DOT/FAA/AR-95/16

Office of Aviation Research
Washington, D.C. 20591

Test and Evaluation Plan for the Explosive Device Detection Baseline (EDDB) Study

J. L. Fobes, Ph.D.
Brenda A. Klock

Airport Security Human
Factors Program, AAR-510
FAA Technical Center
Atlantic City International Airport, NJ 08405

August 1995

Final Report

This document is available to the public
through the National Technical Information
Service Springfield, Virginia 22161

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AR-95/16		

4. Title and Subtitle	5. Report Date
Test and Evaluation Plan for the Explosive Device Detection Baseline (EDDB) Study	August 1995
	6. Performing Organization Code AAR-510
7. Author(s) J. L. Fobes, Ph.D., D. Michael McAnulty, Ph.D., and B. A. Klock of the FAA, and J. Janowitz, Ed.D., E. Neiderman, Ph.D., and R. L. Malone of Galaxy Scientific Corporation.	8. Performing Organization Report No.
9. Performing Organization Name and Address U.S. Department of Transportation Federal Aviation Administration FAA Technical Center Atlantic City International Airport, NJ 08405	10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTFA03-89-C-00043
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Aviation Administration FAA Headquarters 800 Independence Ave., S.W. Washington, DC 20471	13. Type of Report and Period Covered Draft Final 14. Sponsoring Agency Code ACS-1

15. Supplementary Notes

16. Abstract

This document is the Test and Evaluation Plan (TEP) to evaluate Improvised Explosive Device (IED)detection capabilities. Specifically, the testing effort will evaluate the ability of airport security personnel to detect IEDs in carry-on passenger bags.

The test and evaluation (T&E) focuses on determining the baseline performance levels as set forth in the Critical Operational Issues and Criteria (COIC).

The T&E will be conducted at 19 U.S. Category X (CAT X) airports. The results will be analyzed and become part of a later document.

17. Key Words	18. Distribution Statement
X-ray, X-ray screening, IED, improvised explosive device, IED detection, signal detection, Human Factors Engineering	This document is made available to the public through the National Technical Information Service Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	34	

Form DOT F1700.7 (8-72)Reproduction of completed page authorized

PREFACE

This test plan was developed to support investigation of the Critical Operational Issues and Criteria (COICs) set forth by the Federal Aviation Administration (FAA) to assess the capability of airport baggage screeners to detect Improvised Explosive Devices (IEDs) in carry-on (CO) passenger baggage using black/white X-ray equipment. The key FAA personnel supporting this testing effort are J. L. Fobes, Ph.D., Aviation Security Human Factors Program Manager and Engineering Research Psychologist for the Aviation Security Research and Development Division (AAR-510), D. Michael McAnulty, Ph.D., an Engineering Research Psychologist with the Aviation Simulation and Human Factors Division (ACT-500), and Brenda A. Klock, technical specialist for Aviation Security Research and Development Division (AAR-510).

Galaxy Scientific Corporation prepared this document under contract number DTFA03-89-C-00043 with the FAA Technical Center. The Program Manager at Galaxy Scientific Corporation is William Hassler, Jr. The authors of this document are Joan Janowitz, Ed.D., Eric Neiderman, Ph.D, and Robert Malone.

TABLE OF CONTENTS

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

AAR	Aviation Security Research and Development Division
ACS	Associate Administrator for Civil Aviation Security
ANOVA	ANalysis Of VAriance
c	Operator Response Criterion
CAT X	Category X
CO	Carry-On
COIC	Critical Operational Issue and Criteria
d′	d prime, Derived Operator Sensitivity
DOT	Department of Transportation
ED	Explosive Device
EDDB	Explosive Device Detection Baseline
EDDS	Explosive Device Detection System
FAA	Federal Aviation Administration
FAR	Federal Aviation Regulation
IED	Improvised Explosive Device
MOP	Measure of Performance
N_c	Number of Correct Answers
N_{fa}	Number of False Alarms
N_h	Number of Hits
N_i	Number of Incorrect Answers
N_t	Total Number of Test Bag Images
P_d	Probability of Detection
P_{fa}	Probability of False Alarm
P_h	Probability of Hit
P_m	Probability of Miss
ROC	Receiver Operating Characteristic
SDT	Signal Detection Theory
SPEARS	Screener Proficiency Evaluation and Reporting System
T&E	Test and Evaluation
TEP	Test and Evaluation Plan
TER	Test and Evaluation Report
TWA	Trans World Airlines

1. INTRODUCTION.

1.1 PURPOSE.

The Federal Aviation Administration (FAA), in concert with the U.S. aviation industry, is in the process of developing new equipment and procedures to improve the national aviation security system, including detection of Improvised Explosive Devices (IEDs). It is planned to operationally demonstrate some of these new technologies and procedures, at very diverse locations, to verify their performance enhancement benefits. Before this demonstration takes place, however, it is necessary to establish a baseline of current aviation security performance, against which any performance improvements can be compared.

The purpose of this Test and Evaluation Plan (TEP) is to present, explain, and discuss the conduct of the Explosive Device Detection Baseline (EDDB) study to be conducted at the 19 U.S. Category X (CAT X) airports. The study will be conducted in support of the Aviation Security Human Factors Program at the Federal Aviation Administration Technical Center, Atlantic City International Airport, New Jersey, under Research Project Initiative #127 in support of Mission Need Statement #163.

1.2 SCOPE.

The EDDB will evaluate the capability of airport baggage screeners to detect IEDs in carry-on (CO) passenger baggage using black/white X-ray equipment. The X-ray baggage screening environment will be simulated by having operational airport baggage screeners at each of the 19 CAT X airports scan for IEDs in X-ray images presented by a computer-based library of digitized images.

Historically, threat objects such as weapons have resulted in the highest detection rates in operational settings. In contrast, detection rates for IEDs have been comparatively lower. IED detection performance is, therefore, of specific concern to this study and testing will focus on the ability of screeners to detect IEDs

1.3 BACKGROUND.

Since its creation in 1958, the FAA has had the responsibility of ensuring the safety of air travel. Both airports and aircraft are vulnerable to terrorist attacks, but airports pose a particular challenge to security. They must be readily accessible to the public, yet prevent persons with malicious intent from penetrating secure areas. Airport baggage screening is part of the FAA's security concept for airports in which no single system is relied upon to provide security required by regulation. This concept involves a complex system composed of trained personnel, properly maintained and calibrated equipment, and appropriate procedures to provide multiple layers of security from the airport perimeter to the aircraft door. Preboard baggage screening is part of the airport security checkpoint system and part of the total airport security system.

A number of policies affect preboard baggage screening checkpoint operations. Federal Aviation Regulation (FAR) Part 107, Airport Security, section 107.20 indicates that "No person may enter a sterile area without submitting to the screening of his or her person and property in accordance with procedures being applied to control access to that area." FAR Part 108.9 and FAR Part 129.25 present screening policies for U.S. and foreign air carriers. Both policies indicate that airlines may refuse to transport any person who does not consent to a search of his or her person. Furthermore, passenger checked baggage may be examined for the presence of potential threats. Annex 17 to the Convention On International Civil Aviation (1993) describes preventive security measures that must be established to prevent weapons, explosives, or any other dangerous devices that may be used to commit an act of unlawful interference from being introduced onboard an aircraft engaged in international civil aviation.

The threat to civil aviation security has changed dramatically in the last decade. This change in threat has resulted in new challenges with regard to passenger and baggage screening. In the 1980s, the threat was hijacking. The role of the FAA in aviation security against hijacking was greatly expanded, especially after the 1985 hijacking of Trans World Airlines (TWA) Flight 847 in the Middle East. In the 1990s, there has been a shift from threats due to hijacking to concern about sabotage by bombings.

Improvements in technology available to hostile elements, especially in the area of explosive devices, have resulted in increased airliner vulnerability to bombings. Terrorists are reducing their use of prefabricated explosive devices, such as grenades, and opting for less detectable IEDs. An IED can be made from a variety of materials that may resemble innocent or everyday objects, such as batteries, wires, and digital clocks. For example, plastic explosives made with Semtex and C-4 can be shaped and molded into sheets or cubes that, when passed through X-ray screening devices, appear as innocent items such as books or radios. Terrorists have also learned to embed IEDs in electronic devices, as in the Pan American Flight 103 disaster, making detection even more difficult. In addition, miniaturization and digitization of timing devices compound the problem of IED detection with X-ray screening.

Sophisticated terrorists have the knowledge and materials to build difficult to detect IEDs. The potential for complete aircraft destruction, with the loss of hundreds of lives and the disruption of the National Airspace System, has increased. As a result of this shift toward a higher potential for disaster, the focus of civil aviation security has changed from hijackings to methods of countering bombings. This shift has markedly increased the need for improvements in screener systems and operator training. The Associate Administrator for Civil Aviation Security (ACS) and the Office of Civil Aviation Security Policy and Planning have identified the need for research into the performance of Explosive Device Detection Systems (EDDS), particularly the human component in detecting IEDs.

1.4 SYSTEM DESCRIPTION.

IED detection testing will be carried out using a computer-based black/white X-ray image presentation device. Two hundred digitized bag images will be shown; 25 bag images contain simulated IEDs built from modular bomb sets.

1.5 CRITICAL OPERATIONAL ISSUES AND CRITERIA (COIC).

The following sections present the operational issues and associated criteria to be assessed during this study:

1.5.1 Issue 1—Baseline IED Detection Performance.

What is the current screener IED detection performance at each U.S. CAT X airport?

Currently, baseline screener IED detection performance is unknown as measured by administering the same detection task to sample screeners from various airports. This issue will be assessed by administering the same test to all sample screeners and recording the number of hits (N_h) and the number of false alarms (N_{fa}) observed for each screener participating in this study.

The Measure of Performance (MOP) data described in this section will be used to derive values for individual and average screener probability of detection (P_d) and the probability of false alarm (P_{fa}). Values of decision criterion ($\underline{c}$) and operator sensitivity (d') will be derived according to the Signal Detection Theory (SDT) paradigm. A discussion of the SDT paradigm is included in Appendix A.

Criterion None. This issue is investigative in nature.

MOP 1-1. The N_h observed for simulated IEDs.

MOP 1-2. The N_{fa} observed for simulated IEDs.

1.5.2 Issue 2—IED Detection Performance Differences.

Are there significant differences in IED detection performance between CAT X airports?

Based on the data obtained for issue 1, this issue will be assessed by determining any statistically significant differences in IED detection test performance across the screener samples drawn from the 19 CAT X airports. IED detection performance values of P_d, P_{fa}, $\underline{c}$, and d' will be compared across airports.

Criterion. None. This issue is investigative in nature.

MOP 2-1. See MOP 1-1.

MOP 2-2. See MOP 1-2.

1.6 TEST AND EVALUATION LIMITATIONS AND IMPACT.

a. The computerized testing may lack some aspects of operational representativeness compared to actual screening duties at a security checkpoint. The effects of this diminished representativeness are unknown.

b. Performance could be affected as a result of the operators' awareness that they are being observed and the fact that threat item presentation may occur at a higher than normal rate. Conditions which alter arousal level and/or attention to the task may affect c.

1.7 TESTING MILESTONES.

Table 1 shows the milestones for planning and reporting the test and evaluation process.

TABLE 1. TEST AND EVALUATION MILESTONES

Milestone	Date	Responsible Organization
Test Concept/Design Approval	To be determined	AAR-510
TEP Submitted	13 Jul 95	Contractor
Coordinate Airport Sites	Continuous	AAR-510
Test Readiness Review	8 Sep 95	AAR-510
Pilot Test	15 Sep 95	AAR-510, Contractor
TEP Approved	15 Sep 95	AAR-510
Testing Initiated	18 Sep 95	AAR-510, Contractor
Testing Completed	21 Nov 95	AAR-510, Contractor
Test and Evaluation Report	29 Dec 95	AAR-510, Contractor
Evaluation Briefing	29 Dec 95	AAR-510

2. TEST DESCRIPTION.

2.1 TEST PURPOSE.

The purpose of the test is to gather data that will establish an objective baseline of screener IED detection performance for each of the 19 CAT X airports and compare these data across airports to determine any significant difference in performance level.

2.2 SUBJECTS.

The sample will consist of 532 screeners: 28 screeners from each of the 19 CAT X airports. They will be selected in percentages that correspond to the X-ray screening market share held by their employing security company. Each screener will take the IED detection test and a vision test. The screeners will also complete an informed consent form (Appendix B), a personal information questionnaire (Appendix C), and receive test protocol instructions before participating in each test of the study. The test protocols are included in Appendix D.

2.3 TEST ORGANIZATION.

Three test administrators will be required for this study, as shown in Table 2. A detailed description of the duties and schedule of each test administrator is included in Appendix E. There will be three test administrators (numbered 1, 2 and 3 below) who will share various test duties.

TABLE 2. TEST ADMINISTRATORS

Personnel Number	Title	Agency
Test Administrator 1	Test Director	FAA
Test Administrator 2	Test Manager	Galaxy
Test Administrators 1, 2, and 3	IED Detection Tester	FAA and Galaxy
Test Administrators 1, 2, and 3	Vision Tester	FAA and Galaxy

2.4 TEST ORGANIZATION TRAINING.

The test organization will receive training on all tests, procedures, and protocols before conducting the study. The training will cover data collection procedures, methodology, and procedures to be followed in case of an emergency.

2.5 TEST PROCEDURES.

2.5.1 Pilot Study.

To validate appropriate procedures and minimize problems with the testing, a pilot study will be conducted at the FAA Technical Center using local airport personnel as subjects. Two overlapping test cycles will be completed. Eight subjects will be required; four for each test cycle.

2.5.2 Test Protocol.

This protocol describes the manner in which data will be gathered to establish baseline IED detection information. The testing activities at each CAT X airport will require 1.5 days to complete. The test schedule at each airport will consist of seven 3-hour test cycles (see Appendix F). Four screeners will be tested per cycle. It is planned to conduct testing at two adjacent airports during a given work week. The test director will be responsible for ensuring that the screeners are available for each testing cycle. Two rooms will be required at each airport to conduct the testing activities.

A cycle that starts at 8:00 a.m. will proceed as follows. From 8:00 to 8:30, the screeners will receive a briefing and complete an informed consent form and a personal information questionnaire. IED testing will take place from 8:30 to 9:40. Vision testing (see Appendix G) will take place from 0:940 to 10:30. During every test cycle, screeners will receive a 10-minute rest break after approximately each hour of testing.

The IED testing will take place on four computer-based image presentation devices (one device per screener). Before each testing cycle, the test administrator will assign each screener an identification number and will ensure that each test device is serviceable and in the proper configuration. The IED detection test will be run in two separate 30-minute sessions. If a test device malfunctions at any point during testing, the test manager will arrange for maintenance. If it is impossible to rectify the problem immediately, testing on the failed device will be postponed until repairs are made.

Before each test, screeners will be briefed using the instructions contained in Appendix D.

2.5.2.1 IED Testing.

Each screener will be informed that IEDs will be present in some of the bag images presented during the test. The screeners will be unaware of the test bag insertion order. Screeners will carry out normal screening operations and will not receive operational direction from the test administrators. Each computerized test will involve the presentation of 25, black/white, X-ray, CO, test bag images containing simulated IEDs within a total series of 200 black/white X-ray CO bag images. All screeners in the study will view the same 200 test and comparison bag images. Individual screener performance will not be disclosed. To control for presentation order, 14 randomly generated presentation orders for the 200 bags images will be used. Two screeners in each airport sample will receive one of the presentation orders. Each airport sample will receive the same set of 14 presentation orders.

The test administrator will start the test trial and ask the screener to begin the test. The testing device will automatically display the first bag image to the screener. The screener's task is to indicate whether the bag contains a threat image by pressing the appropriate key on the keyboard. Screeners will be given 10 seconds to respond and will be aurally prompted after 6 seconds. At the end of the 10 seconds, the test administrator will urge the screener to make an immediate response. After the screeners have responded to this initial question, screeners will be asked to indicate how confident they were in their response concerning the presence or absence of an IED in the bag image. The confidence rating choices are "very sure" and "not so sure." In combination with the prior yes or no judgment, the following categories can result:

Category 1 - yes, very sure
Category 2 - yes, not so sure
Category 3 - no, not so sure
Category 4 - no, very sure

Screeners will be allowed 5 seconds to make this response. If no response has been given within 5 seconds, the screener will be aurally prompted. The test administrator will prompt the screener to make an immediate response. The testing device will record the response and automatically forward to the next image. The test administrator will stop the test when all bag images have been presented.

2.5.2.2 Vision Testing.

All screeners will be given the Regan High Contrast Acuity Test to ensure that the screeners' vision is within normal range.

2.6 DATA.

2.6.1 IED Detection Test.

The N_h and N_{fa} will be collected for each screener.

2.6.2 Vision Test.

Snellen scores will be collected for each screener using the Regan High Contrast Acuity Chart.

2.7 DATA ANALYSIS.

This section describes the manner in which the data will be analyzed and evaluated to determine airport baggage screener performance to address the two operational issues.

The derived (mean, median, standard deviation) experimental data will be described using descriptive statistics and analyzed using parametric statistics (Analysis of Variance [ANOVA], Regression Analysis, Duncan Test). Sections 2.7.1 through 2.7.3 describe the specific data analyses to be carried out.

2.7.1 Individual and Group Baseline Performance Data.

After completing the IED test trials, the first step in the analysis will be to separate responses to test bags from those to comparison bags and to total these separately for each subject according to the rating response each trial received (see Table 3). The next step will be to convert the raw data into a set of hit and false alarm rates. This will involve starting with the strictest response category (1) and determining P_d and P_{fa} values from the cell data. Two additional P_d and P_{fa} values are determined in a similar manner by collapsing the data across categories 1 and 2 and categories 1, 2, and 3.

TABLE 3. RATING SCALE CALCULATIONS OF P_d AND P_{fa}

Category	1 (High certainty signal)	2	3	4 (High certainty noise)	Total
Threat bags	n_{h1}	n_{h2}	n_{h3}	n_{h4}	N_t
Comparison bags	n_{fa1}	n_{fa2}	n_{fa3}	n_{fa4}	N_C
	$P_{d1}=n_{h1}/N_h$	$P_{d2}=(n_{h1}+n_{h2})/N_h$	$P_{d3}=(n_{h1}+n_{h2}+n_{h3})/N_h$		
	$P_{fa1}=n_{fa1}/N_{fa}$	$P_{fa2}=(n_{fa1}+n_{fa2})/Nf_a$	$P_{fa3}=(n_{fa1}+n_{fa2}+n_{fa3})/Nf_a$		

NOTE: n_{ij} = number of hits or false alarms in each cell

The values of N_h and N_{fa} will be used to derive IED detection performance variables for each screener. The IED derived detection variables are P_d, P_{fa}, d′, and c. For each screener, the N_h will be used to calculate the P_d, as follows:

$$P_d = N_h/N_t$$

Where N_t is the total number of test bag images presented to the screener.

N_{fa} will be used to calculate the P_{fa}, as follows:

$$P_{fa} = N_{fa}/N_C$$

Where N_C is the total number of comparison bags presented.

P_d and P_{fa} will be used to determine a value of c, as follows:

$$\underline{c} = .5(z_{fa} + z_h)$$

Where z_{fa} and z_h are the z-score conversions of number of false alarms and number of hits.

A value of d′ will be calculated from a Receiver Operating Characteristic (ROC) curve plotted for each screener using the rating scale technique applied to the confidence rating data, as depicted in Table 3 (see Swets and Green [1966] for a complete description).

The ROC curve is being used to determine a more stable value of d′ than that determined using the binary "yes-no" response technique. The responses in the yes-no technique are based on a single value of operator response bias. The rating scale technique requires subjects to simultaneously use several response biases and then calculate d′ across these separate biases, thereby stabilizing the determined value of d′.

To obtain the points for the ROC curve, the values of P_d and P_{fa} will be converted into Z_h and Z_{fa} by using the normal tables and plotting against one another. A linear regression will then be determined through the three points. A value of d′ will be determined from the point where the negative diagonal intersects this regression line.

Mean, median, and standard deviation values will be reported for each airport for each of the derived measures.

2.7.2 IED Detection Performance Compared across Airports.

The second issue addressed by the study is determination of any significant difference in IED detection performance across the examined airports. This issue will be addressed using the derived data and parametric statistical analyses.

Four different, single-factor, between-subjects designs will be used to address this issue, as depicted in Table 4. The independent variables will be airport (19 CAT X airports) and the dependent variables will separately be the values of P_d, P_{fa}, d′, and $\underline{c}$ determined for issue 1. A one-way ANOVA will be conducted across airports, according to the ANOVA summary in Table 5, for each of these dependent variables. These ANOVAs will determine if there are significant differences in the values observed for P_d, P_{fa}, d′, and $\underline{c}$ across the 19 different airports. If an overall effect is determined, Duncan post hoc comparisons will be used to isolate significant differences between airports.

TABLE 4. EXPERIMENTAL DESIGN

Airport (U.S. Category X)								
Airport 1 (n = 28)		Airport 2 (n = 28)		Airport n (n = 28)		Airport 19 (n = 28)		
s1	s15	s29	s43	s57	s71	. . .	s505	s519
s2	s16	s30	s44	s58	s72	. . .	s506	s520
s3	s17	s31	s45	s59	s73	. . .	s507	s521
s4	s18	s32	s46	s60	s74	. . .	s508	s522
s5	s19	s33	s47	s61	s75	. . .	s509	s523
s6	s20	s34	s48	s62	s76	. . .	s510	s524
s7	s21	s35	s49	s63	s77	. . .	s511	s525
s8	s22	s36	s50	s64	s78	. . .	s512	s526
s9	s23	s37	s51	s65	s79	. . .	s513	s527
s10	s24	s38	s52	s66	s80	. . .	s514	s528
s11	s25	s39	s53	s67	s81	. . .	s515	s529
s12	s26	s40	s54	s68	s82	. . .	s516	s530
s13	s27	s41v	s55	s69	s83	. . .	s517	s531
s14	s28	s42	s56	s70	s84	. . .	s518	s532
								$N_T = 532$

TABLE 5. ANOVA SUMMARY
(IED DETECTION PERFORMANCE)

Source	df	SS	MS	F	p
A	18	SS_A	MS_A	$MS_A/MS_{S/A}$	-
S/A	513	$SS_{S/A}$	$MS_{S/A}$	-	-
Total	531	SS_{TOT}	-	-	-

NOTE: A = Airport; S = Subjects

2.8 DATABASE MANAGEMENT.

Table 6 shows the database layout for the data to be collected for statistical analyses required to support the evaluation of the baseline data against the COIC.

All data will be retained by the principal investigator for the project for 5 years. In accordance with professional and ethical standards, the principal investigator will maintain separate records of performance data and the names of participating screening personnel. A complete set of data will be provided to the FAA Aviation Security Laboratory for storage as well, without the associated personnel records.

TABLE 6. SCREENER BASELINE DATABASE

Screener ID
Screener Characteristics
Experience
X-ray Equipment Experience
Date
Vision Test Score
IED Detection/Confidence Rating Scores
N_h
Bag Number for Each Hit
N_{fa}
Bag Number for Each False Alarm
N_c
N_i
d'
$\underline{c}$
P_d
P_{fa}

3. PRESENTATION OF OPERATIONAL TEST AND EVALUATION RESULTS.

3.1 IED DETECTION TESTING.

Significant main effects resulting from the ANOVA carried out on P_d, P_{fa}, d′, or c across airports will be reported, along with associated mean and standard deviations for each airport. The results of the post hoc comparisons conducted on any significant main effects will also be reported.

4. REFERENCES.

Green, D.M. and J.A. Swets, Signal Detection Theory and Psychophysics, Wiley, New York, NY, 1966.

APPENDIX A
SIGNAL DETECTION THEORY AND APPLICATION

The Signal Detection Theory (SDT) Paradigm

The Improvised Explosive Device Detection System operation features human operators engaged in tasks to detect an environmental event or signal. SDT is a mathematical representation of human performance in deciding whether or not a signal is present. An operational example of SDT is an airport security guard screening passenger bags for concealed weapons and Improvised Explosive Devices (IEDs).

There are two response categories that represent a screener's detection performance: Yes (a Modular Bomb Set [MBS] signal was present) or No (a MBS signal was not present). There are also two signal presentation states indicating that the MBS signal was present (signal) or absent (noise). A combination of screener responses and the signal state produces a 2 x 2 matrix (figure A-1), generating four classes of operator responses, labeled hits, misses, false alarms, and correct rejections (Wickens 1992).

		State of MBS Image	
		MBS Present	MBS Not Present
Screener Response	Yes	Hit	False Alarm
	No	Miss	Correct Rejection

FIGURE A-1. 2 X 2 MATRIX OF SCREENER RESPONSES AND STATE OF MBS IMAGE

a. A Hit will be recorded when a baggage screener correctly detects an MBS in the scanned baggage.

b. A False Alarm will be recorded when a baggage screener reports a MBS in the scanned baggage when none is present.

As indicated by Wickens (1992), the SDT paradigm assumes that operators perform two stages of information processing in all detection tasks: (1) sensory evidence is aggregated concerning the presence or absence of the signal, and (2) a decision is made about whether this evidence constitutes a signal. According to SDT, external stimuli generate neural activity in the brain. On the average, there will be more sensory or neural evidence in the brain when a signal is present than when it is absent. This neural evidence, X, referred to as the evidence variable, represents the rate of firing of neurons in the brain. The response rate for detecting X increases in magnitude with stimulus (signal) intensity. Therefore, if there is enough neural activity, X exceeds a critical threshold, X_c, and the operator decides "yes." If there is too little, the operator decides "no." Because the amount of energy in the signal is typically low, the average amount of X generated by signals in the environment is not much greater than the average generated when no signals are present (noise). Furthermore, the quantity of X varies continuously, even in the absence of a signal, because of random variations in the environment and the operator's level of neural firing (i.e., the neural "noise" in the operator's sensory channels and brain).

The relationship between the presence and absence of a signal can be seen in the hypothetical noise and signal plus noise distributions contained in figure A-2. The intersection of the two curves represents the location where the probability of a signal equals the probability of noise. The criterion value, X_c, chosen by the operator, is shown by the vertical line. All X values to the right ($X > X_c$) will cause the operator to respond "yes." All X values to the left generate "no" responses.

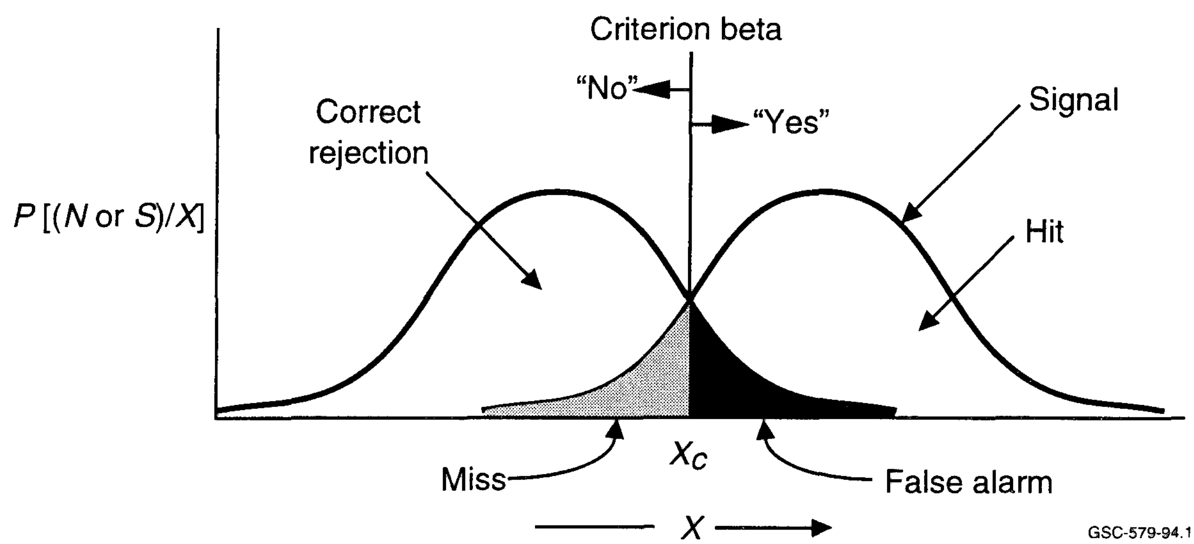


FIGURE A-2. HYPOTHETICAL SDT DISTRIBUTIONS (Wickens 1992)

The different shaded areas represent the occurrences of hits, misses, false alarms, and correct rejections.

Procedures to Calculate SDT Probabilities

a. In SDT, the detection values are expressed as probabilities.

b. The probability of hit (P_h), miss (P_m), false alarm (P_{fa}), and correct rejection (P_{cr}) are determined by dividing the number of occurrences in a cell (figure A-1) by the total number of occurrences in a column.

c. The P_h (also referred as the probability of detection [P_d]) will be calculated by dividing the number of IEDs detected (number of hits) by the total number of hits and misses: $P_m = 1 - P_d$.

d. The P_{fa} will be determined by the number of false alarms divided by the total number of false alarms and correct rejections: $P_{cr} = 1 - P_{fa}$.

Operator Response Criterion

In any signal detection task, operator decision making may be described in terms of an operator response criterion. Operators may use "risky" response strategies by responding yes more often than no. A risky strategy allows operators to detect most of the signals that occur, but also produces many false alarms. Alternatively, operators may use "conservative" strategies, saying no most of the time, making few false alarms, but missing many of the signals.

Different circumstances may require conservative or risky strategies. For example, an appropriate IED detection strategy requires screeners to respond "yes" when there is a question regarding baggage contents. This response may produce false alarms when no threatening objects are present.

One recent parametric measure of response bias is c (Ingham, 1970; Macmillan & Creelman, 1990; Snodgrass & Corwin, 1988). The chief difference between the measure c and its parametric alternative β lies in the manner in which they locate the observer's criterion. Whereas the bias index β locates the observer's criterion by the ratio of the ordinates of the signal-plus - noise (SN) and noise (N) distributions, c locates the criterion by its distance from the intersection of the two distributions measured in z-score units. The intersection defines the point where bias is neutral, and location of the criterion at that point yields a c value of 0. Conservative criteria yield positive c values, and liberal criteria produce negative c values. The measure c is computed as follows:

$$\underline{c} = .5(z_{fa} + z_h) \qquad (2)$$

Sensitivity (d′)

Sensitivity refers to the average amount of operator sensory activity generated by a given signal as compared with the average amount of noise-generated activity (Coren and Ward 1989). Baggage screeners may fail to detect (miss) an IED signal when employing a conservative response criterion. Correspondingly, the signal may be missed because the resolution of the detection process is low in discriminating signals from noise, even if the response criterion is neutral or risky.

The perceptual analog of sensitivity, d′, corresponds to the separation of the means of signal and noise distributions (figure A-2). As the magnitude of the signal increases, the mean of the signal distribution moves to the right. The proportion of signals detected (the P_d) changes as the distance between the signal and noise distributions varies. According to Wickens (1992), if the separation between the distributions is great, sensitivity is great, an operator can readily distinguish a signal plus noise event from a noise only event. Similarly, if the separation between signal and noise is small, d′ measures will be low.

One method of determining d′ is to use an ROC curve determined from a rating scale technique to score screener judgments as to whether IEDs exist in passenger bags passed through an explosive device detection system. When a screener responds as to whether an IED is present within a bag, the response is made in terms of confidence in the response. For example, a four point rating scale may use the following responses:

Certain - signal
Uncertain - signal
Uncertain - noise
Certain - noise

The rating scale technique, in effect, requires the screener to hold several response criteria simultaneously, one criterion for each point on the rating scale. For each criterion, a value of P_d and P_{fa} may be determined. The P_d and P_{fa} values are then converted to z-scores The z-score pairs are then plotted, and a regression line determined. A value of d′ is determined as twice the ordinate value where the regression line intersects the negative diagonal.

The values of P_d and P_{fa} are determined as follows. Using a four point rating scale:

Responses to bags containing IEDs are separated from those to innocent bags and are totaled separately for each subject according to the rating response each trial received (see Table A-2). The next step will is to convert the raw data into a set of hit and false alarm rates. This will involve starting with the strictest response category (1) and determining P_d and P_{fa} values from the cell data. Two additional P_d and P_{fa} values are determined in a similar manner by collapsing the data across categories 1 and 2 and categories 1, 2, and 3.

TABLE A-2. RATING SCALE CALCULATIONS OF P_d AND P_{fa}

	High certainty signal			High certainty noise	
Category	1	2	3	4	Total
Threat bags	n_{h1}	n_{h2}	n_{h3}	n_{h4}	N_T
Comparison bags	n_{fa1}	n_{fa2}	n_{fa3}	n_{fa4}	N_C
	$P_{d1}=n_{h1}/N_h$ $P_{fa1}=n_{fa1}/N_{fa}$	$P_{d2}=(n_{h1}+n_{h2})/N_h$ $P_{fa2}=(n_{fa1}+n_{fa2})/N_{fa}$	$P_{d3}=(n_{h1}+n_{h2}+n_{h3})/N_h$ $P_{fa3}=(n_{fa1}+n_{fa2}+n_{fa3})/N_{fa}$		

Note. n_{ij} = number of hits or false alarms in each cell

APPENDIX B
INFORMED CONSENT

I, ____, have received a briefing by the FAA representative as to the purpose of the FAA study. I fully understand the purpose of the study and have been provided with the opportunity to ask questions of the FAA representative. The FAA representative informed me that the study will require a 30 minute briefing, a 1 hour and 10- minute performance test and a 20-minute vision test.

I understand that this study will impose very little stress. The only stress I may experience in this experiment may be some initial frustration as I learn how to use the testing system. As part of the data analysis, my data will be combined with that of other individuals and I will no longer be identifiable as a participant. I have been informed that my name will remain CONFIDENTIAL.

I have been informed that I have the right to withdraw from the experiment, and that the experiment monitor may terminate my participation in the interest of safety and the experiment. I also certify that I am at least 18 years of age.

I have been informed that if additional details are needed, I may contact any of the test administrators at the airport during the study, or contact James L. Fobes, Ph.D., (609) 485-4944; or Robert L. Malone, (609) 645-0900, upon completion of the study.

Signed: ____

Date: ______ / ______ /______

Witness: ____

Date: ______ / ______ /______

APPENDIX C
FAA

EXPLOSIVE DEVICE DETECTION BASELINE STUDY

SCREENER QUESTIONNAIRE

DATE:______________________ SUBJECT NUMBER:______________________

1. How long have you been a baggage screener?

___________Years ___________Months

2. How long have you been using X-ray equipment to screen baggage?

___________Years ___________Months

3. Circle the highest education level that you have completed.

Elementary School   High School   College or University

APPENDIX D
ED DETECTION TEST PROTOCOL

Read to Screeners:

This is a test of how well X-ray machines can be used for detecting Explosive Devices (EDs). For this activity, we have put X-ray images of passenger bags in this computer. You will view the X-ray images, one bag at a time, and inspect each bag for an ED. The images will be displayed on the video monitor.

INSTRUCTIONS TO SUBJECTS

After you have verified that your subject number appearing on you monitor is correct, press the "ENTER" key on the keyboard to start the practice test. Initially, you will be given about 10 images to practice on before beginning the real test. Once you see the first X-ray image appear on the monitor screen, your task will be twofold.

First, you will respond to the question: "Is there an ED in this X-ray image?"

To indicate "yes" press key labeled "yes" on the keyboard.

To indicate "no" press key labeled "no" on the keyboard.

Please note the label at the bottom center of monitor frame which is there as a reminder of the question to be answered for each image you view: "Is there an ED in the above X-ray Image?" Should you happen to press any key other than the "yes" or "no" key, however, an error message will appear on the monitor telling you that you have, in fact, pressed the wrong key.

It is important that you answer each question to the best of your ability. It is just as important for you to say "no" when you do not see an ED as it is to say "yes" when you do see an ED.

You will have a total of 10 seconds in which to make your response, however, you should answer as quickly as possible. Once you have made your response, you will not allowed to go back and change it.

If you have not answered the question, and 6 seconds have elapsed, an audible alarm will sound, alerting you that you have only 4 seconds left in which to answer the question.

After the full 10 seconds have elapsed and you still have not responded, the image will disappear from the monitor screen and will be replaced by a prompt which advises you that, "Time has expired, please answer."

After you have responded to the first question, a second question will be displayed on the video monitor:

"How sure are you?"
yes = very sure   no = not so sure

Again, you are asked to respond by pressing the "yes" or "no" key on the keyboard or keypad, regardless of whether your answer to the first question was "yes" or "no." In either case, you will use the same two keys that you used to respond to the first question.

You will have 5 seconds to make this second response. If, after 5 seconds have elapsed, and you have not yet answered this second question, the question will disappear from the monitor screen and the same prompt that appeared earlier will reappear telling you "Time has expired, please answer." Please note that the system will not forward to the next image until you have made your response.

You will repeat this same procedure for each image until all images have been viewed. A message will then appear on the monitor indicating that the test has ended. The test should take a little over an hour to complete.

Do you have any questions?

Again, thank you for your willingness to participate in this test and you may now begin by pressing the "ENTER" key on your keyboard to start the practice test. When you have completed the practice test, you may then press the "ENTER" key again to start the real test.

APPENDIX E
EXPLOSIVE DEVICE DETECTION BASELINE STUDY
STATEMENT OF DUTIES FOR ADMINISTRATIVE PERSONNEL

Test Director:

Responsible for directing and overseeing all test activities and personnel.

Federal Aviation Administration

Test Manager:

Ensures screeners are greeted and completes required administration.
Conducts briefs and debriefs.
Manages daily test activities and responsible for starting and stopping test sequence as required.
Plans for and directs contingency activities.
Liaisons with security company administrative personnel.
Liaisons with airline personnel.
Ensures the presence of the proper screener in the required location at the required time.
Escorts screeners between security checkpoint and training and testing rooms.
Resolves any encountered problems with the test director.
Executes all required logistical activities as required.

Galaxy Scientific Corporation

IED Detection Tester

Ensures that the screener's number is recorded on test forms.
Ensures that test forms are given to Galaxy Scientific Corporation personnel.
Briefs and debriefs screeners regarding Improvised Explosive Device (IED) detection test.
Provides guidance and assistance to screeners on matters pertaining to conduct of test.
Administers IED detection test.

Federal Aviation Administration
Galaxy Scientific Corporation

Vision Tester

Ensures that the screener's number is recorded on test forms.
Ensures that test forms are given to Galaxy Scientific Corporation personnel.
Ensures that each screener completes a informed consent form and questionnaire.
Conducts visual acuity tests on screeners using the required test forms.

Federal Aviation Administration
Galaxy Scientific Corporation

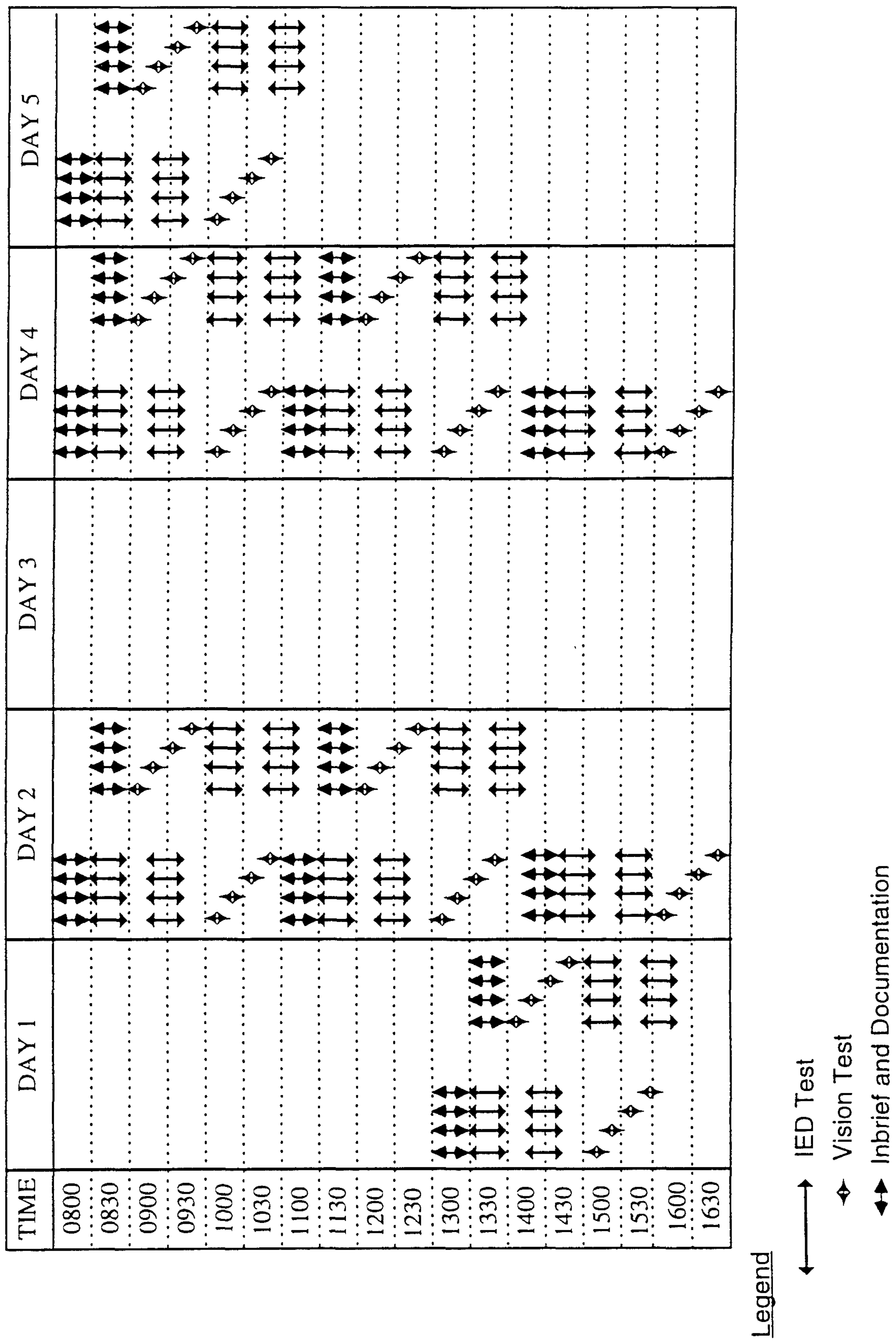
TEST ADMINISTRATOR SCHEDULE

Day	Activity	Time	Location	Administrator
1	Inbriefing	1300-1330	Testing Room 1	Test Manager
		1330-1400	Testing Room 1	Test Manager
1	IED Detection Testing	1330-1440	Testing Room 2	IED Detection Tester
		1500-1610	Testing Room 2	IED Detection Tester
1	Vision Testing	1400-1500	Testing Room 1	Vision Tester
		1500-1600	Testing Room 1	Vision Tester
2	Inbriefing	0800-0830	Testing Room 1	Test Manager
		0830-0900	Testing Room 1	Test Manager
		1100-1130	Testing Room 1	Test Manager
		1130-1200	Testing Room 1	Test Manager
		1410-1440	Testing Room 1	Test Manager
2	IED Detection Testing	0830-0940	Testing Room 2	IED Detection Tester
		1000-1110	Testing Room 2	IED Detection Tester
		1130-1240	Testing Room 2	IED Detection Tester
		1300-1410	Testing Room 2	IED Detection Tester
		1440-1550	Testing Room 2	IED Detection Tester
2	Vision Testing	0900-1000	Testing Room 1	Vision Tester
		1000-1100	Testing Room 1	Vision Tester
		1200-1300	Testing Room 1	Vision Tester
		1300-1400	Testing Room 1	Vision Tester
		1550-1650	Testing Room 1	Vision Tester

4	Inbriefing	0800-0830	Testing Room 1	Test Manager
		0830-0900	Testing Room 1	Test Manager
		1100-1130	Testing Room 1	Test Manager
		1130-1200	Testing Room 1	Test Manager
		1410-1440	Testing Room 1	Test Manager
4	IED Detection Testing	0830-0940	Testing Room 2	IED Detection Tester
		1000-1110	Testing Room 2	IED Detection Tester
		1130-1240	Testing Room 2	IED Detection Tester
		1300-1410	Testing Room 2	IED Detection Tester
		1440-1550	Testing Room 2	IED Detection Tester
4	Vision Testing	0900-1000	Testing Room 1	Vision Tester
		1000-1100	Testing Room 1	Vision Tester
		1200-1300	Testing Room 1	Vision Tester
		1300-1400	Testing Room 1	Vision Tester
		1550-1650	Testing Room 1	Vision Tester
5	Inbriefing	0800-0830	Testing Room 1	Test Manager
		0830-0900	Testing Room 1	Test Manager
5	IED Detection Testing	0830-0940	Testing Room 2	IED Detection Tester
		1000-1110	Testing Room 2	IED Detection Tester
5	Vision Testing	0900-1000	Testing Room 1	Vision Tester
		1000-1100	Testing Room 1	Vision Tester

APPENDIX F
TEST SCHEDULE

AIRPORT BASELINE STUDY SCHEDULE



APPENDIX G

REGAN HIGH CONTRAST ACUITY CHART

SCORE SHEET
SERIES 1

Chart A - 96% Contrast

Patient Name .... Date ....

Left Eye

Z	R	D	O	V	C	N	S	1
H	R	V	C	O	S	K	Z	2
N	D	C	O	H	R	V	S	3
K	V	R	Z	C	O	H	S	4
Z	N	V	K	D	S	O	R	5
D	C	R	V	H	N	Z	K	6
O	S	K	C	V	R	Z	N	7
S	N	H	K	C	D	V	O	8
N	R	D	C	O	K	S	Z	9
V	H	C	O	R	Z	D	N	10
H	R	O	S	C	V	K	N	11

Right Eye

Z	R	D	O	V	C	N	S	1
H	R	V	C	O	S	K	Z	2
N	D	C	O	H	R	V	S	3
K	V	R	Z	C	O	H	S	4
Z	N	V	K	D	S	O	R	5
D	C	R	V	H	N	Z	K	6
O	S	K	C	V	R	Z	N	7
S	N	H	K	C	D	V	O	8
N	R	D	C	O	K	S	Z	9
V	H	C	O	R	Z	D	N	10
H	R	O	S	C	V	K	N	11

Left Eye

Number of Errors	Line Number	Score

Right Eye

Number of Errors	Line Number	Score

It is important to urge the patient to guess each letter, even when uncertain.

Mark each error by crossing out each letter missed.