers. Regardless of any other metrics, the extent to which the SPOT scores accurately identify high-risk travelers is critically important to program validity.

Following criterion-related validity, the next central element of validity is the consistency of implementation of the instrument and program. This will be examined in a variety of ways, including an investigation of the consistency in the operational use of SPOT behavioral indicators Behavior Detection Officers (BDOs) and across locations and time periods, all of which represents reliability assessment. Finally, construct-related validity, or the extent to which the SPOT program behaviors truly represent the expressions of high-risk travelers, will be examined by comparing the SPOT behaviors to similar instruments in use for the same purpose.

¹ Includes TSA's Office of Civil Rights, Office of Chief Counsel, and Privacy Office, and DHS's Policy Office and Transportation Security Laboratory.

S&T's July 2009 study of suicide bomber indicators was the first step in evaluating construct-related validity.

This research is expected to be completed in FY 2011. TSA understands that after this validation is complete, there will be other areas where further research should be conducted, and it is TSA's intention to complete this research.

National Academy of Sciences (NAS) Report Does Not Represent an Exhaustive or Definitive Review of the Research or Operational Literature on Behavioral Screening

TSA would like to specifically address a few comments in the GAO-10-157SU report that we believe are inaccurate. The report draws heavily from a National Academy of Sciences (NAS) report which is being improperly relied upon. As the sponsor of the NAS study, DHS S&T questioned its findings, stating that the study lacked sufficient information for its conclusions because the NAS study principally focused on privacy as it relates to behavioral surveillance—not on behavioral surveillance technology itself. The study was not intended to, and the results do not represent an exhaustive or definitive review of the research or operational literature on behavioral and physiological screening, including recent findings from unpublished DHS, defense, and intelligence community studies. Furthermore, it should be noted that the report did not study the SPOT program, nor did any of the researchers conduct interviews with SPOT program personnel.

Additionally, GAO states that “DHS S&T could not provide us with specific contacts related to the sources of this research.” This statement is not accurate. The record should reflect that DHS S&T provided all requested documents that represented S&T-sponsored research and for which S&T possessed the requisite release authority. DHS was not able to release specific documents related to research for which it was not the originator.

The report further states that the audit team was unable to use the SPOT referral data to assess whether any behavior or combination of SPOT behaviors could be used to reliably predict the final outcome of an incident involving the use of SPOT. However, DHS S&T was able to successfully conduct some preliminary analysis of the SPOT referral database. Prior to analysis of the SPOT reports, S&T worked with TSA to verify the scores assigned to each indicator with the SPOT score sheets and to rescore the pertinent sections and total accordingly for nearly 100,000 operational reports from 2008. While random errors were noted, errors in large databases that require manual entry are not uncommon. Convention suggests that large databases like this typically include an error rate of 3 to 5 percent. As long as such errors are random, the analytical method is robust enough to account for random errors in this range.

In conclusion, TSA strongly believes that behavior detection is a vital layer in its aviation security strategy, and will continue to strengthen as the program matures. Leaders within the community of behavior detection researchers agree. TSA appreciates GAO's work to identify opportunities to enhance the SPOT program, and we will continue to work diligently to address the issues identified by GAO. Our ongoing progress demonstrates our commitment to TSA's mission of securing our Nation's transportation systems.

We also appreciate the opportunity to provide you with, in collaboration with DHS S&T, the attached comments to GAO's audit recommendations.

Sincerely yours,

Jerry Levine
Director
DHS GAO/OIG Liaison Office

Attachment