SAND2015-8539C

# Enhancing Human Performance at the Airport Security Checkpoint:

## Human Factors Research at the Transportation Security Administration

Bonnie Kudrick, TSA

Ann Speed, SNL

Transportation Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

# Agenda

- Human Factors Overview
- Human Factors at TSA
- Overview of Current and Future Human Factors Projects
- Example Data Collection

# HUMAN FACTORS OVERVIEW

# What are the functions of Human Factors?

## Human Factors

- Human Factors (HF) - studies the ways humans relate to the world around them.
- HF accounts for user capabilities and limitations while ensuring tasks, functions, information, processes, procedures and the environment suit the user.

## Cognition

- Visual Attention
- Divided Attention versus Sustained Attention
- Working Memory characteristics
- Decision making processes

## Usability

- How efficiently can users access system functionality to meet their objectives?
- User-centered design
  - iterative process
  - incorporates feedback about usability throughout design, development, and test.

## Performance Optimization

- For both employees and the organization
- Includes transfer of training to the job
- Effective learning is the success measure for the employee and the organization

**TSA's Human Factors engineers provide input on how to improve operational performance with the end user in mind.**



Transportation Security Administration

OSC OFFICE OF SECURITY CAPABILITIES

# Human Factors Example
# Three Mile Island

The accident began with failures in the non-nuclear secondary system, followed by a stuck-open pilot-operated relief valve in the primary system, which allowed large amounts of nuclear reactor coolant to escape. The mechanical failures were compounded by the **initial failure of plant operators to recognize the situation as a loss-of-coolant accident due to inadequate training and human factors**, such as human-computer interaction design oversights relating to ambiguous control room indicators in the power plant's user interface.

# HUMAN FACTORS AT TSA

# HFE Capabilities

The HF team executes data collections in the field and conducts human factors research to gather a multitude of user feedback and information.

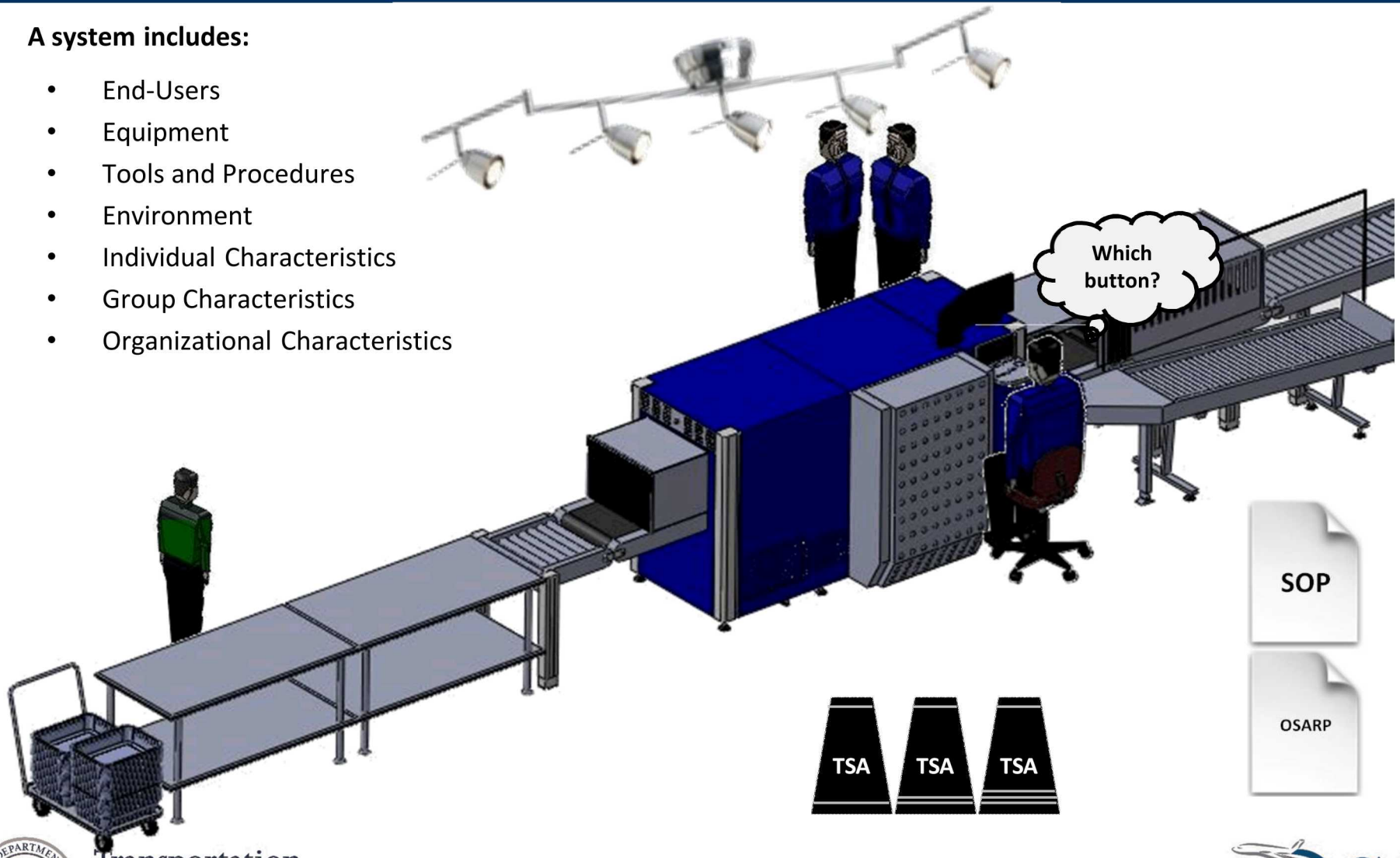
| Type of Data Collected | HF Expertise |
| --- | --- |
| • Eye tracking metrics<br>• Cognitive burden<br>• Optimal duty cycle times for TSO's<br>• Aptitudes for specialization<br>• Best practices for communicating SOP's in the field<br>• Feedback and usability metrics for technologies<br>• Performance metrics | • Human Factor SME support<br>• Data Collection support<br>• Qualitative Analysis<br>  • Focus Groups<br>  • Surveys<br>  • Questionaries'<br>  • Interviews<br>• Quantitative Analysis<br>• Evaluation of current systems/tools to recommend training solutions<br>• Document / Requirements Reviews<br>• Tool creation |

**Understanding perceptual capabilities, cognition, image interpretation, and decision making strategies will inform technology decisions, procedures, and training.**

# Human Factors at the Checkpoint

**A system includes:**

- End-Users
- Equipment
- Tools and Procedures
- Environment
- Individual Characteristics
- Group Characteristics
- Organizational Characteristics




Transportation Security Administration



OSC
OFFICE OF SECURITY CAPABILITIES

# CURRENT HUMAN FACTORS PROJECTS

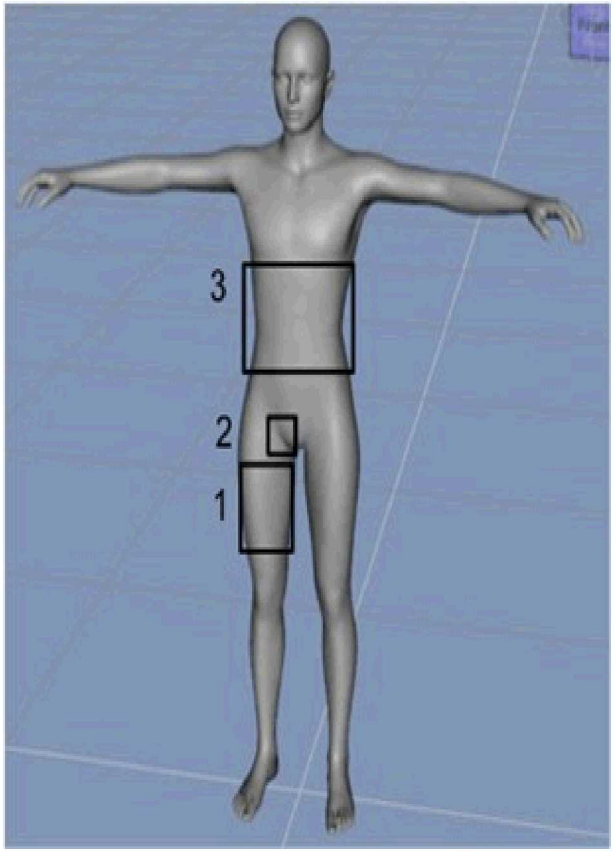
# Pat-down Accuracy Training Tool (PATT)

## Johns Hopkins Applied Physics Lab

# Pat-down Accuracy Training Tool (PATT)
## Johns Hopkins Applied Physics Lab

The pressure sensors are specifically designed to provide TSOs with objective feedback regarding their ability to apply the appropriate amount of pressure and coverage needed to detect prohibited items during a pat-down exam.

Introducing objective training will provide consistency to the process, alleviating some of the stress of conducting pat downs. PATT creates a means to objectively assess if a TSO is applying the appropriate amount of pressure to sufficiently 'clear' a body during a pat-down exam. PATT can also indicate if the proper areas have been covered with an overlapping technique.



# Pat-down Accuracy Training Tool (PATT)
## Johns Hopkins Applied Physics Lab

- Provides TSOs with objective feedback regarding amount of pressure and coverage
- Developed using 3D printer, covered in neoprene, with embedded pressure sensor mats
- Developed pressure standards
- Three demographics
  - Standing male
  - Standing female
  - Seated female
- Data collected and methods
  - Subjects
    - 56 TSOs
  - Data collected
    - TSOs followed the pat-down standard operating procedure and pressure and coverage was measured

# X Ray Scanner
## Kedlin Company

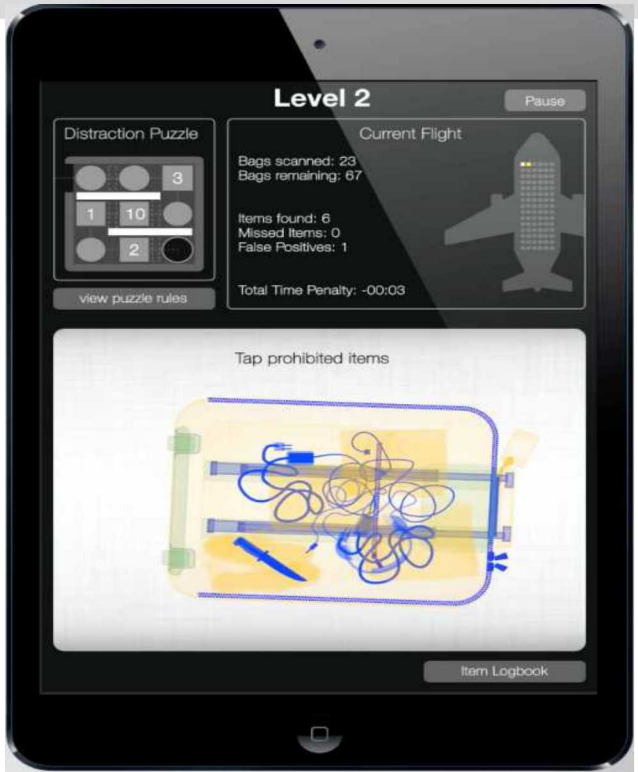
- Based on a app “Airport Scanner” (commercially available in Apple iTunes App store)
- Power is in the amassed data
  - As of 9/4/15 Kedlin Company has collected 2.46 billion trials and has had 8.8 million installs.
  - Data from the general population will be used to supplement the data collected from TSOs.

- Provides a cost effective flexible platform for assessing TSO performance that is not dependent on actual X-ray equipment.
- Linked with a web-based data analysis toolset combined with a massive existing dataset.
- The contractor will conduct analysis on performance data such as screening performance, TSO attributes, and other defined relevant data.


Transportation Security Administration

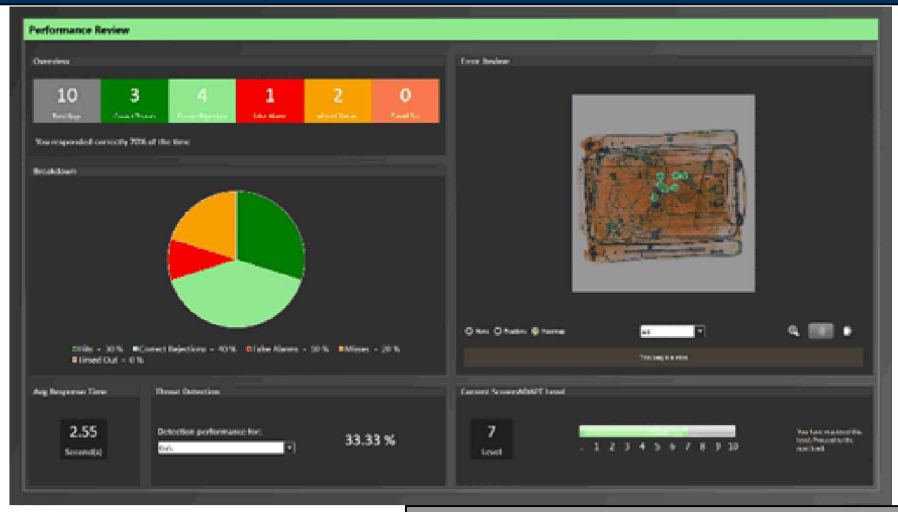
# X Ray Scanner
## Kedlin Company

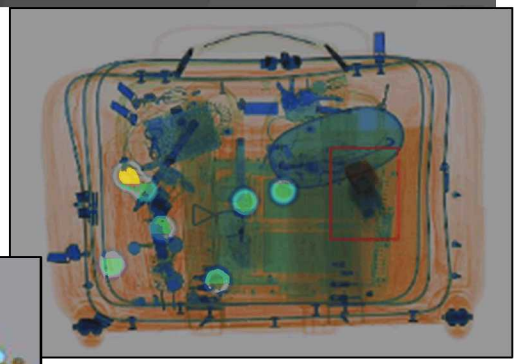
- Airport Scanner modified to allow presentation of more professional “look and feel” app to enable data collection on human performance for visual search and screening pattern recognition.
- Provides a cost effective flexible platform for assessing TSO performance that is not dependent on actual X-ray equipment.
- Linked with a web-based data analysis toolset combined with a massive existing dataset.

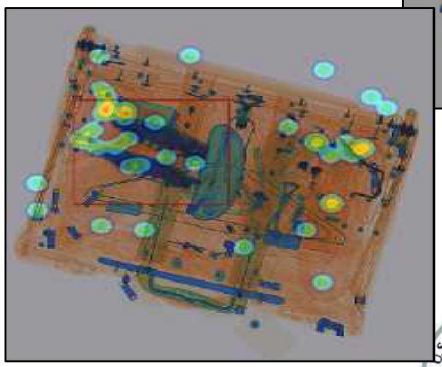




# ScreenADAPT
## Design Interactive, LLC

- X-ray image analysis training system
- Uses eye tracking for:
  - Real-time feedback
  - Customizing training
- Data collected and methods
  - Methods
    - Training effectiveness evaluation was conducted using a standard pre-test-post-test control group design
  - Subjects
    - 115 TSOs participated
    - Four and a half hours of training across five consecutive work days
  - Data collected
    - Image analysis performance (accuracy)
    - Visual search performance (scan patterns)
    - Participant subjective ratings of training value, effectiveness





Transportation Security Administration

# Passenger Threat Detection Resolution (PTDR)
## Sandia National Laboratories

- Determines cognitive and psychological characteristics predictive of successful performance on non-X-Ray checkpoint duties
  - Document checker, divest officer, AIT/WTMD, pat-down
  - Possibly inform selection, hiring, specialization, training
- Data collected and methods
  - Subjects
    - Approximately 390 TSOs from eight airports around the country
  - Data
    - Responses to 33 personality and cognitive measures (25 normed and validated; >1,000 individual items per TSO)
    - TSA-measured job performance on five duties of interest
    - Collected via iPads and Survey Monkey
- Data analysis
  - Factor analysis, principle components analysis, structural equation modeling


Transportation Security Administration

# System Field Evaluation (SFE)
## TSA OSC Human Factors Engineering Team

The Systems Field Evaluation was developed to **help TSA understand both the intended and unanticipated consequences of policy and decision making**.

### Evaluation Background

**1. Gathered Subject Matter Expert Insights**
Met with HQ / Field leadership across divisions to collect their views on pathways for performance improvement

**2. Created Metrics Inventory**
Created initial centralized collection of major HQ and Field metrics

**3. Analyzed Metrics**
Used system dynamics modeling to identify high impact metrics and hypotheses for system improvement

**Systems Field Evaluation**
Facilitated working sessions with Field staff and management to understand high-leverage metric relationships and how they impact decision-making

### Evaluation Objectives

- Better understand operational pressures in the Field and key performance metrics
- Clearly define current system performance and relationships, including RBS technologies and capabilities
- Identify high-leverage points of intervention (metrics that have the greatest impact on the system) to target performance improvement

Transportation Security Administration

# System Field Evaluation (SFE)
## TSA OSC Human Factors Engineering Team

### Data Collection Framework – Six Category X airports

| | |
|---|---|
| **Participants** | The Systems Field Evaluation Team held small activity groups composed of approximately five officers of the same rank, including Transportation Security Officers (TSOs), Lead TSOs (LTSOs), and Supervisory TSOs (STSOs). To encourage honesty, all airport and participant names were kept anonymous. |
| **Session Timing** | Each session was dedicated to eliciting information (SOP deviations and causal factors for those deviations) related to a specific checkpoint screening function and lasted approximately 75 minutes. |
| **Data Collected** | Officers identified deviations and the causes of those deviations in checkpoint screening procedures for the following positions: Advanced Imaging Technology (AIT), Divest Officer (DO), Dynamic Officer, Travel Document Checker (TDC), Walk Through Metal Detector (WTMD), and X-Ray. |
| **Stats** | Over 450 officers were interviewed over the course of this evaluation from across six different airports. Officers were a mix of part-time and full-time employees and ranged from highly experienced to newly hired. |

Transportation Security Administration

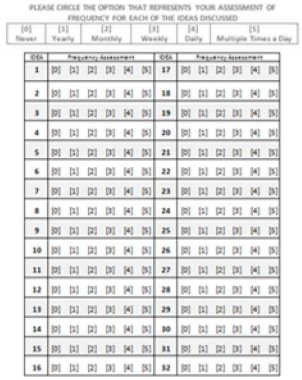
# System Field Evaluation (SFE)
## TSA OSC Human Factors Engineering Team

The team walked the officers through the following structured facilitation framework to identify checkpoint deviations and the causes of those deviations:

### Part 1: Deviations

**Step 1 - Individual**
**Identify Deviations**


**Step 2 - Individual**
**Score Frequency of Deviations**



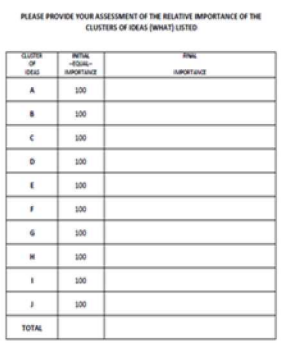
PLEASE CIRCLE THE OPTION THAT REPRESENTS YOUR ASSESSMENT OF FREQUENCY FOR EACH OF THE IDEAS DISCUSSED

| [0] Never | [1] Yearly | [2] Monthly | [3] Weekly | [4] Daily | [5] Multiple Times a Day |
|---|---|---|---|---|---|

| IDEA | Frequency Assessment | IDEA | Frequency Assessment |
|---|---|---|---|
| 1 | [0] [1] [2] [3] [4] [5] | 17 | [0] [1] [2] [3] [4] [5] |
| 2 | [0] [1] [2] [3] [4] [5] | 18 | [0] [1] [2] [3] [4] [5] |
| 3 | [0] [1] [2] [3] [4] [5] | 19 | [0] [1] [2] [3] [4] [5] |
| 4 | [0] [1] [2] [3] [4] [5] | 20 | [0] [1] [2] [3] [4] [5] |
| 5 | [0] [1] [2] [3] [4] [5] | 21 | [0] [1] [2] [3] [4] [5] |
| 6 | [0] [1] [2] [3] [4] [5] | 22 | [0] [1] [2] [3] [4] [5] |
| 7 | [0] [1] [2] [3] [4] [5] | 23 | [0] [1] [2] [3] [4] [5] |
| 8 | [0] [1] [2] [3] [4] [5] | 24 | [0] [1] [2] [3] [4] [5] |
| 9 | [0] [1] [2] [3] [4] [5] | 25 | [0] [1] [2] [3] [4] [5] |
| 10 | [0] [1] [2] [3] [4] [5] | 26 | [0] [1] [2] [3] [4] [5] |
| 11 | [0] [1] [2] [3] [4] [5] | 27 | [0] [1] [2] [3] [4] [5] |
| 12 | [0] [1] [2] [3] [4] [5] | 28 | [0] [1] [2] [3] [4] [5] |
| 13 | [0] [1] [2] [3] [4] [5] | 29 | [0] [1] [2] [3] [4] [5] |
| 14 | [0] [1] [2] [3] [4] [5] | 30 | [0] [1] [2] [3] [4] [5] |
| 15 | [0] [1] [2] [3] [4] [5] | 31 | [0] [1] [2] [3] [4] [5] |
| 16 | [0] [1] [2] [3] [4] [5] | 32 | [0] [1] [2] [3] [4] [5] |

**Step 3 - Group**
**Group Deviations**


**Step 4 - Individual**
**Score Deviation Groupings**



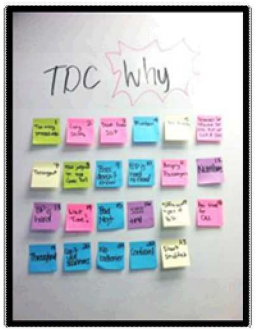
PLEASE PROVIDE YOUR ASSESSMENT OF THE RELATIVE IMPORTANCE OF THE CLUSTERS OF IDEAS (WHAT) LISTED

| CLUSTER OF IDEAS | INITIAL -EQUAL- IMPORTANCE | FINAL IMPORTANCE |
|---|---|---|
| A | 100 | |
| B | 100 | |
| C | 100 | |
| D | 100 | |
| E | 100 | |
| F | 100 | |
| G | 100 | |
| H | 100 | |
| I | 100 | |
| J | 100 | |
| TOTAL | | |

**Step 5 - Group**
**Discuss Deviation Groupings**


### Part 2: Causal Factors

**Step 6 - Individual**
**Identify Causal Factors**




**Step 7 - Group**
**Group Causal Factors**


**Step 8 - Individual**
**Score Causal Factor Groupings**

PLEASE PROVIDE YOUR ASSESSMENT OF THE RELATIVE IMPORTANCE OF THE CLUSTERS OF REASONS (WHY) LISTED

| CLUSTER OF REASONS | INITIAL -EQUAL- IMPORTANCE | FINAL IMPORTANCE |
|---|---|---|
| A | 100 | |
| B | 100 | |
| C | 100 | |
| D | 100 | |
| E | 100 | |
| F | 100 | |
| G | 100 | |
| H | 100 | |
| I | 100 | |
| J | 100 | |
| TOTAL | | |

**Step 9 - Group**
**Discuss Causal Factor Groupings**

# System Field Evaluation (SFE)
## TSA OSC Human Factors Engineering Team

The Systems Field Evaluation was developed to **help TSA understand both the intended and unanticipated consequences of policy and decision making**.

### Evaluation Background

**1. Gathered Subject Matter Expert Insights**
Met with HQ / Field leadership across divisions to collect their views on pathways for performance improvement

**2. Created Metrics Inventory**
Created initial centralized collection of major HQ and Field metrics

**3. Analyzed Metrics**
Used system dynamics modeling to identify high impact metrics and hypotheses for system improvement

**Systems Field Evaluation**
Facilitated working sessions with Field staff and management to understand high-leverage metric relationships and how they impact decision-making

### Evaluation Objectives

- Better understand operational pressures in the Field and key performance metrics
- Clearly define current system performance and relationships, including RBS technologies and capabilities
- Identify high-leverage points of intervention (metrics that have the greatest impact on the system) to target performance improvement

Transportation Security Administration

# FUTURE HUMAN FACTORS PROJECTS

# Proposed Human Factors Initiatives

GOAL: To identify ways to improve operational efficiencies while enhancing security in complex, difficult, and error-prone visual search tasks for both checked and passenger checkpoint screening.

The Human Factors Team is currently reviewing 18 RFI responses, including:

**Satisfaction of Search**

- In radiology, data indicate that detection of one target can reduce detection for a second target.
- Does a Transportation Security Officer's (TSO's) chance of finding one prohibited item (e.g. a true threat) decline if the TSO has already identified another prohibited item (e.g. a water bottle) in the same bag?

**Morphing of Threats and Distractors**

- How do humans respond in a search task where the specific forms of targets and distracting items change over time even if the overall task does not?
- How can TSA train TSOs to minimize the consequences of threat evolution?

**Use of Auto Detect/Auto Assist**

- Evolving technology increasingly uses algorithms to help identify threats and provides bounding boxes on the monitor.
  - Do bounding boxes assist or distract the TSO?
  - What is the optimal presentation of a threat?

# EXAMPLE DATA COLLECTION: PERFORMANCE DECREMENT

# Determining the Optimal Time on X-ray Analysis for Transportation Security Officers

Ann Speed, Austin Silva, Derek Trumbo, David Stracuzzi, Christina Warrender, Michael Trumbo, Kristin Divis
Sandia National Laboratories

Transportation Security Administration

# Experimental Questions

Current duty cycle is 30 minutes

- What happens to performance when TSOs interrogate images for 2 hours?
- Are there individual differences?

# Experimental Design

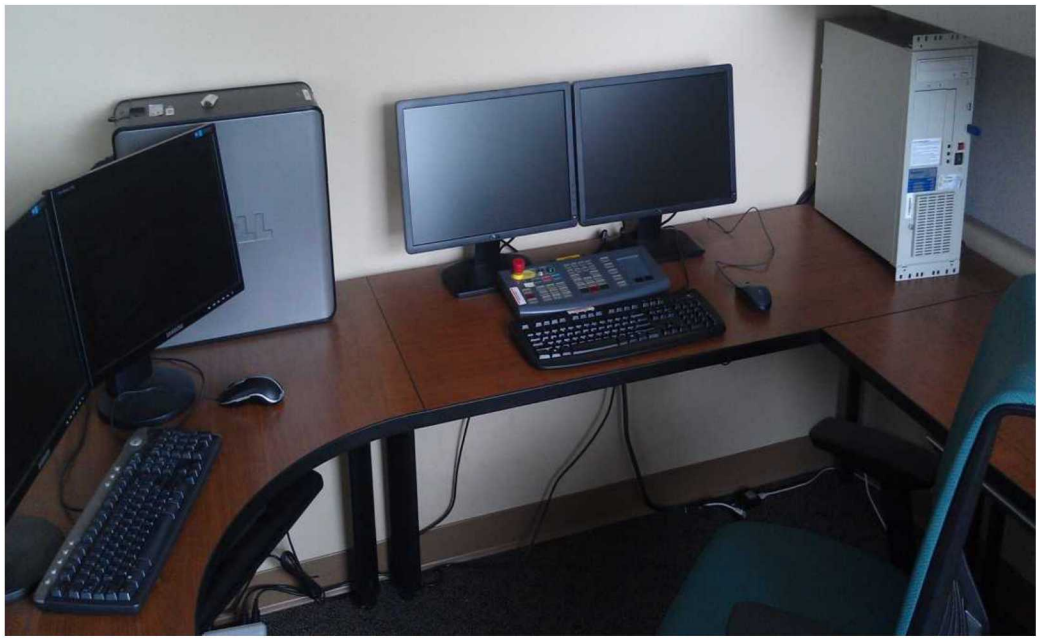
| SOP/Belt (Between Ss) | Threat Type (Within Ss) | 10-minute Epoch (Within Ss) |
|---|---|---|
| TSA Pre✓® - static | Clear | 1-12 |
| | Threat | 1-12 |
| TSA Pre✓® - continuous | Clear | 1-12 |
| | Threat | 1-12 |
| Standard - static | Clear | 1-12 |
| | Threat | 1-12 |
| Standard - continuous | Clear | 1-12 |
| | Threat | 1-12 |

- Stimuli:
  - TSA-generated images of 1,000 carryon items, displayed to TSOs on Sandia-created emulator
  - 99 with threats, same 99 imaged without threats
    - Called Threat and Cleared Threat bags
    - These are the bags used to calculate d prime and response bias
  - 802 clear items
  - Contents (e.g., presence of oversized LGAs) also controlled for, matched stream of commerce

# Stimulus generation

- Generated either 31 or 32 image products for each of the 1,327 .rcf files TSA provided
  - Bounding box images (if generated by Rapiscan)
  - Normal view images
  - Black and white
  - Crystal clear
  - High penetration
  - Inorganic materials
  - Organic materials
  - Inverse colors
  - Variable color
  - Variable density

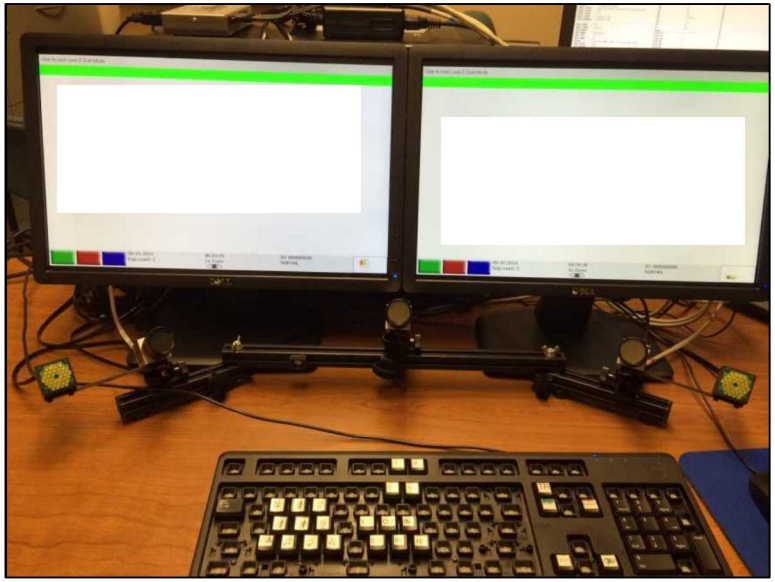
- 62 or 64 total image products per .rcf file (top and side view) captured for a total of **83,624** images

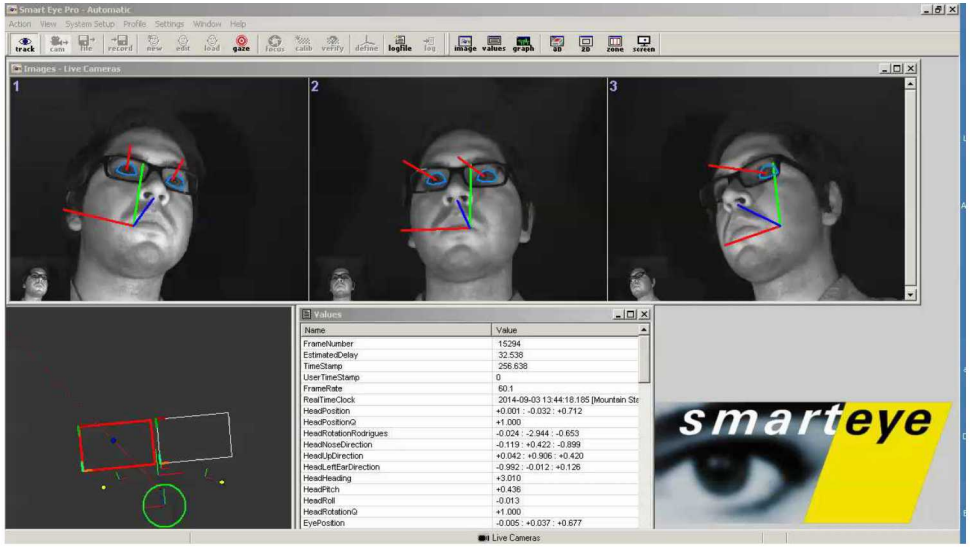


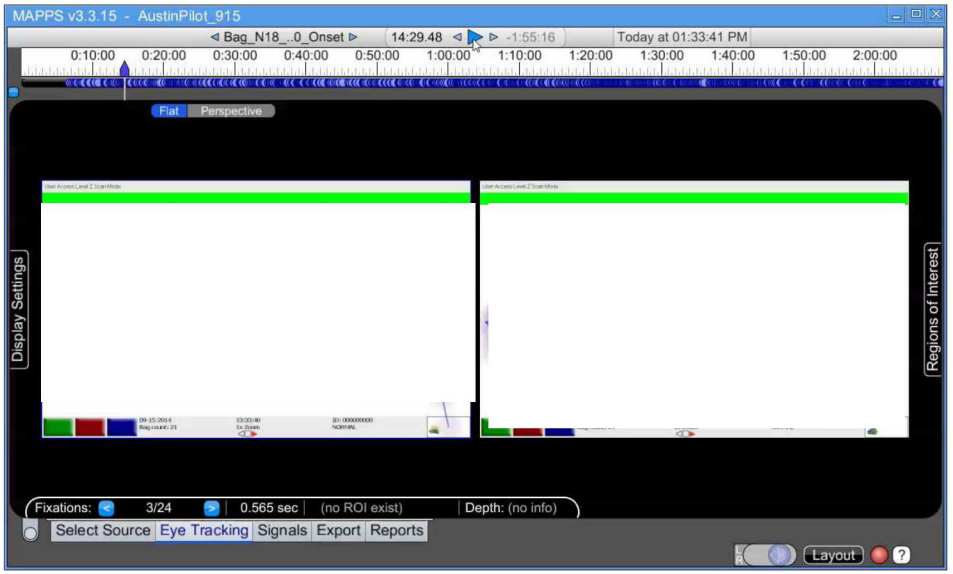
# Three methods of image validation

- Filenames
  - All image products were present for all images
- Hash value
  - No two images across the 83,624 that we generated were identical
- Pixel values of .png versions of the original .rcf files compared to pixel values of normal color .png files SNL captured on emulator
  - Battelle.png file = Sandia-generated .png with the corresponding filename captured using the emulator

# The System



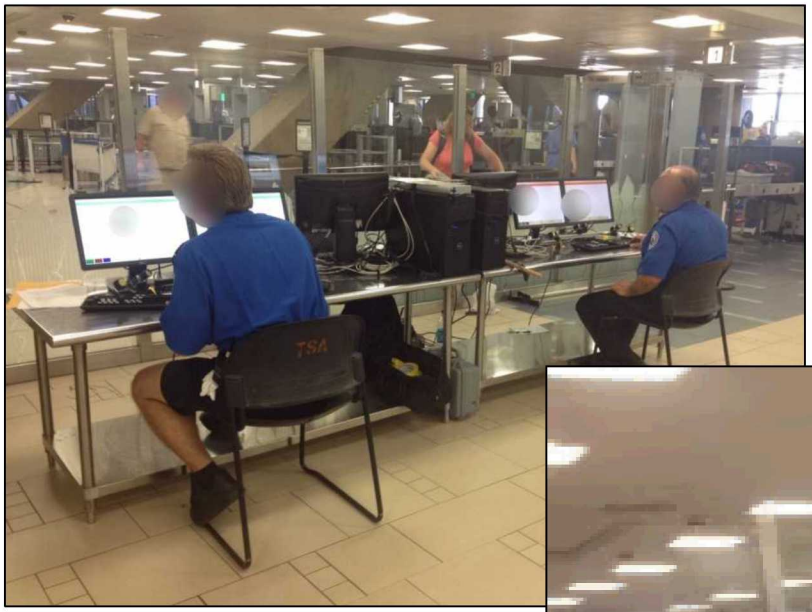


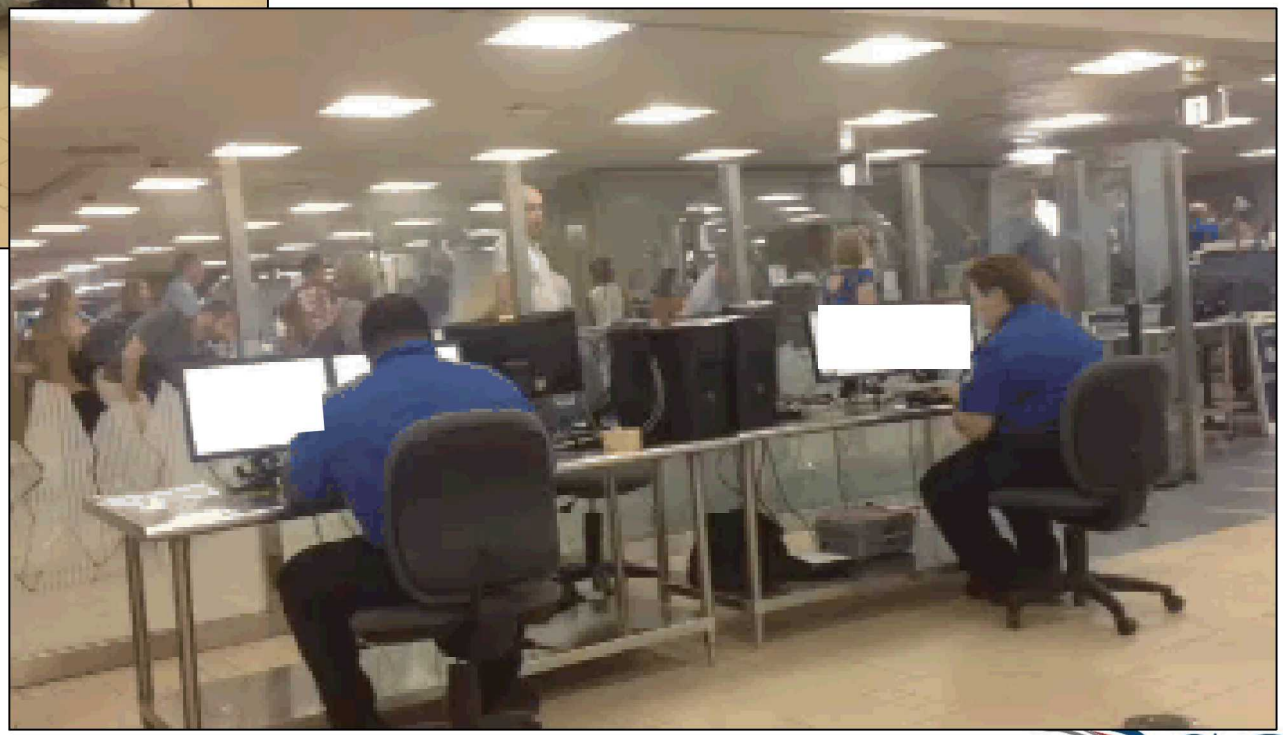




- Monitors identical to those used at checkpoint
- Three infrared eye tracking cameras from SmartEye
- MAAPS eye tracking data analysis software
- Located at the checkpoint
- Emulator captured all user interactions at nanosecond resolution

OSC
OFFICE OF SECURITY CAPABILITIES

# One of the setups



Transportation Security Administration

OSC
OFFICE OF SECURITY CAPABILITIES

# Dependent Variables & Procedure

- All calculated as a function of 5-minute epoch:
  - Pd, Pfa, d', response bias
  - Decision Time
  - Eye Tracking
  - Calculated variables – including d', c, search time consistency
  - Image product use (e.g., order of image manipulation tools, which tools selected, eye tracking patterns associated with each bag, etc.)
- General procedure
  - For each bag, TSO had to clear any bounding boxes, make a decision about the bag itself, indicate the number of threats and benign prohibited items they detected in each bag
  - 187 TSOs analyzed images for 2 hours with no breaks followed by a general cognitive battery

# Resulting dataset

- 187 Subjects across 6 airports
  - 90 female
  - Average age – 41.5
  - Average years experience as TSO – 7
- 2 Hours of main Baggage Screening task
  - 85 to 1000 bags interrogated in 2 hours
  - Mean = 467, SD = 201.5
  - Total: 87,438 observations contributing to each behavioral DV
  - Between 5831 and 8384 observations per epoch
- 45 minutes of domain-general visual cognitive battery
- Eye tracking (60Hz) and user interaction log
  - Over 80 million data points of eye tracking
  - Terabytes of human data

# Data analysis

- Behavioral data
  - Multilevel models for each DV
  - Covariates
    - Multilevel models – each DV
    - Multiple linear regression – random slope and intercepts from primary multilevel models
- Eye tracking analysis
  - Time to first fixation
  - Types of errors
  - Scan patterns
  - Etc.
  - Relationship between eye tracking on general cognitive battery and bag search task
- Machine learning analysis
  - Combining multimodal data stream into one
  - Identifying features predictive of TSO performance

# Questions