Homeland Security

SPOT Referral Report Validation Study Final Report Volume II: Appendices A through E

April 5, 2011

U. S. Department of Homeland Security
Science and Technology Directorate
Washington, D.C. 20528-0207

APPENDIX A:
SPOT REFERRAL REPORT AND
LIST OF SERIOUS PROHIBITED/ILLEGAL ITEMS

Introduction

Appendix A includes the Screening Passengers by Observation Technique (SPOT) Referral Report and a list of Serious Prohibited or Illegal Items with definitions for coding referrals in each database.

SPOT Referral Report

The SPOT Referral Report is the screening instrument used by Transportation Security Administration's (TSA's) SPOT Program during checkpoint screening. The SPOT Referral Report was the original instrument and central element in the SPOT Program when launched in 2006. Since that time, the program has expanded to include SPOT screening operations at other locations throughout the airport, as well as other SPOT playbook operations. However, the Referral Report instrument has remained largely unchanged since program inception and is central to the program's effort to identify high-risk passengers at checkpoint screening. AIR has inserted this document here. The associated SPOT Standard Operations Procedures (SOP) is found in Appendix I or available upon request.

SENSITIVE SECURITY INFORMATION

REV 4

SPOT REFERRAL REPORT

Version 4.0 Revised 02/23/2009

PART I: Behavioral Observation

(b)(3):49 U.S.C. § 114(r)

Section 1 Environmental Baseline (Please provide a brief narrative of the environmental baseline at the time of the referral below)

(b)(3):49 U.S.C. § 114(r)

Section 2 Observation and Behavior Analysis

(b)(3):49 U.S.C. § 114(r)

PART II: SPOT Resolution

(b)(3):49 U.S.C. § 114(r)

Section 3 Unusual Items NONE

(b)(3):49 U.S.C. § 114(r)

Section 4 Signs of Deception NONE

(b)(3):49 U.S.C. § 114(r)

SENSITIVE SECURITY INFORMATION

WARNING: THIS RECORD CONTAINS SENSITIVE SECURITY INFORMATION THAT IS CONTROLLED UNDER 49 C.F.R. PARTS 15 AND 1520. NO PART OF THIS RECORD MAY BE DISCLOSED TO PERSONS WITHOUT A "NEED TO KNOW," AS DEFINED IN 49 C.F.R. PARTS 15 AND 1520, EXCEPT WITH THE WRITTEN PERMISSION OF THE ADMINISTRATOR OF THE TRANSPORTATION SECURITY ADMINISTRATION OR THE SECRETARY OF TRANSPORTATION. UNAUTHORIZED RELEASE MAY RESULT IN CIVIL PENALTIES OR OTHER ACTION. FOR U.S. GOVERNMENT AGENCIES, PUBLIC DISCLOSURE GOVERNED BY 5 U.S.C. 552 AND 49 C.F.R. PARTS 15 AND 1520.

1 of 2

Sensitive Security Information (SSI)

WARNING: THIS RECORD CONTAINS SENSITIVE SECURITY INFORMATION THAT IS CONTROLLED UNDER 49 C.F.R. PARTS 15 AND 1520. NO PART OF THIS RECORD MAY BE DISCLOSED TO PERSONS WITHOUT A "NEED TO KNOW," AS DEFINED IN 49 C.F.R. PARTS 15 AND 1520, EXCEPT WITH THE WRITTEN PERMISSION OF THE ADMINISTRATOR OF THE TRANSPORTATION SECURITY ADMINISTRATION OR THE SECRETARY OF TRANSPORTATION. UNAUTHORIZED RELEASE MAY RESULT IN CIVIL PENALTIES OR OTHER ACTION. FOR U.S. GOVERNMENT AGENCIES, PUBLIC DISCLOSURE GOVERNED BY 5 U.S.C. 552 AND 49 C.F.R. PARTS 15 AND 1520.

SENSITIVE SECURITY INFORMATION

SPOT REFERRAL REPORT

Version 4.0 Revised 02/23/2009

PART II: SPOT Resolution (continued)

Section 5 Automatic LEO Notification — **No LEO Referral**

(b)(3):49 U.S.C. § 114(r)

PART III: Data Collection

Section 6 Passenger Data

(b)(3):49 U.S.C. § 114(r)

Section 7 Prohibited Items — **NONE**

(b)(3):49 U.S.C. § 114(r)

Section 8 Resolution Statement — **Resolution of Behaviors:**

(b)(3):49 U.S.C. § 114(r)

Section 9 LEO Resolution — **LEO Resolution Notes: (Include information such as officer's name and ID# in Notes block below)**

Primary LEO:

- ☐ ATF
- ☐ CBP
- ☐ DEA
- ☐ FAM
- ☐ FBI
- ☐ ICE
- ☐ LOCAL LEO
- ☒ OTHER

List Other LEO:

Reason for Arrest:

- _____ Outstanding warrants
- _____ Undeclared currency
- _____ Suspected drugs
- _____ Illegal alien
- _____ Fraudulent documents

Screening Manager Notified: (Last, First)

CCTV Copied: _____ YES _____ NO

2 of 2

Sensitive Security Information (SSI)

WARNING: THIS RECORD CONTAINS SENSITIVE SECURITY INFORMATION THAT IS CONTROLLED UNDER 49 C.F.R. PARTS 15 AND 1520. NO PART OF THIS RECORD MAY BE DISCLOSED TO PERSONS WITHOUT A "NEED TO KNOW," AS DEFINED IN 49 C.F.R PARTS 15 AND 1520, EXCEPT WITH THE WRITTEN PERMISSION OF THE ADMINISTRATOR OF THE TRANSPORTATION SECURITY ADMINISTRATION OR THE SECRETARY OF TRANSPORTATION. UNAUTHORIZED RELEASE MAY RESULT IN CIVIL PENALTIES OR OTHER ACTION. FOR U.S. GOVERNMENT AGENCIES, PUBLIC DISCLOSURE GOVERNED BY 5 U.S.C. 552 AND 49 C.F.R. PARTS 15 AND 1520.

Serious Prohibited/Illegal Items

The purpose of the SPOT Referral Report is to identify "high-risk" travelers, defined as those individuals who are knowingly and intentionally attempting to defeat the airport security process. The analyses conducted for the Validation Study used four outcome variables to represent high-risk travelers: (1) *LEO* (Law Enforcement Officer) *Arrest*; (2) *Possession of Serious Prohibited of Illegal Items*; (3) *Possession of Fraudulent Documents* (a subset of [2]); and (4) *Combined Outcome* (i.e., LEO arrest and/or possession of serious prohibited/illegal items [including fraudulent documents]).

Figure A-1 provides the list of Serious Prohibited or Illegal Items, which includes fraudulent documents. The list is based on a TSA List of Serious Prohibited Items—that is, items for which the Transportation Security Officer (TSO) must notify a Supervisory Transportation Security Officer (STSO), who will then notify a Bomb Appraisal Officer (BAO) for explosives or a LEO for all other items or if the BAO is not available. The list also includes items that warrant automatic LEO notification (e.g., travel documents, drugs, specific weapons, or serious prohibited/illegal items).

Figure A-1: List of Serious Prohibited/Illegal Items

(b)(3):49 U.S.C. § 114(r)

(b)(3):49 U.S.C. § 114(r)

APPENDIX B:
OPERATIONAL SPOT REFERRAL DATABASE OVERVIEW

Introduction

Appendix B provides information about the Operational SPOT process, the SPOT Referral Report, and the datasets made available by TSA. This appendix first presents an overview of the SPOT process and then the Operational SPOT referral database is described. Finally, we present a brief description of the procedures adopted by AIR to support data preparation activities.

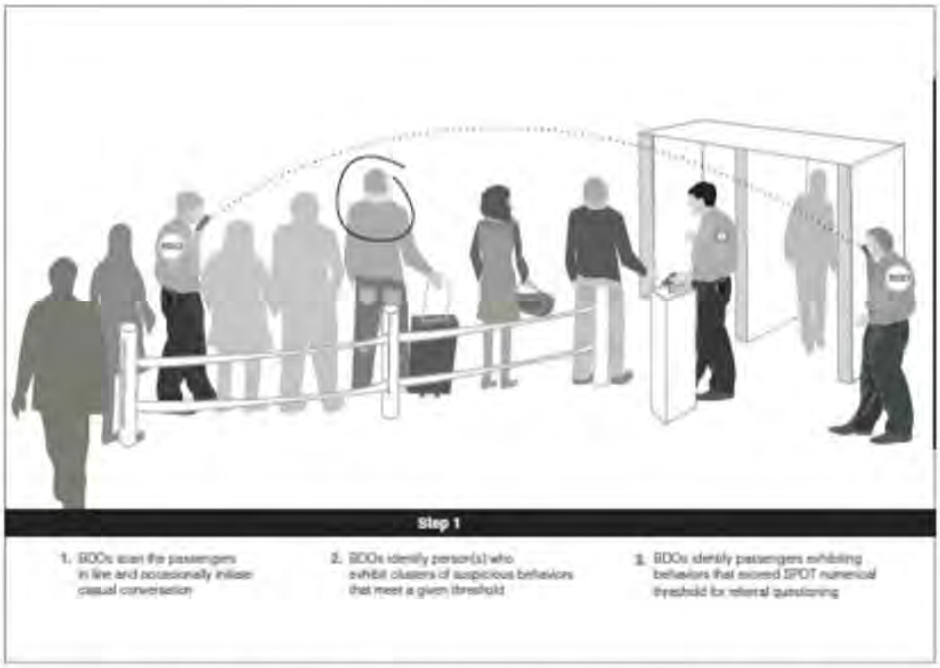
Operational SPOT Process

The SPOT Process

The SPOT process involves the use of behavior observation and analysis techniques by a pair of BDOs to identify potentially high-risk individuals among the traveling population. BDOs conduct SPOT operations at security checkpoints and other designated areas at those airports that have existing SPOT programs. The BDO teams use the SPOT Referral Report to identify and screen suspicious travelers.

The SPOT process begins with the BDO's assessment of the environmental baseline and considers the factors (b)(3):49 U.S.C. § 114(r) that characterize the time and place in which SPOT is being conducted. Once the baseline is established, the BDOs observe the airport checkpoint security lines (starting from the point at which travelers enter the checkpoint) and identify whether any travelers are exhibiting any of the Section 2 indicators on the SPOT Referral Report (see Figure B-1). The focus of the observation is on displays of these indicators that deviate from what might be expected given the existing environmental baseline. Note that it is possible for the environmental baseline to change during a shift or security operation.

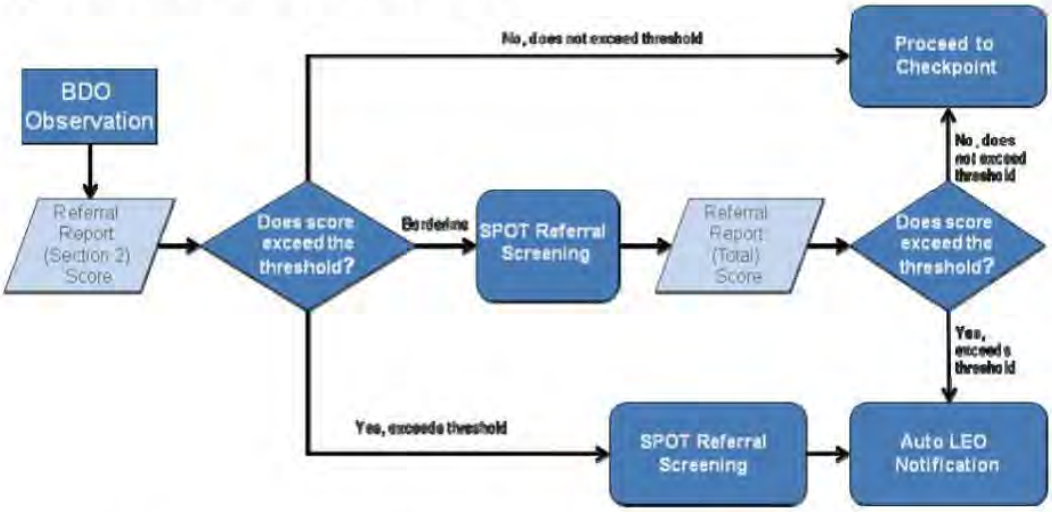
Figure B-1: BDOs Observe Travelers in Checkpoint Queue [1]



Typically, one BDO moves along the checkpoint queue engaging as many individuals as possible in a brief verbal exchange while watching for reactions to the presence and engagement of security personnel. The BDO partner observes reactions to the presence of security personnel, while continuing to observe the entire area for the display of behavioral indicators (i.e., SPOT Referral Report Section 2 indicators). BDOs assign points to individuals whose behavior or appearance deviates from the established baseline. As shown in Figure B-2, to be selected for a SPOT referral, an individual must meet a pre-defined threshold of (b)(3):49 points that are accumulated during the observation. A second threshold of (b)(3):49 U points is used to determine whether or not to involve a LEO and is based on the number of points accrued and/or the type of indicators displayed. During observation, the BDOs mentally calculate the point accrual for an individual. For those individuals who are selected for SPOT Referral Screening, the BDOs typically make notes about the case and formally document the points and other notes in the SPOT Referral database at the end of their shift. The SPOT Referral Report is completed only for those individuals who are selected for SPOT Referral Screening.

[1] United States General Accounting Office. (2010, May). *Aviation security: Efforts to validate TSA's passenger screening behavior detection program underway, but opportunities exist to strengthen validation and address operational challenges* (Publication No. GAO-10-763). Washington, DC: U.S. Government Printing Office.

Figure B-2: Section 2 Screening Thresholds



As Figure B-2 shows, there are three possible outcomes for the SPOT process at the checkpoint. First, if a traveler exhibits no or few indicators, at a level that is below the SPOT threshold, there is no further screening and the traveler continues on to his or her destination. This outcome accounts for about 90% of travelers entering checkpoints where SPOT is in operation. Second, if a traveler exhibits some indicators and reaches a borderline-range score of (b)(3):49 he or she is selected for SPOT Referral Screening. The BDOs use this secondary screening opportunity to attempt to resolve the case. If this borderline case cannot be resolved and/or the traveler's SPOT score increases to (b)(3):49 U.S.C. § a LEO is called to conduct further screening and resolve the case. Finally, if a traveler in line is observed exhibiting enough indicators to exceed a certain SPOT threshold score (b)(3):49 U.S.C. § the traveler is subjected to SPOT Referral Screening and a LEO is automatically called in to conduct further screening to resolve the case. SPOT Referral Screening includes a non-custodial interview with the referred individual (often referred to as the Casual Conversation), as well as a (b)(3):49 U.S.C. search of his or her accessible property. At the end of the SPOT Referral Screening, the BDO pair confer and confirm the details of the case.

A LEO is also brought in to conduct additional screening if prohibited items, fraudulent documents, or other serious issues are identified during the SPOT Referral Screening. Once involved, the LEO is responsible for case resolution. The LEO may determine that the traveler may proceed through the checkpoint and continue to his or her destination, retain the traveler for further questioning, and/or arrest the traveler. The final resolution is documented in the SPOT Referral Report.

Note that although SPOT observations may be conducted throughout the airport, the focus of this Validation Study is to examine the use of the SPOT Referral Report (b)(3):49 U.S.C. § 114(r)

only. At the checkpoint, SPOT observations and the corresponding SPOT Referral Report to document decisions are used from the time a traveler enters the checkpoint line and proceeds through the magnetometer, picks up any belongings or bags, and is subjected to any additional screening, as required. The use of the SPOT Referral Report by the BDO ends for this traveler when he or she has completed screening and is permitted to leave the sterile area. Recently, the SPOT Program expanded its observation area to include locations beyond the checkpoint area. At the time this study was designed and first implemented, the SPOT Referral Report was used at the checkpoint only. Thus, the focus of this analysis was confined to this portion of the SPOT screening process.

SPOT Referral Report

The SPOT Referral Report, which is the instrument under investigation in this Validation Study, is the paper-and-pencil checklist used by BDOs to record key data and observations when conducting SPOT operations in the airport. As shown in Appendix A, the SPOT Referral Report consists of three parts: Part I—*Behavioral Observation*, Part II—*SPOT Resolution*, and Part III—*Data Collection*.[2] These three parts are further divided into a total of nine sections. The following list contains a brief description of the information collected within each of the nine SPOT Referral Report sections.

- *Part I—Behavioral Observation*
  - *Section 1: Environmental Baseline.* This is an open text field for providing a brief narrative of the environmental baseline at the time of the referral. Environmental baseline refers to the typical behaviors and appearances that would reasonably be expected at the time and place SPOT is being conducted. (b)(3):49 U.S.C. § 114 (b)(3):49 U.S.C. § 114(r)
  - *Section 2: Observation and Behavior Analysis.* This section contains the checklist of behavioral and appearance factors used for documenting results of a BDO's visual assessment of individuals, their accessible personal property, and their interaction with others. The visual assessment focuses on identifying factors associated with stress, fear, and deception and assigns points to individuals whose behavior or appearance related to these factors deviates from the established baseline.

[2] As of March 1, 2009, as part of the change in SPOT's SOP, the SPOT Score Sheet was revised and renamed, the SPOT Referral Report. Along with the name change, the sections were reordered to accommodate the inclusion of additional sections. Throughout this report, we use terminology associated with this new SOP.

- *Part II—SPOT Resolution*
  - *Section 3: Unusual Items.* This section is a checklist of items that the individual (b)(3):49 U.S.C. § 114(r)

  - *Section 4: Signs of Deception.* This section contains the checklist of signs of deception the BDO looks for during the SPOT Referral Screening. (b)(3):49 (b)(3):49 U.S.C. § 114(r)

  - *Section 5: Automatic LEO Notification.* This section is a checklist of reasons for an automatic LEO notification, such as discovery of firearms, dangerous weapons, hazardous materials, or explosives; (b)(3):49 U.S.C. § 114(r) (b)(3):49 U.S.C. § 114(r) and discovery of suspected unlawful drugs.

- *Part III—Data Collection*
  - *Section 6: Passenger Data.* This is the location for documenting the passenger's air carrier, flight number, origination airport, next destination airport, and final destination airport. Note that no personally identifiable information (PII) is included in this section or any section of the SPOT Referral Report.

  - *Section 7: Prohibited Items.* This is the location for documenting prohibited items uncovered during the SPOT Referral Screening as well as where these were found (i.e., x-ray, bag search, surrendered before checkpoint).

  - *Section 8: Resolution Statement.* In this section, the BDO enters a statement that documents the nature of the resolution (e.g., resolved by TSA during screening, denied boarding by carrier, questioned and arrested by LEO). This section also includes a text box for providing notes on the resolution of identified behaviors.

  - *Section 9: LEO Resolution.* This is the location to documenting the primary LEO's affiliation (e.g., Federal Bureau of Investigation [FBI], Drug Enforcement Administration [DEA], local law enforcement) and, if appropriate, the reason for an arrest. This section also includes a text box for providing notes on LEO resolution.

Operational SPOT Referrals Database

TSA provided the research team with yearly databases of SPOT records from all operational locations. Data have been made available since 2006 (when the program was being piloted); thus, the analysis for the current study reflects five years of data.[3] The number of yearly cases varied as a function of program expansion as well as varying air travel patterns,[4] ranging from 25,000 to 99,000 records. The research team estimated that a total of 300,000 SPOT cases would be available in this period. The actual total number of cases from January 1, 2006, through October 31, 2010, was 247,630 referrals from 175 airports, which represents data available for this final report.

Database Elements

All records in the SPOT database are positive cases; that is, they represent all travelers stopped during SPOT Program observations. Information collected in these records includes indicators, outcomes, and non-personally identifying traveler and case information (e.g., checkpoint location, time/date, travel itinerary, BDOs involved). The fields contained in the database reflect four general categories of information:

- Background
- SPOT Referral Information
- Flight Information
- Database/Incident Documentation

Table B-1 lists the database fields along with a brief definition of each field.

[3] Data from 2006 represent the pilot testing of the program using a limited number of BDOs and locations.
[4] For example, the FAA estimated a 7.8% decrease in air travel from 2008 to 2009 and then 2% yearly increases in air travel beginning in 2010. Federal Aviation Administration. (2008). *FAA aerospace forecast fiscal years 2008–2025*. Washington, DC: Author.

Table B-1: Database Fields and Definitions

Variable	Definition
Background	
SPOT ID	A unique identification number assigned to each case
SPOT date and time	Date and time of SPOT referral
Airport code	3-letter designator to indicate airport at which SPOT is being conducted
Location	Airport area in which SPOT is being conducted (e.g., checkpoint, terminal)
Team members' names	First and last names of BDOs involved in referral
BDO responsible for initial referral	BDO role (e.g., checkpoint, playbook, gate screening) at time of initial SPOT referral
SPOT Referral Information	
Environmental baseline	(b)(3):49 U.S.C. § 114(r)
Behaviors observed	Documentation of behavior(s) observed for a traveler (from Section 2 of the SPOT Referral Report)
Reason for referral	Brief description of the reason(s) the traveler was referred for SPOT screening (b)(3):49 U.S.C. § 114(r)
Referral notes	Detailed notes to document the circumstances surrounding the traveler referral
Unusual items	(b)(3):49 U.S.C. § 114(r)
Signs of deception	Documentation of signs of deception observed during casual conversation (from Section 4 of the SPOT Referral Report)
Signs of deception notes	Detailed notes to document the signs of deception observed during the casual conversation
Subtotal and total points	(b)(3):49 U.S.C. § 114(r)
Automatic LEO notification	Documentation of traveler situation that met the threshold for automatic law enforcement officer (LEO) notification (b)(3):49 U.S.C. § 114(r) (b)(3):49 U.S.C. § 114(r)
Automatic LEO notification notes	Detailed notes to document the circumstances associated with the automatic LEO notification
Prohibited items—including what they are	Description of prohibited items observed with traveler (e.g., knives; liquids, gels, aerosols)
Prohibited items—how they were discovered	Reason for discovery of prohibited item(s) (e.g., bag search, X-ray, surrender by passenger prior to checkpoint)
Prohibited items—where they	Where the prohibited item was discovered (e.g., bag search, before checkpoint)

[5] Prior to the 2009 change in standard operating procedure (SOP), (b)(3):49 U.S.C. § 114(r) (b)(3):49 U.S.C. § 114(r)

Sensitive Security Information (SSI)

Variable	Definition
were discovered	
Prohibited item notes	Detailed notes to document the circumstances associated with the identification of the prohibited item(s)
Resolution	Culminating decision or action related to the referral (b)(3):49 U.S.C. § 114 (b)(3):49 U.S.C. § 114(r)
Resolution notes	Detailed notes to document the circumstances associated with the referral resolution
Primary LEO	Type of LEO involved in referral (e.g., local LEO, ATF, FBI)
Reason for arrest	Description of reason(s) behind LEO decision to arrest traveler involved in referral process (e.g., outstanding warrants, illegal alien, fraudulent documents)
Illegal alien indicator (traveler)	Yes/No designator to indicate whether or not the traveler is an illegal alien
Self-deporting indicator (traveler)	Yes/No designator to indicate whether or not the traveler is self-deporting
Screening manager notified	Name of screening manager notified if applicable.
LEO notification notes	Detailed notes to document the circumstances associated with the LEO referral/outcome of the referral.
Flight Information	
Air carrier	The airline on which the referred individual intends to travel
Flight number	The flight number designated for the flight on which the referred individual intends to travel
Origination airport code	3-letter designator to indicate airport where referred individual began his/her travel
Next destination airport code	3-letter designator to indicate the next airport where referred individual intends to travel
Final destination airport code	3-letter designator to indicate the airport where referred individual intends to complete his/her travel
Database/Incident Documentation	
Created by	Name of individual who initially populated the database
Created date and time	Date and time individual initially populated the database
Updated by	Name of individual who updated information already contained in the database
Updated day and time	Date and time of any updates to information already contained in the database

The database used does not include any personal identifiable information regarding travelers.

Changes to Operational Database

Note that beginning 1 March 2009, a new standard operating procedure (SOP) went into effect. This resulted in slight changes in indicators, scoring, and some aspects of the SPOT Referral Report. These changes are accounted for in the dataset, where necessary. The difference in the SOP did not change the instrument or items substantially; thus, we did not expect it to have a significant impact on the primary results of this study. Moreover, annual database analyses (reported elsewhere) show little change in patterns of referrals or case information from before versus after the SOP changes. In addition, the database system changed in March 2010. With the new system, certain fields allowed multiple response categories, which were not allowed in the previous data entry system.

Data Preparation

TSA provided AIR with data in either Excel or Access database format. Upon receipt, we transferred the data to SPSS for data preparation and analysis. Data preparation activities involved recoding data, collapsing and creating composite variables, and conducting data checks to ensure quality and accuracy.

Recoding Variables

The data provided by TSA came as string variables (i.e., alphanumeric characters that do not lend themselves to mathematical computation without prior transformation). Part of the data preparation activities involved creating SPSS syntax to transform string to numeric for variables of interest. To illustrate, the variable "illegal alien," in which a passenger is assigned a value of "yes" if an illegal alien and "no" if not an illegal alien, was recoded as 1 and 0, respectively. The numeric codes (1/0) were then labeled accordingly. The transformation to numeric codes allowed subsequent data transformation and analysis activities.

Collapsing Variables

A number of variables lent themselves to within-variable consolidation. For example, the field labeled "Resolution" describes how a SPOT referral is resolved for a particular case. Across cases, the resolution descriptions were readily collapsible into categories (e.g., LEO involved, Denied boarding by carrier, TSA involved). When relevant, we collapsed string variables into categories after recoding them as numeric variables.

Creating Predictor and Outcome Variables

Certain predictor and outcome variables of interest for the Operational SPOT study were derived from the data provided by TSA. When applicable, we created predictor (e.g., total score based on behaviors and unusual items) and outcome (e.g., prohibited items, arrested) variables for use in analysis. For example, TSA provided behavior indicator information in string variables: behavior1, behavior2, behavior3, through behavior8. We used this information to created yes/no indicator variables for every behavior (b)(3): by recoding the string variables into numeric versions (b1–b8), where behaviors were assigned identical values across all eight variables. Then, using programming loops, we were able to identify each specific instance of a particular behavior recorded in variables b1–b8 and assign the dichotomous variable BEH_X a value of 1, indicating the presence of this indicator. If a particular indicator was not observed during the referral, the corresponding BEH_X variable was coded 0, indicating that the indicator was not observed. As way of illustration, the indicator '(b)(3):49 U.S.C. § 114(r) ' became indicator variable BEH_1. First, we assigned every case of '(b)(3):49 U.S.C. § 114(r) ' in variables behavior1 through behavior8 a value of 1 in our numeric variables b1 through b8. Then, we used the looping function to identify every case of 1, or '(b)(3):49 U.S.C. § 114(r) (b)(3):49 in variables b1 through b8, and assigned it a value of 1 in the variable BEH_1. This would indicate that in that specific referral, the passenger (b)(3):49 U.S.C. § 114(r) . If the loop did not find the value 1 (in this example) in a referral case, BEH_1 would be coded 0, indicating that the passenger (b)(3):49 U.S.C. § 114(r)

Data Checks

Throughout the data preparation phase, we conducted a variety of data checks to ensure the accuracy of the data provided by TSA in addition to the new variables created for use in analysis. One data check involved running and comparing frequencies on both the original and recoded variable. Another involved examining the data on individual cases (i.e., case summaries) both before and after recoding, to ensure that the item recoded properly. We conducted checks of all data used for analysis (e.g., airport code checks, date checks, behavior checks, unusual item checks, point total checks). Other such consistency checks included airport (e.g., BDO name with airport should match the airport location listed in the airport field), resolution (e.g., the number of arrests listed under resolution should match the number of entries in the reason for arrest field), reason for referral, and automatic LEO notification.

Further, we checked for consistency across variables provided by TSA. For example, TSA provided a date and time variable. If this date was outside the range of the cases provided (e.g., the dataset spanned cases from March to July 2010 but the date and time variable was December 31, 2009), we noted it and contacted TSA for further explanation.

APPENDIX C:
BASE RATE STUDY METHODOLOGY

Introduction

Appendix C provides information about the Base Rate Study methodology and data preparation activities. First, we describe the study methodology, including information about the sampling, participants and the method used to determine sample size requirements and the specific data collection procedures. Next, we describe the database developed by the Transportation Security Administration (TSA) to contain the Base Rate Study records, as well as summary data collected by the data collection teams. Finally, we present an overview of the data preparation procedures conducted by AIR to prepare data for analysis.

Method

This section details key elements of the method employed for the Base Rate Study. It includes information on study participants, sample size requirements, and data collection procedures.

Sampling Plan

There were two goals for the SPOT Base Rate Study. One was to estimate the base rate of each of the SPOT outcomes among all persons traveling[6] through the 166 US airports where SPOT is in operation; the other was to compare the efficacy of SPOT screening with that of a random screening approach to detecting these outcomes. Both these objectives involved collecting occurrence data from a sample of travelers and combining these collected data to form estimates of population parameters. In the first case, the estimates would be of the occurrence rates within the total population of travelers through airports where SPOT is in operation,[7] while in the second, the estimates would be constructed to match the segments of travelers covered by the SPOT program. In the current study, outcome data were equally weighted in any estimate produced, whether constructing base rates or SPOT comparisons. In this section we discuss this estimation approach in light of the way the sample units were actually selected for the study. For the sake of simplicity, we write in terms of the base rate estimates. However, most of the considerations would be similar for the SPOT Comparison estimates as well, except that in the latter case, the estimates would be constructed to represent only the airport checkpoints and time periods where and when SPOT is routinely in effect.

Ideal Sampling Strategy

For the SPOT base rate estimation goal, an ideal sampling strategy in the sense of minimum sampling variance and bias, without regard to implementation cost or operational constraints, would be as follows. From the population of all N travelers passing through any of the 166 airports where SPOT is in place, at any time during the target months of the study, randomly

[6] A person who takes more than one flight during this period is counted as a separate flight traveler for each one-way trip.

[7] TSA has reported that the set of airports where the SPOT Program is in effect covers 96% of the total U.S. traveling population. Unpublished data (TSA, 2010).

Possession of Prohibited/Illegal Items, and/or Possession of Fraudulent Documents. Then the base count estimate for an outcome would be:

$$\widehat{BC} = \sum_{k=1}^{n} w_k \chi_k$$

where $w_k = 1/p_k$ and $\chi_k = \begin{cases} 1, \text{if the traveler is of interest with respect to the outcome} \\ 0, \text{otherwise} \end{cases}$

The base rate estimate for the outcome is $\widehat{BR} = \widehat{BC}/N$. If N itself is an estimate, then $\widehat{BR} = \widehat{BC}/\widehat{N}$. (This second estimate would have higher sampling variance, accounting for the variance of $\widehat{N}$).

Actual Strategy

This ideal selection strategy, described above, was approximated by implementing a number of steps guided by practical considerations for airport operations. These steps included selecting clusters of travelers by choosing a sample of airports, and by designating time periods and locations within the airports for data collection. AIR organized the airports into strata from a sampling frame of 143 of the 166 current SPOT airports in April 2009 (i.e., at the time when the sampling frame was constructed); this excluded the 23 airports considered as "non-hub primary." The 143 airports were grouped into three explicit strata based on airports' total annual enplanements, according to FAA definitions. Airports were implicitly stratified within these three strata based on throughput and arrest rates. AIR made recommendations about strata allocations (i.e., the proportion of airports that should be selected from each). TSA, with input from local airport facility management selected a sample of airports accordingly. Using the number of airports in the sample and in the frame from each strata, and assuming each airport in a strata had the same chance of being in the sample as any other, the chance of selection for each airport would be the ratio of the two.

Once the airports were selected for inclusion in the study, their managers specified checkpoints as collection sites, as well as time periods during airport operating hours for the collections. These managers were guided by AIR to conduct data collections at times and in locations representative of (i.e., typical for) Operational SPOT. If random selections were spread across all possible checkpoints and time periods proportional to the throughput for those sites and times, then all passengers selected at this stage for an airport would represent the same number of that airport's travelers, and it would be reasonable to assign their data equal weight in the estimates for this stage of sampling. Based on our examination of the Base Rate Study data in comparison to the Operational SPOT data, it appears that data collection times and locations in the two datasets are approximately the same. That is, there were no systematic differences in times or locations where the Base Rate Study data collections occurred as compared to typical Operational SPOT observations. If the results of this limited review[3] were to hold up through an

[3] This analysis of sampling was limited by data availability as noted throughout this section.

in-depth analysis with the complete set of population data, it would further support equally-weighting the outcomes for this stage of sampling.

Data analysis did assume equal sampling weights for each case in the datasets. Because the chances of selections for these sample units at each stage (i.e., airports within strata; locations; and time slots) could not be known precisely given the data available for the Base Rate Study, exact computation of sampling weights was not possible. While TSA could work to get the required data in order for the research team to compute and apply sampling weights, there were several difficulties in doing so that made this analysis beyond the scope of this study. Moreover, considering that the airports were essentially selected proportional to their throughput, and as just described, using the available data, the proportion of collections within airports seem to match the actual throughput across checkpoints and time periods, equal weighting is a reasonable approach to estimation.[9]

Resulting Airports Selected for Base Rate Study

Figure C-1 lists the 43 airports included in the Base Rate Study from 10 September 2009 through 31 October 2010.[10] In September 2009, 24 airports started data collection; 1 airport began data collection on 1 March 2010; and 18 airports were added on 16 July 2010.

Figure C-1: Selected Airports

Airport	Hub Size
Start Date: 10 September 2009	
Atlantic City International Airport (ACY)	Small
Baltimore-Washington International Thurgood Marshall Airport (BWI)	Large
Bob Hope Airport (BUR)	Medium
Cincinnati/Northern Kentucky International Airport (CVG)	Large
Denver International Airport (DEN)	Large
Des Moines International Airport (DSM)	Small
Detroit Metropolitan Wayne County Airport (DTW)	Large
Honolulu International Airport (HNL)	Large
Kona International Airport (KOA)	Small
Lambert-St. Louis International Airport (STL)	Medium
Los Angeles International Airport (LAX)	Large
Memphis International Airport (MEM)	Medium
Newark Liberty International Airport (EWR)	Large
Oakland International Airport (OAK)	Medium
Orlando International Airport (MCO)	Large

[9] Moreover, it is AIR's assessment as well as the assessment of the external review panel that if sampling weights were applied to the data, the primary findings of this study would remain unchanged.
[10] A pilot study was conducted from September through November 2009 using these airports. The primary study was conducted from 01 December 2009 through 31 October 2010.

Airport	Hub Size
Philadelphia International Airport (PHL)	Large
Phoenix Sky Harbor International Airport (PHX)	Large
Pittsburgh International Airport (PIT)	Medium
Portland International Airport (PDX)	Medium
San Diego International Airport (SAN)	Large
Tampa International Airport (TPA)	Large
Tucson International Airport (TUS)	Medium
William P. Hobby Airport (HOU)	Medium
Start Date: 1 March 2010	
Logan International Airport (BOS)	Large
Start Date: 16 July 2010	
Austin-Bergstrom International Airport (AUS)	Medium
Buffalo Niagara International Airport (BUF)	Medium
Charlotte/Douglas International Airport (CLT)	Large
Chicago Midway International Airport (MDW)	Large
Fort Lauderdale-Hollywood International Airport (FLL)	Large
Indianapolis International Airport (IND)	Medium
John F Kennedy International Airport (JFK)	Large
Kahului Airport (OGG)	Medium
La Guardia Airport (LGA)	Large
Louisville International Airport (SDF)	Medium
McCarran International Airport (LAS)	Large
Miami International Airport (MIA)	Large
Minneapolis-Saint Paul International Airport (MSP)	Large
Port Columbus International Airport (CMH)	Medium
Ronald Reagan Washington National Airport (DCA)	Large
Salt Lake City International Airport (SLC)	Large
Southwest Florida International Airport (RSW)	Medium
Ted Stevens Anchorage International Airport (ANC)	Medium

Sample Size Requirements

The research team considered sample size requirements at the level of randomly selected travelers to ensure that the study would provide sufficient data to make inferences about aviation risk in the traveling population. Initial sample size estimation was based on two criteria: (1) with a 95% level of confidence, the study should be able to estimate the incidence of high-risk travelers in the traveling population within a reasonable distance from the actual population parameter; and (2) there should be a 90% chance that the number of positive outcomes (cases that involved a Law Enforcement Official [LEO]) would be at least 20 to ensure that results are not spurious and to provide stable estimates for further analyses.

To satisfy the first criterion, we used certain assumptions about the traveling population based on information from existing SPOT Program data. We assumed that each traveler either was of

interest or was not; no intermediary "level" of interest was considered. Further, we assumed that independent of each other, the travelers had a fixed probability, p, of being of interest. Under these assumptions, the total number of persons of interest has a binomial distribution with parameter p. The number of persons of interest in a sample selected from this population would also have a binomial distribution.[11]

Constructing a confidence interval for a binomial parameter is a basic and important matter in statistical practice. A very common approach, known as deriving the Wald confidence interval for p, is to use a normal distribution to approximate the binomial. This approximation has generally been considered appropriate if the sample size is large enough, say, choosing n so that $np(1-p) > 9$ (were p known).[12] It remained to consider a reasonable, yet conservative, value for p to use in this formula. Lower values of p require higher sample sizes to satisfy the normal approximation requirements, so a conservative conjecture would be a low value. From existing SPOT data, we know that (b) of the cases in which travelers are stopped by the program lead to LEO referrals. Presuming that the SPOT program is better than no screening or random screening, and seeking this conservative value of p, we supposed that only (b)(3):49 of the cases in which travelers were not stopped would have led to LEO referrals. Weighting these incidence rates (b) of those stopped and (b) of those not stopped) by the percentage of the population they refer to (b)(3):4 stopped a (b)(3):49 t stopped) gave us an estimate of (b) for p for the entire population.[13] Using this value in $np(1-p) > 9$ led to a minimum sample size of 45,032.

To satisfy the second of the two criteria (a 90% chance of getting at least 20 LEO-involved cases), we considered various sample sizes, aiming to keep the chance of getting fewer than 20 positives less than 10%. Computing this chance can be performed exactly for a binomial distribution, and using the conjectured p = 0.0002, we estimated that the required sample size was 130,065 travelers. Clearly, this second criterion was the more demanding one, and satisfying it would also satisfy the first. Although this second criterion led to a much greater sample size, as stated above, we felt that it was necessary to support the goals of the study and to lead to stable estimates of the low base rate outcome of interest.

An initial 10-week pilot study was conducted from September through December 2009, in which the rate of LEO referrals for a random selection of travelers was between (b)(3):49 U.S.C. § 114(r) higher rate than the initial parameter estimate of (b) Considering the lower end of the incident rate (0.05%) and using the second criterion for sample size, the adjusted sample size requirement was reduced to 52,000 cases. This is expected to be more than ample to make reasonable inferences about high-risk travelers.[14]

---

[11] Technically, this is true if sampling is done with replacement, but given the large population and a large enough sample size, it is very nearly true even when sampling is done without replacement, as in the current study.

[12] Brownlee, K. A. (1965). *Statistical theory and methodology in science and engineering*. New York: John Wiley & Sons.

[13] Assumptions are based on Operational SPOT analyses (AIR, 2008, 2009) and SPOT Program manager reports (personal communication with TSA, 12/01/2008).

[14] We note that because the proportion of high-risk travelers is so low, the distribution of the sample estimate should really be estimated by a Poisson distribution. Using such an approximation and the currently recommended sample

Procedure

Data Collection

Data collection teams comprised two SPOT Behavior Detection Officers (BDOs) and one Data Collection Coordinator (DCC). At each data collection location (i.e., checkpoints within airports), a specific and predetermined marker was designated as the point where random selection occurred. This selection marker was positioned as close to the entrance of the checkpoint queue as possible so that the selectee was observed during the entire period that they were in the queue. When the data collection team was ready to select a traveler, the next traveler to cross this marker was randomly selected for the study. If multiple travelers crossed the marker simultaneously, then the team was directed to select the traveler on the left. Each randomly selected traveler was observed and scored for each SPOT indicator. Regardless of whether the traveler's score exceeded the threshold, every randomly selected traveler was next directed to the SPOT Referral Screening so that outcomes could be obtained. Once the referral was completed, the BDO team filled out the SPOT Referral Report paperwork and returned to their initial positions at the checkpoint queue. As soon as the team was in position, the next traveler to cross the designated marker was randomly selected.

The random selection method was designed to be as simple as possible in order to assure standardization across data collection teams and locations, avoid bias in selection of travelers, and assure that selections were indeed random. While there may be other methods for random selection (e.g., using a random number generator in the process), the goal here was to minimize complications that could lead to variations across sites. The rule of selecting the individual on the left when multiple travelers crossed the marker simultaneously served to further minimize bias in selection (i.e., to not allow BDOs to decide which person to choose). This sampling method also allowed for variations in time for completion of a case. That is, the length of observation was dictated by the length and speed of the checkpoint queue.[15] Thus, data collection teams were not pressured by a time sampling method (i.e., selection every X minutes) that would have lead to sampling error when observations exceeded the designated sampling period.

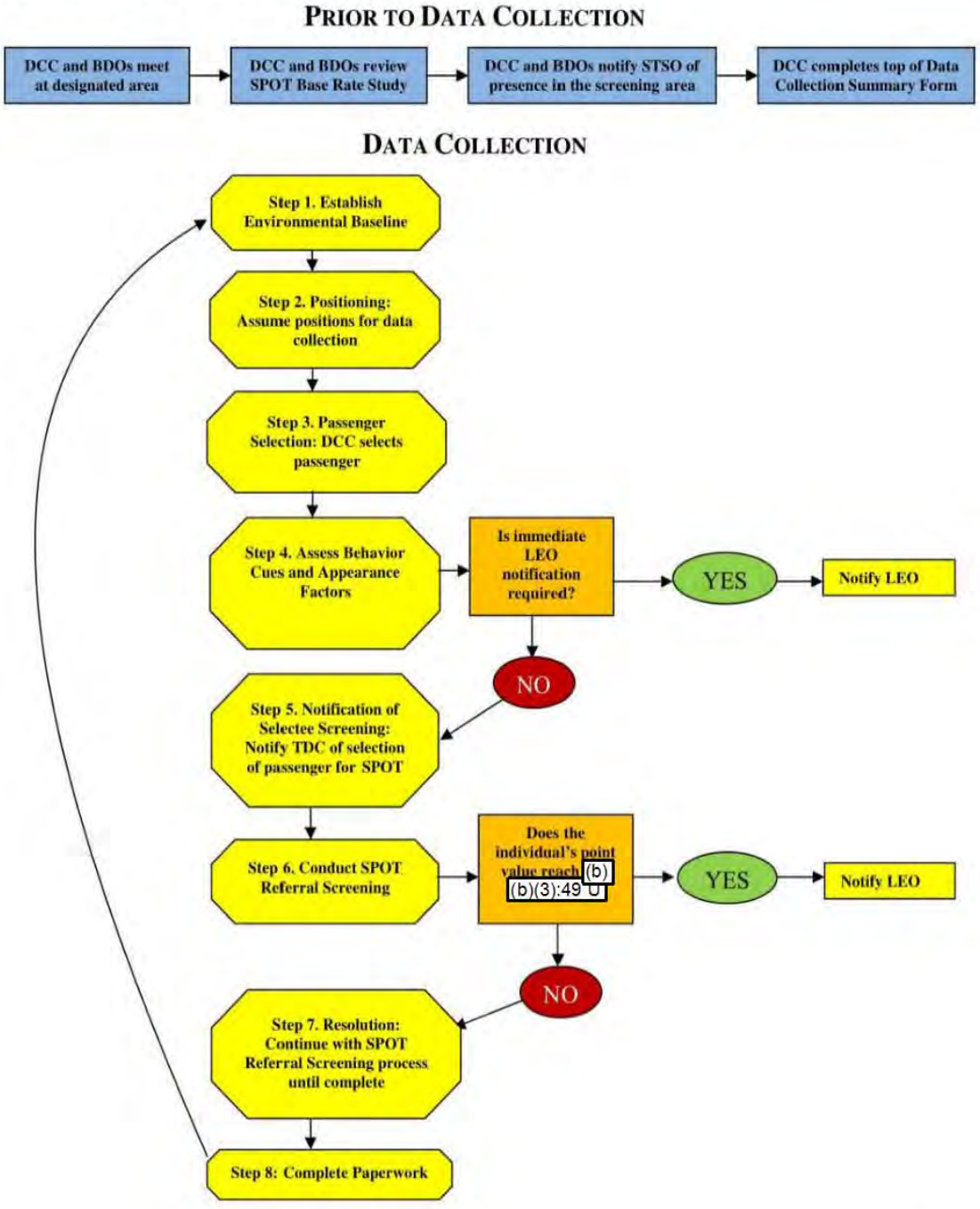
Figure C-2 provides a flow chart of the data collection process. The data collection procedures were designed to mirror standard operational procedures for SPOT observation, screening, and completion of the SPOT Referral Report in order to allow for direct comparisons between SPOT and random screening in the study analyses.

---

size of 52,000, if the estimated proportion is about 0.05%, the width of the 95% confidence interval for the estimate will be about 0.04 percentage points.

[15] This is no different than Operational SPOT, where the length of time the BDO has to observe a traveler in line is dictated by the length and speed of the checkpoint queue.

Figure C-2: Base Rate Study Data Collection Process



Below we provide a detailed description of each step in the data collection process. An asterisk (*) denotes data collection procedures that deviate from the SPOT Standard Operating Procedures (SOP).

1. **DCC and BDOs meet at designated area:** On the day of data collection, and prior to the start of data collection, the BDOs and DCC met at a designated area and confirmed that they had the necessary materials for data collection. These included all materials required of BDOs (as specified in the SPOT SOP), in addition to data collection paperwork:

   Required by SPOT SOP:

   - A TSA-issued cell phone or working radio
   - A working black light and magnification loupe, if available
   - A writing instrument and a notebook
   - A set of SPOT Assist Cards[16]
   - A partner, unless directed otherwise by the SPOT Transportation Security Manager (STSM) or SPOT Coordinator

   Data collection materials:*

   - Paper copies of SPOT Checklist (to be completed by BDOs following each observation)
   - Data Collection Summary Form (to be completed by the DCC)
   - Pens

2. **DCC and BDOs review SPOT Base Rate Study data collection plan:*** The BDOs and the DCC reviewed the SPOT Base Rate Study Data Collection Plan[17] and ensured that everyone understood the specifics of the data collection for the day. This included discussion of:

   - positioning of the DCC and the BDOs,
   - communication strategies to be used between the DCC and BDOs,
   - start time of the data collection, and
   - selection position.

[16] While data collection was under way, the SOP was modified to eliminate the use of SPOT Assist Cards. It is therefore not required as part of materials required for the study.
[17] The Data Collection Plan provided step-by-step instructions on how to conduct the data collection.

3. **DCC and BDOs notify Supervisory Transportation Security Officer (STSO) of presence in the screening area:** The BDOs and the DCC identified themselves to the Travel and Document Checker (TDC) and the Screening Checkpoint STSO and discussed any operational concerns prior to beginning data collection (as specified in the SPOT SOP).

4. **DCC completes top of Data Collection Summary Form:** * The DCC filled in the top section of the Data Collection Summary Form, indicating the current data collection day, time, location, and personnel.

   During data collection, the BDOs used the standard SPOT Program process, outlined in Figure C-3, with an adaptation related to the random selection of travelers for the Base Rate Study and the completion of study paperwork.

Figure C-3: Standard SPOT Process

A. Establish an environmental baseline.
B. Visually assess individuals for SPOT Behavior Cues and Appearance Factors.
C. Conduct casual conversation.
D. Conduct SPOT Behavior Detection and Analysis.
E. Notify a LEO and others, as required.
F. Conduct SPOT Secondary Screening.
G. Complete the SPOT screening process by resolving SPOT behaviors, or requesting LEO assistance.
H. Complete all appropriate paperwork.

5. **Establish environmental baseline:** The data collection team established the environmental baseline by determining the typical behaviors and appearance that would reasonably be expected at the time and place SPOT was being conducted.

6. **Assume positions for data collection:** The DCC and the BDOs assumed their positions.

(b)(3):49 U.S.C. § 114(r)

7. **DCC selects the first passenger:*** Prior to the start of data collection, the BDOs and the DCC discussed the start time of the data collection (see Step 1, item 2c). For example, if their shift started at 7 a.m., the team would have decided ahead of time to start the data collection promptly at 7:15 a.m. At the designated start time, the DCC selected the first passenger who passed the selection position/marker. This position was identified in a way that was unobtrusive to the checkpoint (e.g., a tape mark on the floor, the entrance to the queue, the position of the second stanchion in the queue). In addition, the selection marker was placed as close to the entrance of the checkpoint queue as possible so that the selectee was observed during the entire period that they are in the queue. The DCC communicated the selected passenger to the BDOs, using the agreed-on communication strategy.

8. **Assess behavior cues and appearance factors:** BDO 1 continued to observe the queue, with particular attention to completing the SPOT Referral Report on the selected passenger. He or she observed this selected passenger, and any traveling companions, in the same way he or she would if this person had been selected by SPOT. The goal of this observation was to complete Section 2 of the SPOT Referral Report while the passenger was in line.

9. **Notify TDC of selection of passenger for SPOT:** As the passenger approached the TDC, BDO 2 discreetly notified the TDC that the passenger had been selected for screening, as specified in the SPOT SOP. When the passenger was at the TDC, the TDC referred the individual (and any traveling companions) to SPOT Secondary Screening (also referred to as SPOT Referral Screening), as specified in the SPOT SOP.

10. **Conduct SPOT Referral Screening:** BDO 1 and BDO 2 conducted a SPOT Secondary Screening using the established SPOT procedures.

11. **Complete paperwork:*** Once a resolution was reached, the BDOs completed a paper copy of the SPOT Referral Report. On the checklist, the BDOs indicated which behaviors were observed while the passenger was in line/interacting with the TDC and which behaviors were observed during the SPOT Referral Screening. Once the paperwork was complete and placed in a secure location, they returned to their positions and communicated their return to the DCC.

12. **Select next traveler:*** As soon as the BDOs returned to their positions, the DCC selected the next passenger who passed the selection position/marker. Data collection continued in this manner until the predetermined stop time.

Base Rate Study Procedural Differences from SPOT SOP

As noted above, there were some small differences between the Base Rate Study data collection and the SPOT SOP that were necessary for the study protocol to be standardized across locations. These are summarized here:

1. In addition to the typical system of two BDOs observing the line, a third BDO served as the DCC. The DCC randomly selected travelers, tracked the daily data collection, and reported any incidents that required suspending the data collection.
2. At the start of each data collection period, the BDO team designated a selection position or marker. This naturally varied by location because checkpoints have varied layouts. Regardless, the traveler selected as the first study participant was the first one to cross the designated marker when data collection started. Each subsequently selected traveler was the next to cross the marker once the BDO team had completed the data collection for the previous randomly selected traveler and had repositioned at the checkpoint.
3. The selected traveler was observed in line and then selected for SPOT Referral Screening, regardless of the SPOT observation score. Then, the screening followed the usual procedure to obtain a SPOT score and case outcomes.
4. The BDOs completed the paper-based SPOT Referral Report as soon as the SPOT Referral Screening was complete. This timing minimized errors in data collection. The BDOs noted on the Referral Report (and subsequently in the database) where the various indicators were observed. There were four possible locations for observation: Travel Document Checker (TDC) Queue, Divest Area, Walkthrough Metal Detector (WTMD) to Referral Screening Area, and Referral Screening.
5. If, during the data collection period, the BDO team observed another traveler exhibiting behaviors that warranted further observation, the BDOs stopped data collection to address the potential security issue. The DCC noted this break from protocol on the Data Collection Summary Form. Figure C-4 provides examples of unusual scenarios that could arise and could affect the data collection, procedures that were developed to address these potential issues, and information on where these data were entered into the databases.

Figure C-4: Protocol Deviation Procedures

Issue	Mitigation Plan	Data Entry
Fits Data Collection Protocol		
Result of observation: Passenger is high risk.	Proceed as required by SPOT SOP. Once case is resolved, complete SPOT Referral Report paperwork.	Data entered into the Performance Management Information System (PMIS)[18]. Data entered into Base Rate Study Database
Result of SPOT Referral Screening: Passenger is high risk.	Proceed as required by SPOT SOP. Once case is resolved, complete SPOT Referral Report paperwork.	Data entered into PMIS Data entered into Base Rate Study Database
Selected passenger is traveling with companions.	Follow data collection procedure, and observe passenger. Follow SPOT SOP for handling companions. Complete data collection paperwork for the whole party (as in SPOT SOP).	Data entered into Base Rate Study Database
Deviation		
Observation of passenger in queue (not part of random selection) exhibiting anomalous behavior	BDO will indicate break in data collection. Passenger is observed in accordance with SPOT SOP. DCC will complete Protocol Deviation Form.	Data entered into PMIS

Data Collection Training

Prior to the start of data collection, AIR trained BDOs who were designated to serve as DCCs on the data collection protocol and their role as DCC. Designed as a half-day session, the training included classroom time during which an AIR team member presented an overview of the Base Rate Study. This overview included a discussion of the importance of the study; prior analyses of SPOT incident data; why the BDOs were selected for participation; their role in supervising the data collection process; and data collection procedures.

The overview of data collection procedures specifically included the following elements:

- Data collection logistics, including the composition and roles of the data collection team
- Data collection procedures, including what to do prior to, during, and after data collection
- Contingency planning, (b)(3):49 U.S.C. § 114(r)
  (b)(3):49
- Data collection errors or key activities that will make up the Base Rate Study outcomes (e.g., falsifying data, not following procedures).

[18] Performance Management Information System (PMIS) is the Operational SPOT database.

During the training, the AIR team reinforced the importance of following the random selection procedures as outlined in the protocol, and although variations in checkpoint layout and throughput would require such things as location of the marker to differ, the specific procedures for random selection should remain standardized across locations. The AIR team emphasized that deviating from the prescribed protocol could increase the likelihood of introducing systematic bias into the study and limiting the results.

The training concluded with a practice data collection session at a checkpoint. During this practice session, participating BDOs divided into data collection teams and conducted a number of data collection rounds with actual travelers. The BDOs took turns serving in the DCC function. AIR team members remained on the floor with the BDOs to observe the practice session, provide feedback to each team, help troubleshoot airport-specific issues or concerns, and answer any questions. Following the practice data collection session, the participating BDOs and the AIR team debriefed and reviewed key points.

AIR conducted data monitoring visits to at least half of the airports in the study. These monitoring visits were conducted to ensure that all data collection procedures for the SPOT Base Rate Study were being implemented in a standardized manner throughout all participating airports. To this end, the AIR team observed the data collection process to assess the integrity of data collection, provided feedback, and answered questions. (See Appendix D for more information on data monitoring visits).

Data Preparation

This section provides details of the elements contained in the Base Rate Study database as well as AIR's data preparation activities.

Base Rate Study Database Elements

TSA provided AIR with a database of Base Rate Study data from all study locations. The database included data from 1 December 2009 through 31 October 2010.[19] All records in the Base Rate Study database are cases that represent the travelers selected randomly during the data collection. Information collected in these records includes indicators, outcomes, and non-personally identifying traveler and case information. The total number of cases in the database for this time period is 71,589, which represent the data available for this report. Figure C-5 presents the Base Rate Study database fields with a definition for each field.

[19] An initial 10-week pilot study was conducted between 10 September 2009 and 30 November 2009. The database containing the records for that time period was provided to AIR in December 2009 for initial review and reassessment of sample size requirements.

Figure C-5: Database Fields and Definitions

Variable	Definition
Background	
SPOT ID	A unique identification number assigned to each case
SPOT date and time	Date and time of SPOT referral
Airport code	3-letter designator to indicate airport at which SPOT is being conducted
Location	Airport area in which SPOT is being conducted (e.g., checkpoint, terminal)
Team members' names	First and last names of BDOs involved in referral
BDO responsible for initial referral	BDO role (e.g., checkpoint, playbook, gate screening) at time of initial SPOT referral
SPOT Referral Information	
Environmental baseline	(b)(3):49 U.S.C. § 114(r)
Behaviors observed	Documentation of behavior(s) observed for a traveler (from Section 2 of the SPOT Referral Report)
Reason for referral	Brief description of the reason(s) the traveler was referred for SPOT screening (b)(3):49 U.S.C. § 114(r)
Behavior observed at	Documentation of where the behavior (s) was observed (b)(3):49 (b)(3):49 U.S.C. § 114(r)
Referral notes	Detailed notes to document the circumstances surrounding the traveler referral
Unusual items	(b)(3):49 U.S.C. § 114(r)
Signs of deception	Documentation of signs of deception observed during casual conversation (from Section 4 of the SPOT Referral Report)
Deception observed at	Documentation of where the sign (s) of deception were observed (b)(3):49 (b)(3):49 U.S.C. § 114(r)
Signs of deception notes	Detailed notes to document the signs of deception observed during the casual conversation
Subtotal and total points	(b)(3):49 U.S.C. § 114(r)
Automatic LEO notification	Documentation of traveler situation that met the threshold for automatic law enforcement officer (LEO) notification (b)(3):49 U.S.C. § 114(r) (b)(3):49 U.S.C. § 114(r)
Automatic LEO notification notes	Detailed notes to document the circumstances associated with the automatic LEO notification
Prohibited items—including what they are	Description of prohibited items observed with traveler (e.g., knives; liquids, gels, aerosols)
Prohibited items—how they	Reason for discovery of prohibited item(s) (e.g., bag search, X-ray, surrender by

Variable	Definition
were discovered	passenger prior to checkpoint)
Prohibited items—where they were discovered	Where the prohibited item was discovered (e.g., bag search, before checkpoint)
Prohibited item notes	Detailed notes to document the circumstances associated with the identification of the prohibited item(s)
Resolution	Culminating decision or action related to the referral (b)(3):49 U.S.C. § 114((b)(3):49 U.S.C. § 114(r)
Resolution notes	Detailed notes to document the circumstances associated with the referral resolution
Primary LEO	Type of LEO involved in referral (e.g., local LEO, ATF, FBI)
Reason for arrest	Description of reason(s) behind LEO decision to arrest traveler involved in referral process (e.g., outstanding warrants, illegal alien, fraudulent documents)
Illegal alien indicator (traveler)	Yes/No designator to indicate whether or not the traveler is an illegal alien
Self-deporting indicator (traveler)	Yes/No designator to indicate whether or not the traveler is self-deporting
Screening manager notified	Name of screening manager notified if applicable
LEO notification notes	Detailed notes to document the circumstances associated with the LEO referral/outcome of the referral
Flight Information	
Air carrier	The airline on which the referred individual intends to travel
Flight number	The flight number designated for the flight on which the referred individual intends to travel
Origination airport code	3-letter designator to indicate airport where referred individual began his/her travel
Next destination airport code	3-letter designator to indicate the next airport where referred individual intends to travel
Final destination airport code	3-letter designator to indicate the airport where referred individual intends to complete his/her travel
Database/Incident Documentation	
Created by	Name of individual who initially populated the database
Created date and time	Date and time individual initially populated the database
Updated by	Name of individual who updated information already contained in the database
Updated day and time	Date and time of any updates to information already contained in the database

In addition to the Base Rate Study records, traveler throughput information was obtained for each data collection period and location. A separate database with this information, which is routinely collected at checkpoints, was provided by TSA for use in the analysis.

Data Preparation

TSA provided the research team with a database of Base Rate Study records from all study locations for the period 1 December 2009 through 31 October 2010. TSA provided the data in either Excel or Access database format. Upon receipt, we transferred the data to SPSS for data preparation and analysis. Data preparation activities involved recoding data, collapsing and creating composite variables, and conducting data checks to ensure quality and accuracy.

Recoding Variables

The data provided by TSA came as string variables (i.e., alphanumeric characters that do not lend themselves to mathematical computation without prior transformation). Part of the data preparation activities involved creating SPSS syntax to transform string to numeric for variables of interest. To illustrate, the variable "illegal alien," in which a passenger is assigned a value of "yes" if an illegal alien and "no" if not an illegal alien, was recoded as 1 and 0, respectively. The numeric codes (1/0) were then labeled accordingly. The transformation to numeric codes allowed subsequent data transformation and analysis activities.

Collapsing Variables

A number of variables lent themselves to within-variable consolidation. For example, the field labeled "Resolution" describes how a SPOT referral is resolved for a particular case. Across cases, the resolution descriptions were readily collapsible into categories (e.g., LEO involved, Denied boarding by carrier, TSA involved). When relevant, we collapsed string variables into categories after recoding them as numeric variables.

Creating Predictor and Outcome Variables

The predictor and outcome variables of interest for the Validation Study were derived from the data provided by TSA. When applicable, we created predictor (e.g., total score based on behaviors and unusual items) and outcome (e.g., *LEO Arrest, Possession of Serious Prohibited/Illegal Items*) variables for use in the analysis. For example, TSA provided behavior indicator information in string variables: behavior1, behavior2, behavior3, through behavior8. We used this information to create yes/no indicator variables for every behavior (b) by recoding the string variables into numeric versions (b1–b8), where behaviors were assigned identical values across all eight variables. Then, using programming loops, we were able to identify each specific instance of a particular behavior recorded in variables b1–b8 and assign the dichotomous variable BEH_X a value of 1, indicating the presence of this indicator. If a particular indicator was not observed during the referral, the corresponding BEH_X variable was coded 0, indicating that the indicator was not observed. As way of illustration, the indicator (b)(3):49 U.S.C. § 114(r) indicator variable BEH_1. First, we assigned every case of (b)(3):49 U.S.C. § 114(r) in variables behavior1 through behavior8 a value of 1 in our numeric variables b1 through b8. Then, we used the looping function to identify every case of 1, or (b)(3):49 U.S.C. § 114(r) in variables b1 through b8, and assigned it a value of 1 in the variable BEH_1. This would indicate that in that specific referral, the passenger (b)(3):49 U.S.C. § 114(r) If the loop did not find the value 1 (in this example) in a referral case, BEH_1 would be coded 0, indicating that the passenger (b)(3):49 U.S.C. § 114(r) (b)(3):

Data Checks

Throughout the data preparation phase, we conducted a variety of data checks to ensure the accuracy of the data provided by TSA in addition to the new variables created for use in analysis. One data check involved running and comparing frequencies on both the original and recoded variable. Another involved examining the data on individual cases (i.e., case summaries) both before and after recoding to ensure that the item recoded properly. We conducted checks of all

data used for analysis (e.g., airport code checks, date checks, behavior checks, unusual item checks, point total checks). Further, we checked for consistency across variables provided by TSA. For example, TSA provided a date and time variable. If this date was outside the range of the cases provided (e.g., the dataset spanned cases from March to July 2010, but the date and time variable was December 31, 2009), we noted it and contacted TSA for further explanation.

Other such consistency checks included airport (e.g., BDO name with airport should match the airport location listed in the airport field), resolution (e.g., the number of arrests listed under resolution should match the number of entrees in the reason for arrest field), reason for referral, and automatic LEO notification.

Data Collection Summary Forms

The Data Collection Summary Form was completed by the DCC at the start of every data collection session. The fields to be completed on the form follow:

- Date
- Airport and screening location
- Team member names
- Start time for the data collection
- End time for the data collection
- Total number of passengers screened

In addition, the Data Collection Summary Form provided space for the team to note whether any incidents occurred during the data collection that required the team to deviate from the data collection protocol. The form included prompts to provide information on

- the particular incident,
- the time it was identified, and
- how it was resolved.

The DCCs were provided with specific instructions on how to fill in these forms accurately and completely. A copy of the Data Collection Summary Form is provided below.

Figure C-6: Data Collection Summary Form

Sensitive Security Information (SSI)

SPOT Base Rate Study
Data Collection Summary Form

Instructions: This form is to be completed by the Data Collection Coordinator (DCC) for each day of data collection. Please fill in the information as completely as you can.

Date: _____ (mm/dd/yyyy)

Airport Location: _____ **Screening Location:** _____

Data Collection Coordinator (Last, First): _____

SPOT Team Member 1 (Last, First): _____

SPOT Team Member 2 (Last, First): _____

Start Time for Data Collection: _____ **End Time for Data Collection:** _____

Total Number of Passengers Selected/Screened: _____

Protocol Deviation:
Please indicate any incidents that occurred during data collection that required the team to deviate from the data collection protocol. Describe the particular incident(s), the time it was identified, and how it was resolved. Make sure you indicate time data collection was resumed.

Incident Description:

Time Incident Noted:

Resolution:

Time Data Collection Resumed:

Incident Description:

Time Incident Noted:

Resolution:

Time Data Collection Resumed:

Sensitive Security Information (SSI) *Please continue on back of page if necessary.*

Sensitive Security Information (SSI)

Data Preparation

Each airport location was asked to provide AIR with printed copies of the Data Collection Summary Forms as well as the completed SPOT Referral Reports for the Base Rate Study. Upon receipt of these forms, the research team reviewed the information for general completeness and entered the data into an Excel database. AIR encountered various issues when entering the data from the Data Collection Summary Forms. Common issues and steps taken to resolve these are listed in Figure C-6.

Figure C-6: Common Data Preparation Issues

Issue	Action
Screening location was changed midway through data collection session but the DCC did not start a new Data Collection Summary Form (per established procedures).	Entered the multiple screening locations as written on the Form in the Screening Location field and entered "Multiple" in the Throughput field.
Data Collection Summary Forms were completed and sent to AIR when data collection did not take place (i.e., DCC started to complete form but zero passengers were selected for screening because BDO team had to stop data collection).	Removed affected Data Collection Summary Forms with zero selected passengers from the data set.
End time was missing from Data Collection Summary Form.	Entered an approximate end time by using the time of the last Referral Report completed during the data collection session. In instances where the Referral Reports could not be located, AIR referred to the Base Rate Study database to ascertain this approximate end time (also from the last Referral Report of the session). In a few instances, records of completed Referral Reports for a particular data collection session could not be located (in hard copy or as entries in the database). These end times were entered as "Missing."
Screening location was missing from Data Collection Summary Form.	Entered screening location based on location information printed on corresponding Referral Reports. In a few instances, hard-copy Referral Reports could not be located. These screening locations were entered as "Missing."
Number of passengers selected/screened was missing from Data Collection Summary Form.	Entered number of passengers based on the number of Referral Reports that correspond to the Data Collection Summary Form. The Base Rate Study database was also referenced to obtain the number of passengers screened during a specific window of time.
Two separate start and end times were recorded on a single Data Collection Summary Form.	Entered first and second sets of start and end times as two separate entries (as if two Data Collection Summary Forms had been completed). To determine the number of passengers screened during each of the two sessions (as opposed to the total number for both sessions, as indicated on the Form), AIR referenced the corresponding Referral Reports and/or entries in the Base Rate Study database.

In instances where information was missing from the Data Collection Summary Form but AIR was able to obtain these data elsewhere (e.g., Referral Reports, Base Rate Study database), handwritten notes were made on the affected form to indicate the source of the supplementary data.

Calculating Throughput

As noted in Volume 1, Chapter 6, the calculation of throughput was critical to calculating classification accuracy for random selection in the base rate study and for Operational SPOT selection methods. Throughput was defined as the total number of persons passing into airport checkpoints. This total value represents travelers making their initial entry into the aviation system, as well as any airport, airline, security, or other personnel who must enter the system. TSA throughput data are collected electronically at each checkpoint within each airport in the United States and account for entries and departures. That is, when an individual crosses through the walk through metal detector (WTMD) and into the sterile part of the checkpoint area, that individual count is added to the total throughput. When an individual crosses through the WTMD and out of the sterile area, that individual count is subtracted from the total throughput.[20]

TSA provided AIR with these data, summarized on an hourly basis for each checkpoint location. AIR used these data, along with the Base Rate Study collection period start and stop times at each collection checkpoint (documented on the Base Rate Study Data Collection Summary forms; see Figure C-6), to estimate total passenger throughput for each of those checkpoints. For cases in which a data collection period did not start or end precisely on the hour, the throughput for any fractions of an hour at the beginning or the end of the period were imputed from the TSA hourly throughput data. For example, if a data collection period ran from 8:35 a.m. to 12:12 p.m. at a given location, we multiplied the location's 8:00 a.m. hour throughput number by .42 (the fraction of the hour represented by 25 minutes, the amount of time elapsed between 8:35 a.m. and 9:00 a.m.) and the 12:00 p.m. hour throughput total by .20 (the fraction of the hour represented by 12 minutes). These counts were then added to the hourly throughput counts for the rest of the period from the 9:00 a.m. hour through the 11:00 a.m. hour (i.e., from 9:00 a.m. to 11:59 a.m.). Although this calculation is slightly imperfect because it assumes a steady flow of travelers through a checkpoint within each hour, it is the most precise estimate of actual throughput that can be reasonably obtained for a study of this nature and given the size of the traveling population of approximately 59,000,000 per month.[21] The total number of travelers summed across these data collection periods served as the total throughput for the Base Rate

[20] Note that when travelers arrive at their destination, they leave the system through an exit that is separate from the sterile checkpoint area and that does not involve a WTMD. Therefore, arrival at a destination does not result in subtraction from the throughput total. Rather, those exiting through the WTMD are, for example, travelers who have to go through it a second time because of detected metal or security personnel who cross in and out of the sterile area throughout the day.

[21] Federal Aviation Administration. (2008). *FAA aerospace forecast fiscal years 2008–2025*. Washington, DC: Author.

Study. Using this method, the total throughput estimate for the Base Rate Study was 5,920,166 travelers.

APPENDIX D:
BASE RATE SITE VISIT REPORT

Introduction

In 2007, the U.S. Department of Homeland Security's Science and Technology Directorate (DHS S&T) initiated a research program to examine the validity of the Screening Passengers by Observation Techniques (SPOT) Program's SPOT Referral Report in the context of checkpoint screening. The research, conducted by the American Institutes for Research (AIR), examined the use of the instrument in its current form in order to answer the primary research question: *To what extent does the use of the existing SPOT Referral Report lead to valid inferences about the traveling population, with a focus on high-risk travelers, or persons knowingly and intentionally trying to defeat the security process?*

The primary focus of the research that was carried out by AIR was an examination of criterion-related validity (defined as the extent to which the use of the instrument results in the correct selection of high-risk travelers). AIR examined criterion-related largely through data collected in the Base Rate Study. To ensure standardization in data collection, AIR along with the Transportation Security Administration (TSA) conducted Data Collection Monitoring Site Visits (referred to herein as data monitoring visits). This methodological standardization was critical to ensuring that the study results would generalize appropriately to the all airports using the SPOT Referral Report and would not be contaminated by variability in how the Base Rate Study was conducted during the study.

In this report we describe data monitoring visits conducted from February 2010 through September 2010. This includes the methodology employed for the monitoring visits, the results, and recommended next steps.

Methodology

Data Collection Procedure

For the Base Rate Study, data were collected on a systematic random sample of travelers, selected as they crossed a specific, predetermined marker by experienced SPOT Behavior Detection Officers (BDOs) who had completed validation study training conducted by AIR (described in the following section). Managers recommended BDOs for participation in the study on the basis of their track record as dedicated, responsible, detail-oriented, and knowledgeable personnel. Each randomly selected traveler was observed and scored for each of the SPOT indicators. Regardless of whether or not a traveler's score exceeded the threshold, all randomly selected travelers were next directed to the SPOT Referral Screening so that outcomes could be obtained. To allow direct comparisons to the standards of the SPOT Program, the data collection procedures were designed to mirror standard procedures for observation, screening, and completion of the Referral Report. A few exceptions to standard operational procedures were required for this data collection's protocol standardization:

6. In addition to the typical system of two BDOs observing the line, a third BDO served as the Data Collection Coordinator (DCC). The DCC randomly selected travelers, tracked the daily data collection, and reported any incidents that required suspending the data collection.
7. The BDO team designated a selection marker. This marker naturally varied by airport because checkpoints have varied layouts. Regardless of the location of the selection marker, the first traveler selected was the first to cross the designated marker when data collection started. Each subsequently selected traveler was the next to cross the marker once the BDO team had completed the data collection for the previous randomly selected traveler (i.e., SPOT Referral Screening and paperwork were complete).
8. The selected traveler was observed in line and then selected for SPOT Referral Screening, regardless of his or her SPOT observation score. Then, the screening followed the usual procedure to obtain a SPOT score and case outcomes.
9. The BDOs completed the paper-based SPOT Referral Report as soon as the SPOT Referral Screening was completed. This timing minimized errors in data collection and recording. BDOs were instructed to note on the Referral Report (and subsequently in the database discussed below) where the various indicators were observed. There were four possible locations for observation. This extra step was included to allow for an analysis of linkages between indicators and outcomes as a function of when a behavior was first observed.
10. If, during the data collection period, the BDO team observed another traveler exhibiting behaviors that warranted further observation, the BDOs were instructed to stop data collection to address the security issue. The DCC noted this break from protocol on the data collection tracking form.
11. Data collected on each subject was entered into the Data Collection database developed specifically for the SPOT Base Rate Study. If a randomly selected traveler was determined to be high risk as a result of referral screening, data on the passenger were entered into both the Data Collection database and the SPOT database.

The completion of each subject's data collection—from observation through resolution—averaged 20 minutes.

In addition to the SPOT Referral Report records, traveler throughput information was obtained for each data collection period and location. TSA provided a database with this information for use in the analysis.

Validation Study Training

Prior to the start of data collection at each airport location, the AIR research team, along with Government representatives, trained BDO teams on the data collection protocol (referred to herein as validation study training). Designed as a half-day session, validation study training included classroom time during which an AIR team member presented an overview of the Base

Rate Study. This overview included a discussion of the reasons the study was needed; prior analyses of SPOT incident data; the criteria for selecting BDOs to participate; and data collection procedures. The overview of data collection procedures included the following elements:

- Data collection logistics, including the composition and roles of the data collection team
- Data collection procedures, including required steps prior to, during, and after data collection
- Contingency planning, such as the procedure for selecting the passenger whenever he or she was traveling with companions
- Data collection errors or key activities that would make up the Base Rate Study outcomes (e.g., falsifying data, not following procedures).

The training materials were developed by AIR and reviewed, edited, and approved by DHS and TSA.

During training, the AIR team reinforced the position that airport security must remain the top priority. Although this study was designed for an operational setting, the regular screening of travelers was not compromised because (1) Operational SPOT would not occur at the same time as the data collection and (2) BDOs were instructed to suspend data collection activities if they observed any anomalous behavior in travelers who were not selected for screening by the Base Rate Study procedures. Training also included a discussion of procedures for formally addressing deviations from the data collection protocol.

The training session concluded with a practice data collection session at a checkpoint. During this practice session, participating BDOs divided into data collection teams and conducted a number of data collection rounds with live passengers. The BDOs took turns serving in the Data Collection Coordinator (DCC) function. AIR team members remained on the floor with the BDOs to observe the practice session, provide feedback to each team, help troubleshoot issues or concerns that were airport-specific (e.g., which marker to use when a checkpoint divides passengers into multiple different lines), and answer any questions. Following the practice data collection session, the participating BDOs and the AIR team debriefed and review key points.

Data Monitoring Visits

The main objective of the data monitoring visits was to ensure that all data collection procedures for the SPOT Base Rate Study were being implemented in a standardized manner throughout all participating airports. To this end, we observed the data collection process, as it normally occurred. This methodological standardization was critical to ensuring that the study results would generalize appropriately to the SPOT Program and would not be contaminated by variability in data collection procedures during the study. The AIR research team conducted data monitoring visits to assess the integrity of data collection, provide feedback, and answer questions.

Site Selection

TSA selected forty-three airports for inclusion in the study to represent the range of hub sizes, traveler throughput levels, and level of SPOT Program activity. Twenty-four were represented in the earliest group with staff to be trained and were thus able to begin data collection in September 2009. One additional airport was added in March 2010, and 18 more were added in July 2010. Data monitoring visits were conducted at least one time at most participating airports. The selected airports are listed in Table D-1.

Table D-1: Airports Selected by TSA as Data Collection Sites

Airport	Hub Size
Start Date: 10 September 2009	
Atlantic City International Airport (ACY)	Small
Baltimore-Washington International Thurgood Marshall Airport (BWI)	Large
Bob Hope Airport (BUR)	Medium
Cincinnati/Northern Kentucky International Airport (CVG)	Large
Denver International Airport (DEN)	Large
Des Moines International Airport (DSM)	Small
Detroit Metropolitan Wayne County Airport (DTW)	Large
Honolulu International Airport (HNL)	Large
Kona International Airport (KOA)	Small
Lambert-St. Louis International Airport (STL)	Medium
Los Angeles International Airport (LAX)	Large
Memphis International Airport (MEM)	Medium
Newark Liberty International Airport (EWR)	Large
Oakland International Airport (OAK)	Medium
Orlando International Airport (MCO)	Large
Philadelphia International Airport (PHL)	Large
Phoenix Sky Harbor International Airport (PHX)	Large
Pittsburgh International Airport (PIT)	Medium
Portland International Airport (PDX)	Medium
San Antonio International Airport (SAT)	Medium
San Diego International Airport (SAN)	Large
Tampa International Airport (TPA)	Large
Tucson International Airport (TUS)	Medium
William P. Hobby Airport (HOU)	Medium

Airport	Hub Size
Start Date: 1 March 2010	
Logan International Airport (BOS)	Large
Start Date: 16 July 2010	
Austin-Bergstrom International Airport (AUS)	Medium
Buffalo Niagara International Airport (BUF)	Medium
Charlotte/Douglas International Airport (CLT)	Large
Chicago Midway International Airport (MDW)	Large
Fort Lauderdale-Hollywood International Airport (FLL)	Large
Indianapolis International Airport (IND)	Medium
John F Kennedy International Airport (JFK)	Large
Kahului Airport (OGG)	Medium
La Guardia Airport (LGA)	Large
Louisville International Airport (SDF)	Medium
McCarran International Airport (LAS)	Large
Miami International Airport (MIA)	Large
Minneapolis-Saint Paul International Airport (MSP)	Large
Port Columbus International Airport (CMH)	Medium
Ronald Reagan Washington National Airport (DCA)	Large
Salt Lake City International Airport (SLC)	Large
Southwest Florida International Airport (RSW)	Medium
Ted Stevens Anchorage International Airport (ANC)	Medium

Logistics

At some airports, the SPOT Base Rate validation study data collection training was combined with the data monitoring visit (see Exhibit 2 in the *Results* section for a list).

The planning and scheduling for each data monitoring visit and the validation study training were conducted via a conference call that included representatives from AIR, TSA, and the participating airport. The information gathered during this call included airport location, the proposed date of the visit, point of contact (POC) information, and other pertinent information regarding logistics. During these calls, the team and the airport POC agreed that two shifts would be observed during the visit in order to capture the most comprehensive data. Exhibit D-1, at the end of this appendix, shows the Data Monitoring Visit Initial Call Sheet.

Base Rate Study Data Monitoring Checklist

The SPOT Base Rate Study Data Monitoring Checklist (referred to herein as the data monitoring checklist) was devised to be used by the research team while observing BDOs engaged in collecting data during the data monitoring visits. The data monitoring checklist included background information to be completed for each observation, as well as items that addressed

specific data collection procedures (prior, during, and after data collection) that would need to be observed. Exhibit D-2, at the end of this appendix, shows the Data Monitoring Checklist. All AIR researchers who participated in the site visits were trained in executing the data monitoring process, including instruction on the procedure for completing the data monitoring checklist. No concerns or issues were raised about the use of the data monitoring checklist. Had questions or issues come up as the AIR researchers were completing the data monitoring checklist, procedures had been put in place to ensure that they were addressed immediately following the monitoring visit.

Prior to Observation of Data Monitoring Process

For each site visit, two researchers from AIR were deployed to the respective airports.

Upon arrival at the designated airport, the research team met with the POC and/or BDO team to discuss the following:

- Where best to position the research team in order to observe the BDO team unobtrusively
- How to observe the areas in front of and behind the security checkpoint

The research team then observed the BDO team meeting, which took place prior to data collection, noting the following:

- Positioning of the DCC and the BDOs
- The specific communication strategy BDOs used during data collection (If no communication strategy was discussed, the research team would confer with the DCCs to inform them of the omission, note the reason given, and determine the communication strategy .[22])
- Start time of the data collection
- Selection position/marker to be used

Prior to the data monitoring process, the research team documented as much of the logistical information as possible on the data monitoring checklist: date, airport location, screening location, names of the DCC and the BDOs, and names of the research team.[23] The research team also completed the Prior to Data Collection section of the data monitoring checklist. This section essentially confirmed to the research team that all protocols for data collection/monitoring were in place (e.g., start time, communication, selection marker).

[22] The prevailing reason BDOs may not discuss a communication strategy is due to the tacit understanding many BDOs have with their co-workers; that is, they have worked together for a period of time and implicitly know what communication strategy will be used during data collection.

[23] The remaining logistical information on the data monitoring checklist is to be completed at the end of the observation (i.e., end time of observation, total number of passengers observed, and any anomalous observations).

Observation of Data Monitoring Process

After these pre-observation issues were agreed upon and the logistical information was documented, the research team began the observation. The observation lasted approximately two to three hours. The research team monitored the BDOs' SPOT Program observations as outlined in the Base Rate Study. The data collection checklist allowed the research team to check off all items for each passenger observed. The research team took careful notes and remained as inconspicuous as possible.

After Data Collection/Summary

Following the observation of the data monitoring process, the research team conferred with each other and completed the After Data Collection/Summary section of the data monitoring checklist, noting the data collection team's adherence to the following procedures:

(b)(3):49 U.S.C. § 114(r)

The research team completed the logistical information on the data monitoring checklist and summarized their data monitoring observations, noting any issues and/or problems that arose during the observation.

Debriefing

Following the observation, the research team conducted a debrief meeting with the data collection team. This meeting was an opportunity for the data collection team to ask any questions and for the research team to provide feedback on their observations. The issues and questions that arose during the debriefing of the first eight data monitoring visits are addressed in the following *Results* section.

Results

The research team conducted seven large-scale validation study trainings, each at different airport locations: Boston Logan International, Denver International, Detroit Metropolitan, Los Angeles International, Orlando International, Pittsburgh International, and Tucson International. A training session at Orlando International was conducted at the onset of data collection and allowed the standardized training of a large number of BDOs from different airports at one time. Two additional trainings were conducted when the study began in fall 2009, and four trainings were offered in the first quarter of 2010. These six trainings were small and involved local BDOs only. They were offered to help airports increase their staffing for the data collection. These smaller trainings were coupled with a data monitoring visit to maximize efficiency (e.g., during the two-day visit to TUS, AIR researchers conducted a one-day training session for additional local BDOs and then conducted data monitoring on the second day). The research team conducted 17 data monitoring visits. Table D-2 indicates the airports at which trainings and data monitoring visits took place.

Table D-2: Data Monitoring Visits and Validation Study Training Sites

Location	Date of Visit	Monitoring	Training
Baltimore Washington International Airport (BWI)	9/14/2010	Y	N
Boston Logan International (BOS)	3/11/2010	N	Y
Chicago Midway International Airport (MDW)	10/5/2010	Y	N
Denver International (DEN)	2/22–23/10	Y	Y
Detroit Metropolitan Wayne County Airport (DTW)	2/16/2010	Y	Y
Indianapolis International Airport (IND)	10/4/2010	Y	N
John F Kennedy International Airport (JFK)	9/21/2010	Y	N
La Guardia Airport (LGA)	9/22/2010	Y	N
Los Angeles International (LAX)	12/17/2009	N	Y
Metropolitan Oakland International (OAK)	2/24/2010	Y	N
Minneapolis - St. Paul International Airport (MSP)	10/7/2010	Y	N
Newark Liberty International Airport (EWR)	9/15/2010	Y	N
Orlando International (MCO)	9/9–11/09	N	Y
Philadelphia International Airport (PHL)	9/16/2010	Y	N
Phoenix Sky Harbor International (PHX)	2/17/2010	Y	N
Pittsburgh International (PIT)	11/12/2009	N	Y
Port Columbus International Airport (CMH)	10/6/2010	Y	N
Portland International (PDX)	3/3/2010	Y	N
San Antonio International (SAT)	4/28–29/10	Y	N
Tampa International (TPA)	4/28–29/10	Y	N
Tucson International Airport (TUS)	2/16/2010	Y	Y

The overall process of arranging and conducting the site visits worked effectively for the eight sites that were monitored. The only exception with respect to arrangements was for the DEN site

visit in which the agreed-upon schedule was changed after the team arrived. The DEN change was most likely due to a misunderstanding of the purpose of the visit. Overall, airport POCs and BDOs, including those at DEN, were extremely receptive and willing to assist with the data monitoring process.

In general, the BDO teams followed the data collection process as specified. However, some issues were clarified to maintain data integrity and ensure standardization across data collection sites.

Specific Issues Identified

As would be expected in any data collection using multiple data collection teams at different sites, some deviation from specified procedures had resulted over time. Each deviation was discussed with the relevant team members to remind them of the procedures and the importance of standardization of method. These issues concerned the following:

1. Random selection of passengers
2. Positioning of DCC and BDOs
3. Communication
4. Recording of the location of observed SPOT indicators
5. Application of inclusion/exclusion criteria
6. Continuation of data collection throughout the data collection shift
7. Understanding of the purpose of monitoring site visits

Below we briefly discuss each identified issue.

Random Selection of Passengers

The most crucial identified issue was the deviation of some data collection teams from the random selection procedures as outlined in the protocol. Although the airports in question implemented what they deemed to be a random selection procedure, this deviation did not actually meet criteria for a systematic random selection and as such, if left uncorrected, could increase the likelihood of introducing systematic bias into the study. Because of variations in checkpoint layout and throughput, it would be expected that such things as location of the marker would differ. However, the specific procedures for random selection should remain standardized across locations.

One airport implemented a numbering system in which the DCC would call out a "random" number (n) and the data collection team would then choose the nth passenger to cross the selection position/marker. Although this method may have appeared to the team to be random, humans are not able to generate truly random numbers. Other adaptations of the selection procedure were more problematic in terms of the potential to introduce bias. In a few airports, the DCC walked back to his or her starting point, watching the passengers approaching the marker. This provided too much opportunity for the DCC's expectation about which passenger would cross the marker and when, to affect, without his or her knowledge, his or her speed of return to position. In fact, it was noted that in one of these checkpoints, the DCC would linger for some unspecified and variable amount of time before concluding that he was ready and in

position, and only then was the next passenger selected. Another DCC misunderstood the selection protocol, thinking that he was not permitted to communicate the identity of the selectee verbally. He taught his team that he would give them a nonverbal signal when he was ready, and the BDO closest to the marker was to identify the next passenger crossing the mark.

(b)(3):49 U.S.C. § 114(r)

Sometimes the deviation had to do with the use of the selection marker. Data collection protocol specifies that the marker used for selection should be unmovable and used throughout the data collection session. At one airport, the DCC would "randomly" alternate between selection markers, each for a separate line of passengers feeding into a single line up to the Travel and Document Checker (TDC). The clear problem with this strategy is that there is nothing actually random about the selected time to change between markers, so any decision to switch between markers could reflect a systematic bias.

Across airport sites, the data monitors frequently heard BDOs discussing types of travelers (b)(3):4 (b)(3):49 U.S.C. § 114(r) and the challenge some were likely to present upon being randomly selected for a secondary screening. Thus, any deviation from the selection protocol could increase the risk of the insertion of some judgment into the selection process.

Positioning of DCC and BDOs

(b)(3):49 U.S.C. § 114(r)

Communication

Generally speaking, data collection teams did well at communicating. However, a few communication problems were noteworthy. In a few cases, there was some confusion among team members about precisely who the selectee was. For one of these teams, this was directly related to the fact that the BDO erroneously thought that he was supposed to communicate this information nonverbally.

Recording of the Location of Observed SPOT Indicators

This issue concerned the requirement for the data collection team to record the location at which each SPOT behavior was observed (b)(3):49 U.S.C. § 114(r) (b)(3):49 U.S.C. § 114(r) However, at one airport, the DCCs were not recording the location at which the behaviors(s) were observed, but rather marking the behavior with an X.

Application of Inclusion/Exclusion Criteria

(b)(3):49 U.S.C. § 114(r)

Continuation of Data Collection Throughout the Data Collection Shift

Data collection was expected to continue uninterrupted throughout the shift unless some operational imperative required a break in, or termination of, data collection. A data collection team at one airport ceased data collection after each randomly selected passenger to conduct operational SPOT activities. Because this deviated from the specific steps described in the data collection protocol, AIR recommended reiterating that protocol must be followed in order to maintain integrity of data.

Understanding of the Purpose of Monitoring Site Visits

Data collection site monitors did note some confusion on the part of data collection teams at a few airports concerning the research team's goals at site visits. As emphasized with the airport POCs, the research team was there to observe the data collection process *as it normally occurred*.

Specific Questions Asked by Data Collection Teams

As with the application of inclusion/exclusion criteria issue described above, many of the questions asked during the debriefing could be addressed by referring back to standard SPOT protocol. Examples of this type of question follow:

(b)(3):49 U.S.C. § 114(r)

(b)(3):49 U.S.C. § 114(r)

Exhibit D-1: Data Monitoring Visit Initial Call Sheet

BASE RATE STUDY MONITORING VISITS
INITIAL CALL WITH POCS

Airport Location: _____
Data Monitoring Date (mm/dd/yyyy):_____
Point of Contact (POC):_____
POC Title:_____
POC Role (relative to validation study):_____
POC Phone Number(s):_____
POC email:_____
AIR Data Monitoring Team:_____
_____

Data Collection Shift(s) (AM or PM)?_____
Shift Start Time:_____ Shift End Time _____
Shift Start Time:_____ Shift End Time _____

If shifts overlap, can observation be done for both shifts?

What time would the data monitoring team have to be there?

Where in the airport would the data monitoring team meet the POC?

What information is required for entering the sterile area? When is the information needed?

How easy would it be to have the observers moving in and out of the sterile area?

If not, will data monitoring team be able to be positioned, one in sterile area and one outside?

Will there be time for a debriefing following the data monitoring?

Is there a room available to have debriefing?

If there is training scheduled:

How many BDOs will attend?_____

Will we have access to a projector?_____

Is there a particular hotel that you would recommend?

Exhibit D-2: Base Rate Study Data Monitoring Checklist

BASE RATE STUDY DATA MONITORING CHECKLIST

Overview

Purpose

The Data Monitoring Checklist is to be used during the observation of BDOs engaged in collecting data for the SPOT Base Rate Study at each site visit. The main goal of the observation is to ensure that all data collection procedures are being implemented in a standardized and consistent manner.

Instructions for Observation

Please use the Data Monitoring Checklist as a guide to your observation of the data collection activities. The checklist includes background information to be completed for each observation, as well as items that address specific data collection procedures that need to be observed.

• Before the observation period, spend a few minutes talking to the airport POC or the BDO team. Especially focus on discussing your positioning for optimal observation. Remember, we want to capture as much data as possible with minimal disruptions. With respect to position, discuss

o where to best position yourself so that you can observe the BDO team unobtrusively; and

o how to observe both in front of and behind the security checkpoint.

• In order to observe steps taken prior to the data collection, start the observation by attending the initial BDO team meeting.

• Make sure you are clear on all items, particularly the communication strategy being used by the BDO team. (If communication strategy is not discussed, make sure you discuss with the DCC that (a) you did not observe them discussing a communication strategy (note the reason given) and (b) you will need to know what the communication strategy will be.)

• You will be observing a number of passengers over a period of 2 to 3 hours. The Data Monitoring Checklist allows you to check off all items for each passenger observed. (The Data Monitoring Checklist allows recording observations for 10 passengers. Use additional sheets as necessary.)

• While observing, take careful notes. Again, it is important to be as inconspicuous as possible. Consider bringing a small notebook and taking notes in stages. Periodically check your notes for accuracy. After observing, check all of your notes to be sure you captured accurately all that you observed.

• You will have a debriefing with the BDO team after your observation. This is an opportunity for the BDO team to ask questions and for you to provide feedback on what was observed.

SPOT Base Rate Study Data Monitoring Checklist		
Date (mm/dd/yyyy):	Start Time:	End Time:
Time: Airport Location:	Screening Location:	
Data Collection Coordinator (Last, First):		
BDO 1 (Last, First):		
BDO 2 (Last, First):		
AIR Observer 1 (Last, First):		
AIR Observer 2 (Last, First):		
Total Number of Passengers Observed:		
Observations:		

ITEM	Y/N/Partial	Comments
Prior to Data Collection		
The DCC and BDOs meet to ensure that everyone understands the specifics of the data collection for the day. Discussion includes:		
Positioning of the DCC and the BDOs		
Communication strategies to be used between the DCC and BDOs		
Start time of the data collection		
Selection Position/Marker to be used		
The DCC and BDOs have all required materials.		
TSA-issued cell phone or working radio		
Working black light and magnification loupe, if available		
Writing instrument and a notebook		
Set of SPOT Assist Cards		
The Selection Marker selected is unobtrusive and is stationary.		
The BDOs and the DCC identify themselves to the Travel and Document Checker (TDC) and the Screening Checkpoint STSO.		
The DCC completes the top section of the Data Collection Summary Form prior to the start of the data collection and provides complete information on:		
• the current data collection day		
• start time		
• location		
• team members		

ITEM	PAX1	PAX2	PAX3	PAX4	PA□5	COMMENTS
Dur□ng Data Collection (Y/□)						
The DCC/BDOs establish the environmental base□ine.						
(b)(3):49 U.S.C. § 114(r)						

ITEM	PAX1	PAX2	PAX3	PAX4	PAX5	COMMENTS
(b)(3):49 U.S.C. § 114(r)						
Once a resolution has been reached, the DCC and BDOs confirm and concur on the resolutio.						
The D completes a paper copy of the SPOT Referral Form imme iately following resolution.						
DCC notes on the SPOT Referral Form which behaviors were observed (b)(3):49 U.S.C. § 114(r) (b)(3):49 U.S.C. § 114(r)						
The DCC and BDOs return to their positions after completing the SPOT paperwork.						
DCC selects the next passenger who passes the Selection Position.						

ITEM	Y/N/Partial	Comments
After Data Collection/Summary		
DCC is randomly selecting passengers using the preselected Selection Position at all times.		
The Selection Position was not changed throughout shift.		
DCC completes a SPOT Referral Form (b)(3):49 U.S. (b)(3):49 U.S.C. § 114(r)		

Summary Observations

APPENDIX E:
EXPLORATORY FACTOR ANALYSIS

Introduction

Appendix E presents the findings of a preliminary examination of the construct-related validity of the SPOT Referral Report's Section 2 indicators. Specifically, we used exploratory factor analysis (EFA) as a means to begin to examine the underlying structure of the SPOT Referral Report. This analysis focused on an examination of whether there are some interpretable constructs that might underlie the items in Section 2 of the SPOT Referral Report. A comprehensive examination of the construct-related validity of this measure could not be conducted because of limitations in the data, primarily due to how the instrument is scored and the rare event nature of the items. Nonetheless, there was merit to conducting EFA and correlation analysis as a means to begin the examination of construct-related validity.

Results of these analyses may provide initial insight about the meaning and function of the SPOT Section 2 indicators, as well as the manner in which the indicator set is used in the SPOT Program. Moreover, results might suggest ways to better optimize the list. For example, if certain indicator pairs were found to be highly correlated, it might be possible to combine them or reduce such pairs to a single item. If strong patterns were to emerge in these analyses, such findings could have important implications for operational procedures or training. Finally, results will be informative in developing the next steps for research to conduct a more complete examination of construct-related validity.

It is important to again emphasize that the nature of these analyses and resulting findings are exploratory and additional studies are required to provide supporting evidence. The potential challenges are discussed in the following *Method* section, and include the consistent issue related to the nature of the data in that it deals with rare events, and thus rare observances of the indicators. Furthermore, as the principal goal of this study was not an investigation of construct-related validity as it pertains to item representation of the constructs purported in the SPOT Referral Report, the data were not collected for this purpose. Given these limitations, the results are subject to cautious interpretation, but still provide a foundation for future studies.

Method

The research team examined the extent to which there may be underlying constructs among the set of indicators for the Operational SPOT dataset by using two approaches to analysis: correlation and factor analysis. First, Pearson correlations were computed for each pair of indicators. The correlation value (r), significance level (p), and effect size (r^2) were each used to determine the strength of correlations among pairs of items. It was expected that indicators of the same underlying construct would be more highly intercorrelated than indicators of other distinct constructs. Given the large size of the Operational SPOT dataset, we conducted independent analyses on the Operational SPOT stratified split-half subsets. These data subsets are described previously and used for the analyses of indicator-level predictive utility, reported in

Chapter 7.[24] The use of the stratified subsets allowed us to examine stability in results and address possible random associations in the data.

Next, an exploratory factor analysis (EFA) was conducted to further examine underlying constructs, or latent variables, that may exist among the indicator set.[25] We did not restrict the factor analysis to the three constructs of stress, fear, and deception,[26] but rather considered the best fit for the data. The EFA was run in MPlus,[27] which has a procedure appropriate for categorical (i.e., binary) data. Given the exploratory nature of this analysis, we used the split subsets to develop the factor analysis solution and subsequently test its stability. The EFA model was developed on the first split subset of the Operational SPOT dataset, varying several parameters to derive the best solution. The team considered various factor extraction methods (i.e., eigenvalues >1 [Kaiser criterion] and Scree plot), as well as corresponding factor solutions and various rotation methods (i.e., orthogonal and oblique types). Factor solutions were evaluated on the basis of percent of variance accounted for, factor loadings, parsimony, and solution interpretability. Any reasonable EFA model (factor solutions) was then tested on the second split subset to examine model stability.

For each stable model, we then examined the extent to which each of the factors related to outcomes. First, factor-based scores from Subset 1 results were computed by multiplying each indicator by its factor loading. Then, following an examination of factor-based score descriptives and distributions, correlations between factor-based scores were examined. Finally, t-tests were conducted to examine differences in factor-based scores for the presence or absence of each outcome.[28] The difference in group sizes, with less than 10% of cases in any single outcome group, was expected to bias any analysis of group differences. Thus, we selected a random sample of 5% of non-target cases (absence of outcomes) to compare to all target cases (presence of outcomes). Balancing this dataset meant that any significant results could then be attributed to meaningful group differences rather than to unbalanced samples. The two Operational SPOT subsets were used to investigate stability in results.

Anticipated Challenges

(b)(3):49 U.S.C. § 114(r)

[24] See Appendix F, Tables F-22 toF-26 for characteristics of split Subsets 1 and 2.
[25] The three indicators that have negative points-values (i.e., lower risk) were excluded from this analysis.
[26] The SPOT Referral Report groups the Section 2 indicators into three categories of stress factors, fear factors and deception factors, which correspond to 1-, 2-, and 3-points values, respectively.
[27] Muthén, L. K., &. Muthén, B. O. (1998-2010). *Mplus user's guide (6th Ed.)*. Los Angeles, CA: Author.
[28] Homogeneity of variance was examined using Levene's test. Results of any t-tests that failed the assumption of equal variances (at p<.01) were adjusted accordingly.
[29] American Institutes for Research (2010). *SPOT validation research overview*. Washington, DC: Author.

(b)(3):49 U.S.C. § 114(r)

Moreover, unlike a diagnostic scale in which an individual who exhibits one symptom would be expected to exhibit several additional symptoms of the same construct (e.g., anxiety or depression), the SPOT Referral Report includes the large indicator set because of the many and varied signs of deception. A positive finding on the SPOT Referral Report instrument could be a result of any number of combinations of indicators or signs. In contrast, there are a finite number of expected patterns of symptom endorsement on a more typical diagnostic instrument. This difference in the SPOT instrument as compared to other types of instruments may affect results of correlation and factor analyses, which are based on communalities, or shared variance, among items in that there may not be sufficient and consistent patterns of relationships (communalities) among the items to produce strong factor analytic results.

Results

Correlations

(b)(3):49 U.S.C. § 114(r)

Table E-1: Section 2 Indicator Pairs With Correlations Greater Than |0.10|

(b)(3):49 U.S.C. § 114(r)

Exploratory Factor Analysis

The foregoing correlation analysis was also conducted as a first step in the EFA to examine the extent of shared variance among the indicators. The greater the degree of correlation in a correlation matrix, the more likely it is that the factor analysis will produce a meaningful model. The matrices for correlations among (b)(3):49 U.S.C. § 114(r) showed only (b)(3):4 (b)(3):49 U.S.C. § 114(r) therefore any factor solution was expected to be limited; moreover, factor solutions would likely not represent all variables included in the model. However, despite the small correlations, the matrices were significantly different from the identity matrix (correlation matrix of (b) therefore, it was appropriate to proceed with the (b)(3):49 U.S.C. § 114(r)

(b)(3):49 U.S.C. § 114(r)

Table E-3: Variance Accounted For by Factor Solutions

		Number of Factors					
		3	4	5	6	7	15
Variance Accounted For by Factor	1	(b)(3):49 U.S.C. § 114(r)					
	2						
	3						
	4						
	5						
	6						
	7						
	8						
	9						
	10						
	11						
	12						
	13						
	14						
	15						
Total Variance Accounted For							

Similar results were obtained for obliquely rotated models. This was expected given (b)(3):49 (b)(3):49 U.S.C. § 114(r) among obliquely rotated factors. Factor correlations ranged from (b) to (b) depending on the number of factors in the model. (b)(3):49 U.S.C. § 114(r) indicated that the orthogonally rotated solutions were preferred.

Model Stability

The best EFA solutions—(b)(3):49 U.S.C. § 114(r) —were next tested on the second subset of Operational SPOT data to examine model stability. Overall, (b)(3):49 (b)(3):49 appeared the most stable. While there was some natural fluctuation in factor loadings across samples, this model had the same sets of items loading together on each factor.

(b)(3):49 U.S.C. § 114 (r) showed stability in (b)(3):49 U.S.C. § however, (b)(3):49 U.S.C. § 114(r) (b) great variability, with substantially different loadings (b)(3):49 U.S.C. § 11. Therefore this factor model was found to be less stable and not worth further consideration.

Model Interpretation

The resulting (b)(3):49 U.S.C. § 114(r) accounted for (b) (3):4 of variance and included (b) ndicators with loadings (b) (3):49 see Appendix Table G-2 for factor loadings). There is variation in guidance about interpretation of factor loadings, with recommended ranges of

minimal loadings from (b)(3):49 U.S.C. § 114 (r) ue to lack of information in the data (i.e., data sparseness), we considered lower factor loadings in model interpretation (including interpretation of some items with factor loadings lower than (b) Thus, the factor solution must be considered with caution and as only a first step in a process aimed to understand the extent to which there may be underlying constructs among the set of indicators for the Operational SPOT dataset. Again, the intent of this analyses and proposed explanation of the emerging (b)(3):49 U.S.C. § 114(r) is to provide a first look at the ways in which the indicators are being used in combination. Moreover, it is important to note that the explanation of any resulting model from an exploratory factor analysis is subject to interpretation. The following interpretation, using the factor loadings for Subset 1,[32] was developed by qualified team members with experience related to: (1) the Section 2 indicators, (2) an educational background in human behavior, and (3) deception detection theory and research. Of course, other interpretations are worth consideration and future research should include further testing of interpretations and theories related to any factor analysis results (these or other analyses).

(b)(3):49 U.S.C. § 114(r)

Several of these negative loadings suggest (b)(3):49 U.S.C. § 114(r)
(b)(3):49 U.S.C. § 114(r) Also, the factor loading for other hostile and uncooperative behaviors such as (b)(3):49 U.S.C. § 114(r)
(b)(3):49 U.S.C. § 114(r) are quite low, as are the loadings for (b)(3):49 U.S.C. § 114(r)

[32] Factor loadings for Subset 2 are included in Appendix Table G-2 and are virtually identical to Subset 1 loadings.

(b)(3):49 U.S.C. § 114(r)

(b)(3):49 U.S.C. § 114(r)

(b)(3):49 U.S.C. § 114(r)

(b)(3):49 U.S.C. § 114(r)

(b)(3):49 U.S.C. § 114(r)

(b)(3):49 U.S.C. § 114(r)

Summary

The analyses reported in this appendix explored the extent to which there may be underlying constructs among the set of Section 2 indicators on the SPOT Referral Report. The breadth of the analyses, consisting of correlations and EFA, was limited by the available data. Several challenges were anticipated with this analysis including the low frequency of occurrence of the indicators and the highly varied patterns within referred cases, and thus caution is necessary in interpreting the results. Despite these limitations at the outset of this exploratory analysis, some interesting findings emerged that are worth considering and investigating further in future research on the construct-related validity of the SPOT Referral Report.

Despite low correlations and communalities among indicators, results of the EFA found a relatively meaningful and (b)(3):49 U.S.C. § solution. This (b)(3):49 U.S.C. § is, of course, subject to interpretation and should be considered as only a first step to understanding how the Section 2 indicators may be inter-related and representative of underlying constructs. Our interpretation of the solution is that these (b)(3):49 U.S.C. § 114(r) traveler profiles: (b)(3):49 U.S.C. § 114(r) (b)(3):49 U.S.C. § 114(r) (b)(3):49 U.S.C. § represented only 21% of the variance, or correlation, among indicators and included only 25 (61%) indicators with moderate loadings, the factors did meaningfully relate to outcomes. Specifically, relatively more positive factor-based scores (lower negative values) on the factor of (b)(3):49 U.S.C. § 114(r) (b)(3):49 U.S.C. § 114(r) In addition, for the factor (b)(3):49 U.S.C. § (b)(3):49 U.S.C. § 114(r) (b)(3):49 U.S.C. § 114(r) were significantly related to the occurrence of a *LEO Arrest*. These findings suggest that there may be underlying constructs that are represented by Section 2 indicators, and these factors differentially relate to outcomes. Thus, the results of the analyses are contribute to our understanding of the indicators by suggesting that the diverse set of items can be characterized into groups, or factors, though additional analyses are required. This analysis is informative as a first step in exploring construct-related validity of the SPOT Referral Report. Future research is recommended in this area to further instrument development and validation.