Scientific Substantiation of Behavioral Indicators

August 17, 2015
Fiscal Year 2015 Report to Congress

Homeland Security

Transportation Security Administration

Message from the Administrator

August 17, 2015

I am pleased to submit the following report, "Scientific Substantiation of Behavioral Indicators," prepared by the Transportation Security Administration.

This report was compiled pursuant to the *Fiscal Year 2015 Department of Homeland Security Appropriations Act* (P.L. 114-4), its accompanying Explanatory Statement, and House Report 113-481. The report provides the scientific substantiation for the use of behavioral indicators and behavior detection as a security capability and includes the current state of the implementation of a new protocol and subsequent test strategies.

Pursuant to congressional requirements, this report is being provided to the following Members of Congress:

> The Honorable John R. Carter
> Chairman, House Appropriations Subcommittee on Homeland Security
>
> The Honorable Lucille Roybal-Allard
> Ranking Member, House Appropriations Subcommittee on Homeland Security
>
> The Honorable John Hoeven
> Chairman, Senate Appropriations Subcommittee on Homeland Security
>
> The Honorable Jeanne Shaheen
> Ranking Member, Senate Appropriations Subcommittee on Homeland Security

If I may be of further assistance, please do not hesitate to contact me at (571) 227-2801 or the Department's Deputy Under Secretary for Management and Chief Financial Officer, Chip Fulghum, at (202) 447-5751.

Sincerely yours,

Peter V. Neffenger
Administrator

Executive Summary

This report contains a detailed discussion on the scientific evidence for the continued use of behavior indicators as a method to identify terrorists. It provides the background for the Transportation Security Administration (TSA) Behavior Detection program and the scientific substantiation. This report supplies the current state of the Behavior Detection program, the implementation of a revised behavior detection protocol, and subsequent plans to test behavior detection rigorously.

As history has shown, terrorists have a variety of means to attempt to inflict harm on an aircraft—everything from explosives in shoes to liquids—but what is consistent across these methods is the actor's malicious intent. TSA's behavior detection approach is designed to identify and engage individuals who may be high-risk (e.g., possess malicious intent) on the basis of an objective process using behavioral indicators and thresholds, and then route them to additional security screening. Behavior Detection is threat-agnostic, and unlike technology, does not become obsolete when the adversary develops a new weapon or tactic. It is one element of TSA's efforts to mitigate threats against the traveling public, and is critical to TSA's layered approach to deter, detect, and disrupt individuals who pose a threat to aviation.

TSA has leveraged the latest research, experience from TSA field operations, and the expertise of leading scientists and law enforcement officials from around the globe to substantiate and improve the behavior detection protocol. TSA's 2014 internal reference guide, Behavior Detection Capability: Discussion on Empirical Support (referred to herein as the Empirical Document), lists 189 scientific and operational references that provide evidentiary support for behavior detection. The guide has been provided to the Government Accountability Office and is available to the Committees. A revised behavior detection protocol is in the early stages of a pilot assessment to support a national deployment decision. If the decision is made to deploy nationwide, TSA will begin a rigorous operational effectiveness testing phase. Such testing would be the first of its kind and has not been attempted by any other domestic anti-terrorism or law enforcement agency. TSA is collaborating with other agencies to pioneer methods to assess the security effectiveness and performance attributes of the behavior detection protocol.

TSA will initiate another outcome-based study to provide further scientific support for behavioral indicators. This effort also will test for disparity issues to address concerns regarding racial, ethnic, or religious profiling. This study, known as the Benchmark Study, is expected to begin at the same time as the operational effectiveness testing efforts. TSA is also allowing its behavior detection protocol and test plans to be reviewed by community advocacy organizations to provide transparency into how TSA

prevents profiling within the behavior detection program. TSA instituted a Community Advisory Panel, comprised of key stakeholders and community advocacy groups, and granted them access at the Sensitive Security Information level to show firsthand how TSA strives to protect privacy and civil liberties, as well as to use their expertise in designing aspects of the Benchmark Study.

Behavior detection methods are based on techniques that have been used by law enforcement and defense organizations for years. The research and scientific basis to support the use of behavioral indicators is both sound and substantial. This report addresses the scientific basis and provides a summary of current state and future activities to further the scientific substantiation of using behavioral indicators as a security measure.

Scientific Substantiation of Behavioral Indicators

Table of Contents

I. Legislative Language

This report is submitted pursuant to the *Fiscal Year 2015 Department of Homeland Security Appropriations Act* (P.L. 114-4), its accompanying Explanatory Statement, and House Report 113-481.

P.L 114-4 states:

TRANSPORTATION SECURITY SUPPORT

For necessary expenses of the Transportation Security Administration related to transportation security support pursuant to the Aviation and Transportation Security Act (Public Law 107–71; 115 Stat. 597; 49 U.S.C. 40101 note), $917,226,000, to remain available until September 30, 2016: *Provided*, That not later than 90 days after the date of enactment of this Act, the Administrator of the Transportation Security Administration shall submit to the Committees on Appropriations of the Senate and the House of Representatives—

> (1) a report providing evidence demonstrating that behavioral indicators can be used to identify passengers who may pose a threat to aviation security and the plans that will be put into place to collect additional performance data; and …

Provided further, That of the funds provided under this heading, $25,000,000 shall be withheld from obligation for Headquarters Administration until the submission of the reports required by paragraphs (1) and (2) of the preceding proviso.

The Explanatory Statement includes the following provision:

> The bill withholds $25,000,000 from obligation until TSA submits to the Committees a report providing evidence that behavioral indicators can be successfully used to identify passengers who may pose a threat to aviation security …

House Report 113-481 states:

Behavior Detection Officers

The Committee believes that questions remain over the value of the BDO program, which has not been sufficiently validated and for which few measures have been developed to prove its intrinsic value to the aviation

security environment. In November 2013, GAO recommended that Congress limit future funding for the BDO program, concluding, among other things, that available evidence does not support whether the behavioral indicators used in TSA's Screening of Passengers by Observation Techniques (SPOT) program can be used to identify persons who may pose a risk to aviation security. Accordingly, to help ensure that security-related funding is directed to programs that have demonstrated their effectiveness, the bill withholds $25,000,000 from obligation for Headquarters Administration until TSA submits to the Committee, not later than 90 days after the date of enactment of this Act, a report providing evidence demonstrating that behavioral indicators can be used to identify passengers who may pose a threat to aviation security and the plans TSA will put into place to collect additional performance data.

II. Background

This report provides evidence demonstrating that behavioral indicators can be used to identify passengers who may pose a threat to aviation security and discusses the Transportation Security Administration's (TSA) plans to continue collecting performance data on the effectiveness of the behavior detection program. The evidence includes 189 scientific and operational references, including the 2011 study, Screening of Passengers by Observation Techniques (SPOT) Referral Report Validation Study completed by the Department of Homeland Security's Science & Technology Directorate. These references are cited and discussed in the Empirical Document. All of these materials have been provided to the Government Accountability Office (GAO) and are available to the Committees.

Since its creation following the terrorist attacks of September 11, 2001, TSA has made great strides in advancing aviation security through investments in innovative technology and human capital, and a continuous iterative approach to enhance security capabilities.

Congress directed TSA to conduct security screening of passengers and baggage, and it does so at approximately 440 airports in the United States that facilitate air travel for approximately 1.9 million people per day, in addition to prescreening more than 14 million passengers each week. TSA also conducts security regulation compliance inspections and enforcement activities at airports for domestic and foreign air carriers and for air cargo screening operations throughout the United States and at last-point-of-departure locations internationally.

TSA is committed to improving security in the most cost-effective manner possible. Through advancements in technology and workforce efficiency, TSA has been able to address known terrorist threats, to include the implementation of the restrictions on liquids, aerosols, and gels, while adapting to changes in aviation travel, such as the growing number of carry-on bags at checkpoints (due to airlines charging fees for checked baggage), and the screening required for the significant increase in the number of laptops carried by passengers. By employing smarter security practices in developing and deploying our people, processes, and technologies, TSA is delivering more effective security in a more efficient manner.

TSA employs risk-based, intelligence-driven principles to reduce the vulnerability of the Nation's transportation system to terrorism. The goal at all times is to maximize transportation security to stay ahead of evolving terrorist threats while protecting privacy, civil rights, and civil liberties, and facilitating the flow of travel and commerce. TSA's security measures create a multi-layered system of transportation security that substantially mitigates risk. Moreover, to remain ahead of those who seek to do harm,

TSA continues to evolve and improve its security approach by constantly evaluating the security procedures and technologies used to protect transportation security.

The Behavior Detection and Analysis (BDA) program, formerly known as the SPOT program, is one of the many capabilities TSA employs to protect the aviation system. The goal of behavior detection is to identify higher-risk individuals and route them toward additional screening, while protecting each passenger's privacy and civil liberties, and preventing unlawful profiling activities. Unlike other security capabilities, behavior detection techniques are unobtrusive, applied in real time, free of large equipment footprints, and threat-agnostic (i.e., focused on the adversary versus the adversary's weapon), allowing implementation in a variety of settings and configurations. This capability fills the gap in other layers of security by helping to identify travelers who potentially pose a high risk of terrorist activity and directing them for additional screening.

The TSA BDA program employs a dedicated and specialized team of Behavior Detection Officers (BDO) who observe and assess passengers for specified behavioral indicators. BDOs conduct operations primarily at airport security checkpoints, positioned where they can optimally observe and engage travelers, especially at or near "stress points" (e.g., Travel Document Checker). These officers look for behaviors that are anomalous to the environmental baseline (defined as typical behaviors that would be expected at the time and place), and that may signal a need for further screening. BDOs look for verbal and nonverbal behaviors associated with a fear of discovery, as well as behaviors related to the stress or anxiety that comes with the execution of a terrorist attack. Additionally, BDOs look for behaviors that are associated with concealment (physical items or knowledge) that could be used in an impending terrorist attack.

While no one system or layer of security can identify every high-risk individual, the behavior detection protocol provides an objective, standardized process in which BDOs observe large numbers of people, identify those who may be high-risk, and devote additional resources (e.g., time, concentration) to them. The use of a standardized set of indicators helps to prevent BDOs from applying biased decision-making or relying on personal hunches or heuristics. BDOs are trained on these standardized indicators and use them when observing passenger behavior to make objective security-related decisions, such as referring an individual to additional screening when that person reaches a predefined threshold. Established thresholds (number of indicators required to qualify for a referral) also create a standardized process that ensures that BDOs have a set of objective criteria that can allow them to take action before an attack is carried out; it is proactive rather than strictly reactive.

Modern behavior detection techniques are an accepted practice within the law enforcement, customs and border enforcement, defense, security, and anti-terrorism communities, and have been for many years. Individual and crowd scrutiny is a hallmark

of the vigilant patrolman, the diligent soldier on the battlefield, and the watchful guard tasked to protect entry to a building. However, despite how commonplace behavior detection methods are, there have been few objective measures applied to behavior detection as practiced by these entities. TSA has been at the forefront of designing the methodology to assess behavior detection operationally and validate the scientific basis. As will be discussed, the evidence supports the use of behavioral indicators to identify potentially high-risk passengers.

III. Discussion and Initiatives

This section provides a detailed discussion of the evidence contained in the aforementioned Screening of Passengers by Observation Techniques (SPOT) Referral Report Validation Study and the Empirical Document reference guide. While regularly testing the effectiveness of any security layer is central to improving upon its potential, developing tests for evaluating the effectiveness of behavior detection as an anti-terrorist tool is difficult due to engineering psychology planning constraints. The low frequency of domestic terrorist attacks represents a challenge for testing because it means that any program that has a component of behavior detection is unable to measure systematically the capability against a true terrorist's behavior. In relation to the Transportation Security Administration's (TSA) program, unless a terrorist is operational at the time of travel, the likelihood of behaviors presenting themselves is quite low. It is the operational terrorists, or the ones who are there to carry out an attack, who are more likely to display the indicators that Behavior Detection Officers (BDO) are trained to detect. Because of this, it is extremely challenging to design a test scenario in a laboratory or operational setting that can fully replicate the circumstances that would stimulate a BDO to respond as if the individual were a real terrorist rather than a test subject.

Despite the methodological challenges, TSA recognized the need for the further study of behavior detection as it relates to anti-terrorism operations. In 2007, TSA partnered with the Department of Homeland Security's Science & Technology (S&T) Directorate to initiate research and studies of behavior detection. In 2011, the S&T analysis, SPOT Referral Report Validation Study, was completed. The results of this study confirmed that the use of behavior detection is over nine times more effective at identifying high-risk passengers than random screening. In the study, high-risk passengers were defined as those passengers possessing dangerous, prohibited, or illegal items; or fraudulent documentation; or who were arrested for other illegal conduct, such as an outstanding warrant. As explained in the study, when the high-risk outcomes were combined, behavior detection is over 54 times more likely than random screening to identify high-risk passengers who were subsequently arrested. While this study was not without limitations, as is any large-scale study of this nature, a thorough examination of the BDOs' impact from these limitations on the results led TSA and S&T to the conclusion that the 2011 study actually underreported the value of behavior detection when compared to a random selection process. A thorough discussion on this topic is included in the Empirical Document.

In 2012, TSA initiated another round of research aimed at improving the security effectiveness, efficiency, and suitability of the behavioral indicators used in the program. This second round of research, expanding upon the existing scientific basis, was concluded in 2014, and the results of this effort are included in the Empirical Document,

which is described below. The scientific knowledge gleaned from this work was used to develop an optimized set of behavioral indicators and a more efficient process to identify potentially high-risk passengers. In addition, operational effectiveness testing will further strengthen the scientific foundation for this capability. These efforts – compilation of the research, optimization of the behavioral indicators, and operational testing – are discussed in more detail below.

IV. Compilation of Relevant Research Material

In October 2014, the Empirical Document was completed. The purpose was to compile the available scientific reference material from academic, open source and other government agencies that was used by the Transportation Security Administration (TSA) to identify and describe the behavior detection processes and indicators relevant to its security mission. The Empirical Document was written to provide internal current and future TSA researchers with a discussion on the analysis of the research materials and how it formed the logical basis for the Behavior Detection Reference Guide (BDRG), which also was completed in October 2014. The BDRG describes the Optimized Behavior Detection (BD) protocol and indicators, and provides the accompanying exemplars, as well as the specific process by which Behavior Detection Officers (BDO) should conduct behavior detection referrals. The Empirical Document includes 189 cited reports and references relating to the scientific research and empirical publications that support the use of behavior detection. This document outlines each behavioral indicator used in the newly developed BD protocol and, for each indicator, provides the specific source that scientifically substantiates the use of that indicator. An index also has been created to show how the original Screening of Passengers by Observation Techniques (SPOT) indicators were integrated into the revised BD list. Based on the analysis of the latest research compilation, most of the original SPOT indicators were either combined, condensed, or revised, with a small subset being removed. In December 2014, TSA delivered these two documents, as well as the Optimized Behavior Detection Concept of Operations, the Optimization Pilot Standard Operating Procedures, and the Optimization Pilot Operational Handbook to the Government Accountability Office (GAO). On February 3, 2015, TSA provided all of the 189 archived research documents included in the Empirical Document to the GAO as well. As previously stated, all of these documents can be made available to the Committees if the GAO has not already provided them.

The 189 references cited in the Empirical Document were obtained from various organizations. These reports include published research studies from academia, government publications documenting research studies of actual suicide bombing attacks, industry-developed research materials, and eyewitness and news media accounts of actual terrorist attacks. Most of these citations are very specific in scope and are directly applicable to behavior detection, as opposed to meta-analyses of aggregated reports, most of which originate from research studies using college students. Some critics have used meta-analyses to question the effectiveness of TSA's Behavior Detection program. Although the research included in these meta-analyses is only tangentially applicable to behavior detection capabilities, there is support for elements of TSA's behavior detection process contained in these reports. Detailed discussions of the applicability and limitations of the meta-analyses also are included in the Empirical Document.

The science behind the continued use of behavior detection is sound. The scientific body of knowledge regarding behavior detection will be further expanded through several projects underway by Department of Homeland Security's (DHS) Science & Technology (S&T) Directorate. This research includes examining how technology can capitalize on the known physiological responses that a person with malicious intent will exhibit. TSA and organizations within the Department of Defense (DOD) also are discussing ways to leverage their respective best practices to combat terrorism. Additionally, TSA is engaging with other countries on behavior detection analyses. TSA and the European Civil Aviation Conference (ECAC) have been meeting for the past several years and exchanging preliminary information through the Study Group on Behavior Detection in Aviation Security (BDIAS-SG), which includes the United Kingdom, France, the Netherlands, Switzerland, and Spain. In June 2015, TSA behavior detection researchers attended the BDIAS-SG conference in Zurich, Switzerland, and briefed TSA's progress with the new optimized protocol and provided an overview of the analysis of current research. Furthermore, TSA has participated in bilateral discussions with the United Kingdom's Department of Transport, exchanging operational and test concepts. Our European partners are actively engaged with the United States in all facets of behavior detection. The BDIAS-SG member states share similar research and validation goals with the United States. This collaboration with the European nations is expanding beyond the ECAC with the inclusion of countries such as South Korea, India, and possibly Singapore. TSA also has observed behavior detection used by Israeli Security Forces at Ben Gurion Airport and other checkpoints in non-aviation settings. In summary, behavior detection is recognized domestically and around the globe as an important security layer.

A. Optimized Behavior Detection Protocol

Behavior Detection Optimization evolved as research materials were compiled and as it became evident that TSA had an opportunity to reexamine the behavioral indicators and improve upon the referral determination process. The two primary sources of material used to revise the indicators included TSA-sponsored efforts related to Optimization and an S&T study related to suicide bombers. The S&T study on suicide bombers was a comprehensive review of suicide bombing attacks around the world. It included analysis of video and interviews with failed suicide attackers. This rigorous study of real-life terrorist attacks yielded critical experiential data that either corroborated or refuted academic studies in terms of what had occurred behaviorally prior to an attack. This work was also valuable in optimizing the behavioral indicators used in TSA's behavior detection program. Additionally, TSA sponsored an effort by the American Institute for Research (AIR), which included a multi-pronged study that canvassed all of the updated research literature that either did not exist or was otherwise unavailable when TSA initiated the SPOT program. This work also provided information on the feasibility of future behavior detection testing and a design for a more robust outcome-based

examination of individual indicators and demographic disparities. AIR held focus groups with approximately 75 BDOs from around the country as well as separate behavior detection subject-matter expert panels. These panels included experts from TSA, the Federal Bureau of Investigation, academia (national and international representatives), and the United Kingdom's Department of Transport. All of this material was used to refine the behavioral indicators and revamp the assessment methodology used when referring an individual for additional screening. In developing the BD protocol, TSA's goals were to increase security effectiveness, improve operational efficiency, and ensure the protocol's use was suitable for BDOs.

TSA completed the revised protocol in August 2014, and the training materials were ready for the first large-scale deployment in October 2014. The plan includes piloting the BD protocol at three to five airports. The success criteria for this pilot include the BDOs' ability to apply the protocol with no significant adverse effects on screening operations (e.g., resource constraints). Portland International Airport (PDX) and Seattle-Tacoma International Airport (SEA) were the first two airports to receive the new protocol. The training at PDX began in October 2014 and concluded 6 weeks later. SEA training commenced in January 2015 and concluded in May 2015, the extra time at SEA being attributed to a larger population of BDOs as compared to PDX. The remaining three pilot airports are Fort Lauderdale Hollywood Airport (FLL), Miami International Airport (MIA), and Tampa International Airport (TPA). Training commenced at FLL in June 2015, with the remaining two sites tentatively planned for later in calendar year (CY) 2015. Preliminary results are very promising and TSA may conclude the pilot phase prior to the implementation at MIA or TPA. If it is determined that the pilot objectives are satisfied, TSA will deploy the BD protocol nationally and a rigorous testing phase will begin. The results of the pilot will be issued in a final report at the conclusion of the pilot.

The revised protocol seeks to increase BDO efficiency and effectiveness by simplifying, eliminating, and consolidating behaviors originally listed in the program. While the pilot testing of the BD protocol is underway, the continued use of the current indicators in airports does not cause security concerns and is not problematic. Because BDOs are already accustomed to using these indicators (which are also supported by the cited research), and the rigor associated with the new training, TSA believes that it is better to provide comprehensive training to BDOs rather than altering the protocol section by section. Accordingly, TSA will introduce all of the changes to each airport when the airport is retrained on the new BD protocol rather than implementing changes to the current protocol in a piecemeal manner.

B. Behavior Detection Test Concepts

At the successful conclusion of the current BD pilot, the former protocol will be replaced around the country with the revised protocol. As BD rolls out, testing will begin at select

locations and will involve two main tracks that encompass several areas of research: operational testing and the Benchmark Study. The operational testing track will examine the security effectiveness of Behavior Detection, while the Benchmark Study will examine whether disparity issues exist within the program. The TSA Office of Security Capabilities Operational Support Division will conduct the testing. This organization has been recognized as an independent Operational Test Agent by the DHS Director of Operational Test and Evaluation for programs that have been designated as Acquisition Level 1 or Special Oversight categories. The testing conducted on behavior detection will be applied with the same level of rigor as is given to these categories of acquisition programs. Both of the two tests will be discussed in greater detail below.

The timing of the two interrelated test events follows:

> (1) Operational Test: The 10 operational test sites include the five pilot airports and the next five locations scheduled to receive the revised protocol. Each operational test site is expected to be subjected to data collection activities for 5 to 7 months. Since operational testing will not commence simultaneously at each of the 10 sites, data collection for all 10 sites could span 1 year. Assuming no delays, it is estimated the operational test data collection can be completed by the end of CY 2016.

> (2) The Benchmark Study: The outcome-based data set to evaluate disparity requires the results, or "outcomes," of over 90,000 referrals for a significant sample size. These will be obtained from at least 50 airports during an entire calendar year at each site to account for seasonal variations. It is expected that this data collection period will require, at a minimum, 3 years to conclude, as this study coincides with the timing of the national implementation of the BD protocol.

There are three behavior detection test objectives that TSA aims to satisfy with the two overarching test strategies: The first objective is the pioneering effort to quantify operational security effectiveness. Despite the widespread use of behavior detection techniques, there are no relevant significant historical operational test results from any domestic or international defense, security, or law enforcement agency that can illustrate the security effectiveness of behavioral observations. In fact, TSA could not identify any metrics that specify an acceptable level of performance for behavioral observation currently in use by any organization. Following the operational test, TSA will analyze the data and compare those results to the estimates used in risk models and establish the required thresholds for behavior detection security effectiveness. The second test objective is to capture significantly more data than is currently available for analysis from operational sources or previous studies so that each individual indicator can be examined for continuous refinement. The third objective is to examine the behavior detection protocol to ensure that it does not systemically lead to the unwarranted targeting of

individuals due to demographic, ethnic, or religious categorizations (Benchmark Study outcomes).

C. Threat-Inject Testing

Two different test strategies will be used to meet these three objectives. The first strategy involves threat-inject methodologies that are similar to tests that TSA and other agencies conduct on technologies in support of acquisition decisions. This involves sending mock threats through the environment in order to simulate adversary behavior. The second test strategy will be a robust outcome-based study (Benchmark Study methodology) that uses standard Behavior Detection outcomes, as well as outcomes stemming from random selections for referrals. Specific mention of the exact objective measures, and their results for behavior detection must be treated as Sensitive Security Information (SSI).

Operational data collected by TSA will be analyzed as they relate to proximate measures. Due to the low base rate of known actual terrorists transiting through our Nation's airports, reliance must be placed upon representative measures. There is a strong correlation between many of the indicators that an active terrorist would display and the behaviors that another individual in a heightened state of anxiety or stress would display, such as an individual fearing engagement or apprehension by lawful authorities (i.e., a criminal, a smuggler, a fraudulent identification holder). Additionally, individuals who are low-risk but are nevertheless experiencing heightened levels of stress or anxiety, possibly due to a recent significant event, may also display behavioral indicators. Though it is beyond the scope of this report, TSA acknowledges that the similarities between behaviors displayed by terrorists and nonterrorists make prior differentiation between the two groups difficult. However, by measuring the rates at which BDOs accurately assess behavioral indicators, refer individuals to additional screening, and the frequency with which these referrals lead to high-risk outcomes, it can be derived that performance metrics can provide some indication of security effectiveness. These measures allow us to assess how effective the behavior detection protocols are at enabling BDOs to observe and properly assess behaviors.

TSA is collaborating with S&T, DOD, other domestic and international organizations, and experts on devising protocols, whereby behavior detection can be tested operationally. Unlike an x-ray system where actual or simulated explosives can be used to test its explosives-detection functionality, there are no mature processes that fully simulate a terrorist and can sufficiently stimulate a security officer without introducing artificial bias in the outcome in use by any organization. One method currently in development involves scenario-based testing. This testing will be done in a manner similar to that used to test transportation security equipment and primarily will involve threat-inject scenarios or covert tests. This type of testing, however, is challenging. The objective is to examine the likelihood that behavior detection is able to identify a terrorist about to commit an imminent attack. The individual who is to conduct the threat inject

trial then must be able to exhibit the behaviors in such a manner that the BDOs under test will not be able to discern that it is a test subject they are observing, rather than a high-risk passenger or terrorist. This is extremely difficult to do, because in most instances, the genuine display of behavior is difficult to simulate.

S&T and TSA have studied the feasibility of using actors to simulate such behaviors. It seems unlikely that acting or simulation alone will be sufficient. In the initial feasibility trials, it did not appear there were a sufficient number of behaviors to use for testing that can be reasonably simulated in the clusters (or combinations of behaviors) required to meet the threshold for a behavior detection referral. The few clusters that were demonstrated via acting have the potential to quickly become predictable to BDOs and negate their utility for covert security effectiveness tests. However, TSA is continuing to refine this method and another round of feasibility trials are to be run prior to including simulated behaviors within the testing space. Techniques using simulated behaviors also will provide a useful capability that TSA can use to support training and supervisory or quality assurance assessments.

The only other viable test option is to induce the tester to exhibit behaviors naturally. This method is difficult to achieve without conditioning the subject using arguably nefarious scenarios to believe that they are participating in an actual terrorist attack. While academic circles for years have been using deception under the oversight of Institutional Review Boards, TSA will not use deception and instead will attempt to induce the desired state of consequence or trepidation in the covert testers. TSA will use methods that create a sense of consequence through stress, fear of failure, or anxiety surrounding the test, where the subject knows they are testing security, but do not believe they are part of an active terrorist attack. The test community within TSA is actively collaborating with S&T, DOD covert test teams, international allied government agencies, and academic and leading industry experts on this challenge. These specific methodologies can be shared with Congress in an SSI-level briefing.

Our European partners have experienced some success with this method, and there is a reasonable expectation that a subset of behaviors can be exhibited naturally by a covert tester. However, it is very likely that only a small sample of the individual behavioral indicators will be exhibited by the testers in a reasonable test period due to the lack of perceived consequence that the testers experience, as well as variations in the temperament and behavior of the testers. It is equally likely that the resulting BDO referral rates obtained from both methods, acting or simulated behaviors and the induced naturally exhibited behaviors of a tester, cannot be definitively proven to correlate with the actual BDO referral rates against actual active terrorists. It ultimately can be concluded that covert tests provide a reference measure of security effectiveness that is useful for risk models but not for measuring performance against an operational terrorist. It is anticipated that the Behavior Detection and Analysis (BDA) program will be able to use the data to define a reasonable range within which the actual performance parameters

should fall to further support risk assessment calculations. Additionally, due to this uncertainty over the correlation between test subject and terrorist behavior, a test of this nature will not be able to validate each behavioral indicator; instead, it will serve as a mechanism by which to judge the efficacy of behavior detection as a system, as opposed to individual subcomponent indicators.

It must be understood that this is a pioneer endeavor and these test practices have not previously been attempted with scientific rigor by any agency in the United States or overseas. No other law enforcement or defense agency has been asked to test their officers' efficacy in identifying terrorists, criminals, or other high-risk individuals in this manner. Past test events that have used actors or other covert test methods have significant limitations and test biases, and cannot provide a definitive engineering resolution to the question of behavior detection security effectiveness as they cannot withstand scientific scrutiny. However, these previous covert or "red team" tests still hold utility for operations planners to estimate the likelihood of defeating an adversary.

Furthermore, none of the tests discussed in this report will be able to measure the deterrence value associated with Behavior Detection. As with a military operation, a terrorist operation will attempt to avoid defenses that will impede the accomplishment of its objective. Deterrence, by its very nature, is difficult to measure. The deterrence value of any security protocol, procedure, or defensive barrier can be discussed, but a metric has not been devised to scientifically measure deterrence effectiveness. However, based upon current intelligence, there is a deterrence effect. Recent law enforcement bulletins have indicated that the Islamic State in Iraq and Syria have begun instructing members on potential methods to avoid notice by security forces employing behavior detection, even rudimentary observation.

D. Comparing Operational and Random Referral Data

In addition to the security effectiveness testing, TSA is also conducting a broader study, the aforementioned Benchmark Study that gathers data on the outcomes of Behavior Detection referrals and the outcomes of randomly selected passengers for screening. This study will commence at the start of the operational test. This data set will be analyzed with two purposes: to understand potential disparity issues and identify protocol improvement opportunities. The prevention of systemic, unlawful profiling has been at the nucleus of TSA behavior detection prior to the development of SPOT. TSA is conducting the Benchmark Study to ensure that behavior detection protocols do not systemically target individuals based upon demographic or ethnic characteristics. The data set that is required for the Benchmark Study is so significantly scoped that it affords TSA the opportunity to investigate the relationships between many of the behavior indicators, high-risk outcomes, airport and seasonal environmental impacts, populations, officer human factors engineering, and other considerations.

This study is designed to mitigate the limitations that TSA, DHS, and the GAO identified in the 2011 SPOT Referral Report Validation Study. The Benchmark Study data requirements are significantly higher than the requirements for a typical TSA operational test. This is due to the need for sufficient data in order to analyze the corollary relationships between outcomes of varying types, analyze passenger demographics of varying types, and conduct an adequate comparison of the referred and randomly selected populations. The design of experiments for the Benchmark Study requires a minimum of 90,000 random screenings and as many operational referral screenings that are conducted in the same time period. These data points will be gathered across the minimum 50 airports that have been converted to the BD Protocol and ideally across the remaining locations that will be converted. To account for seasonal variations, the data will be collected across 12 to 15 months. This equates to an average of four random screenings per day at each airport. The number is expected to vary with the passenger volume of each airport, for example, with more random referrals occurring at a Category X airport versus a Category I airport.

The randomly screened population is necessary to develop the baseline against which to compare the BDO referrals. The analysis of the randomly selected population and research into each airport's passenger demographic population distribution will provide the basis upon which to weigh the Behavior Detection referrals. The researchers will examine the distribution of referrals against the distribution of the randomly selected population to see if there are occurrences where a segment of the population is over- or under-referred, as compared to the population norms. This robust data set will allow for the analysis of behavior indicators that may be more prevalent with certain populations based upon cultural norms. For instance, there are certain cultures that are averse to direct contact with law enforcement or security officials. In such a circumstance, further investigation would be needed to examine the underlying reasons for any deviations from normal distribution patterns to determine if there is a systemic issue with the protocol's design; geographical, seasonal, or other environmental unforeseen causal factors; cultural factors; or improper activities on behalf of the workforce. The resolution of any findings will depend on the particular factor(s) involved and the subsequent impact on security and the traveling public. These could range from simple acceptance of the deviation, changes to the exemplars within the training curricula, adjustment to the referral mechanisms within the protocol, or the discontinued use of the particular indicator in question.

TSA has taken great strides to ensure that the workforce receives clear guidance and training on preventing any form of unlawful profiling. Additionally, TSA takes any and all allegations of unlawful profiling seriously. The offensive and objectionable use of race, ethnicity, religion, nationality, gender, sexual orientation, or gender identity as a criterion in conducting BDA Referral Screening or other screening activities is strictly prohibited and will not be tolerated. Officers must immediately notify management if they observe or believe that unlawful profiling has occurred. Additionally, Officers must

not retaliate against any employee, contractor, or member of the public who has made a civil rights/liberties complaint, engaged in a protected activity, or has shared civil rights concerns with others. Any employee found in violation of this policy will receive disciplinary action, up to and including termination.

TSA has had a long-term relationship with the community it serves and has established the TSA Disability and Multicultural Coalition, consisting of multicultural, disability, religious, and civil rights and liberties advocacy organizations that meet regularly with TSA to discuss issues of concern to their respective constituencies. TSA's Office of Security Capabilities, in conjunction with the TSA Office of Civil Rights & Liberties, Ombudsman, and Traveler Engagement's Multicultural Branch, also established a Community Advisory Panel (CAP) to provide transparency over the Benchmark Study and to engage representatives from various nationally recognized multicultural, religious, and civil rights and liberties advocacy organizations.

The CAP presents an opportunity for TSA and interested external stakeholders to hold SSI-level discussions specifically centered on the Benchmark Study and its eventual findings. In the summer of 2014, the TSA Deputy Administrator approved the conditional release of SSI information to the CAP membership provided that the individuals passed the SSI pre-clearance checks, conducted by TSA's Office of Intelligence and Analysis, and signed nondisclosure agreements. The first meeting of the CAP whereby SSI materials were shared was held on October 28, 2014, followed by two additional meetings in the subsequent months. The CAP is an unprecedented invitation to the nationally recognized advocacy organizations to see the actual BDA protocols and learn how TSA conducts behavior detection. The CAP members view this information and provide individual opinions from their own constituencies' perspectives on the Benchmark Study's methodology, results, data, and implications. TSA is not seeking collective, consensus advice from the CAP though, because TSA understands that each constituency will have different yet equally valid opinions to assist TSA in making sure that unlawful profiling is prevented.

V. Conclusion

Behavior detection is a vital component of the Transportation Security Administration's (TSA) overall security posture. In order to stay ahead of evolving threats, there must be a security capability that transcends current technological solutions and can remain flexible in order to protect the homeland. Behavior detection is such a capability and provides the means to identify potentially high-risk individuals when the method of attack is not readily known or may otherwise escape detection.

TSA has shown that there is a significant body of scientific evidence and operational literature that supports the use of behavior detection indicators to identify high-risk passengers. TSA has compiled over 189 documents that include scientific research studies and exemplars from operational events that illustrate the reasoning for the use of the indicators in identifying terrorists who are an imminent threat. TSA is using iterative systems engineering approaches in the establishment of benchmark metrics, and the development and continued improvement of the behavior detection protocols. TSA is currently assessing an improved protocol with the goal of commencing a national deployment by the end of calendar year 2015.

A comprehensive and pioneering series of operational tests will be conducted on this revised protocol to evaluate and optimize security effectiveness, as well as to answer disparity questions. TSA will continue to collaborate on the security effectiveness test methodologies within the Department of Homeland Security, and with the Department of Defense and academic and international experts to ensure continued transparency on the subject. TSA welcomes the continued dialogue with Congress and the Committees on the subject of behavior detection.

VI. Abbreviations/Acronyms

Abbreviation/Acronym	Definition
AIR	American Institutes for Research
BDA	Behavior Detection and Analysis
BDO	Behavior Detection Officer
BDIAS-SG	Study Group on Behavior Detection in Aviation Security
BDRG	Behavior Detection Reference Guide
CAP	Community Advisory Panel
CY	Calendar Year
DHS	Department of Homeland Security
DOD	Department of Defense
ECAC	European Civil Aviation Conference
FLL	Fort Lauderdale Hollywood Airport
GAO	Government Accountability Office
PDX	Portland International Airport
S&T	Science & Technology
SPOT	Screening of Passengers by Observation Techniques
SSI	Sensitive Security Information
TPA	Tampa International Airport
TSA	Transportation Security Administration