Development and Analytic Process Used to Create a 3-Dimensional Graphical User Interface System for Baggage Screening

15 May 2020

Transportation Security Administration

RCA | REQUIREMENTS & CAPABILITIES ANALYSIS

Presenters

Charles "Chas" McKee
Program Manager
Global Systems Technologies

Simone McKnight (Erchov), Ph.D
Sr. Human Factors Engineer
Global Systems Technologies

Transportation Security Administration

RCA | REQUIREMENTS & CAPABILITIES ANALYSIS

Presentation Overview

Overview

Background

Research Process

Supplementary Content

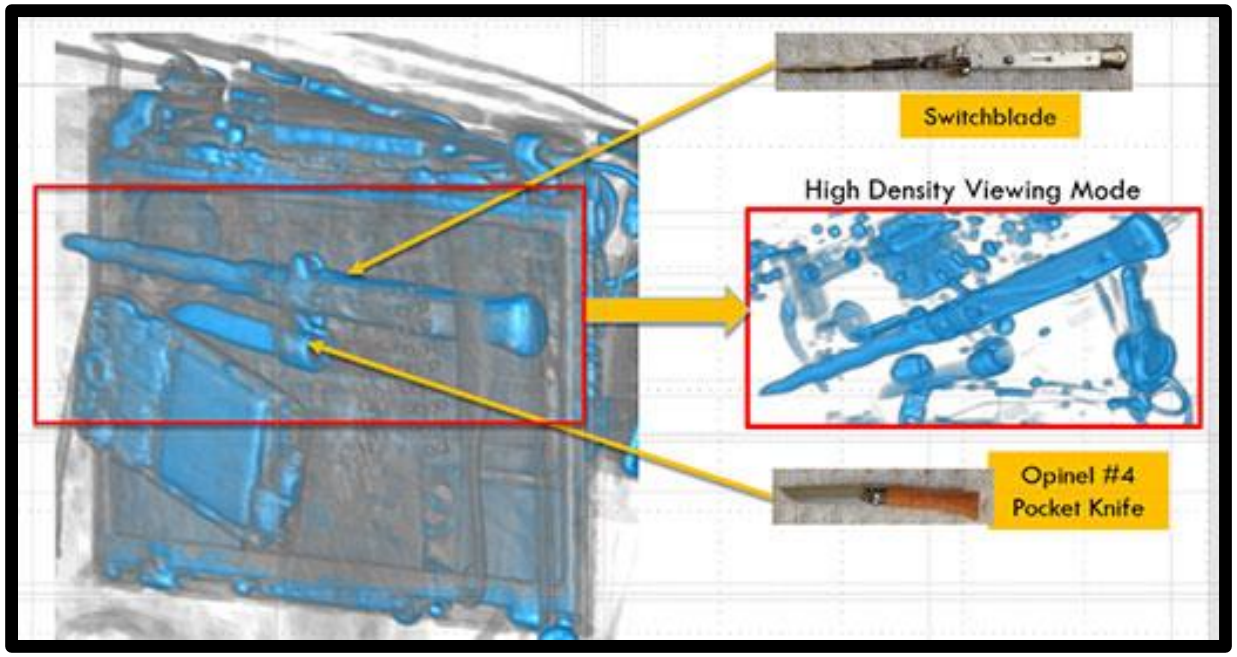
Purpose & Goal

- ***The inclusion of Human Factors (HF) research principles can help support more efficient, effective, and sustainable outcomes for both institutional and technological programs at TSA***

- **Purpose**: Review the HF research process for developing a common graphical user interface for CT systems (CT-APSS CGUI) and discuss broader implications for TSA research procedures

- **Goal:** Communicate a general method and set of considerations for incorporating HF research into the TSA project life cycle.

Disclaimer



- ***We have included content approved for public consumption***
  - ***The results presented are limited in the depth of detail we can discuss about the data***
  - ***Several images have been altered according to security guidelines***
    - ***Some images readily available to the general public (open source) by the CT vendors will be shared***

Presentation Overview

Overview

Background

Research Process

Supplementary Content

Human Factors Research

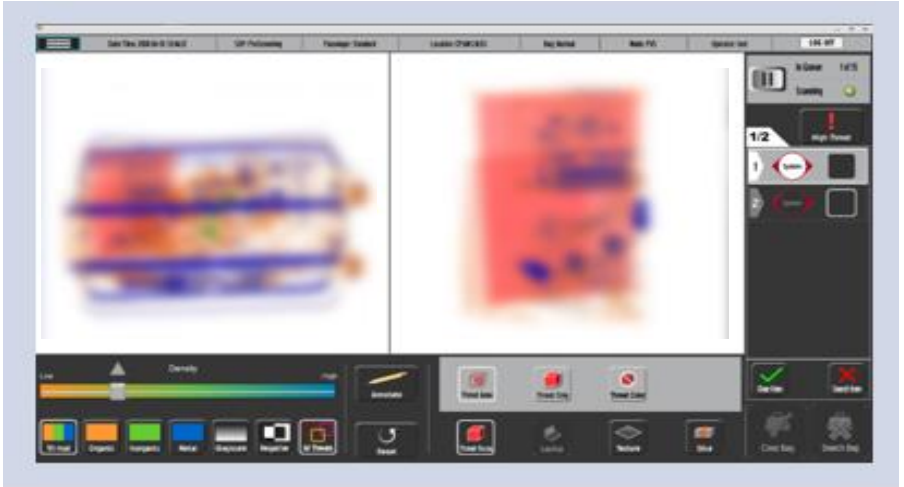
Human Factors Value at TSA

- ***Human Factors Offers***
  - The process of Human Factors research and the benefits associated from a problem- and team-oriented approach.
  - The range of tools Human Factors research offers TSA
  - How Human Factors research enhances the quality of outcomes for TSA projects across the organization.
- ***Human Factors Addresses***
  - Human performance capabilities
  - Cognitive and physical limitations
  - General patterns of human functioning
    - How people navigate, work, attend to stimuli, manage risks, etc.
- ***How Human Factors Addresses Questions***
  - Multi-method and multi-informant approach
    - Problem-oriented approach
    - Team-based approach
    - Qualitative and quantitative methods
    - Use of historical and actively collected data

Human Factors Research

Human Factors Value at TSA

- ***How Human Factors Addresses Questions***
  - Example process modeled with CT-APSS CGUI project
    - Problem:
      - Create a Common Graphical User Interface (CGUI) for CT Systems introduced at the Checkpoint that supports:
        - More effective and efficient screening
        - Ease of switching
        - Training efficiency
        - System architecture flexibility



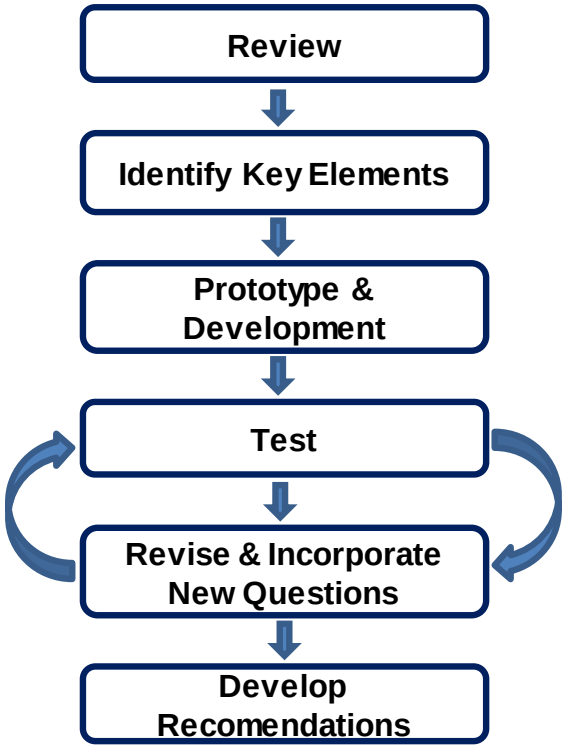
HF Research Process via CT-APSS CGUI

The CT-APSS CGUI Development models a standard HF research process for future TSA projects

CT-APSS CGUI Project Aims

- ***More effective and efficient screening***
- ***Ease of switching***
- ***Training efficiency***
- ***System architecture flexibility***

HF Research Process



Presentation Overview

Overview

Background

Research Process

Supplementary Content

Step 1: Review

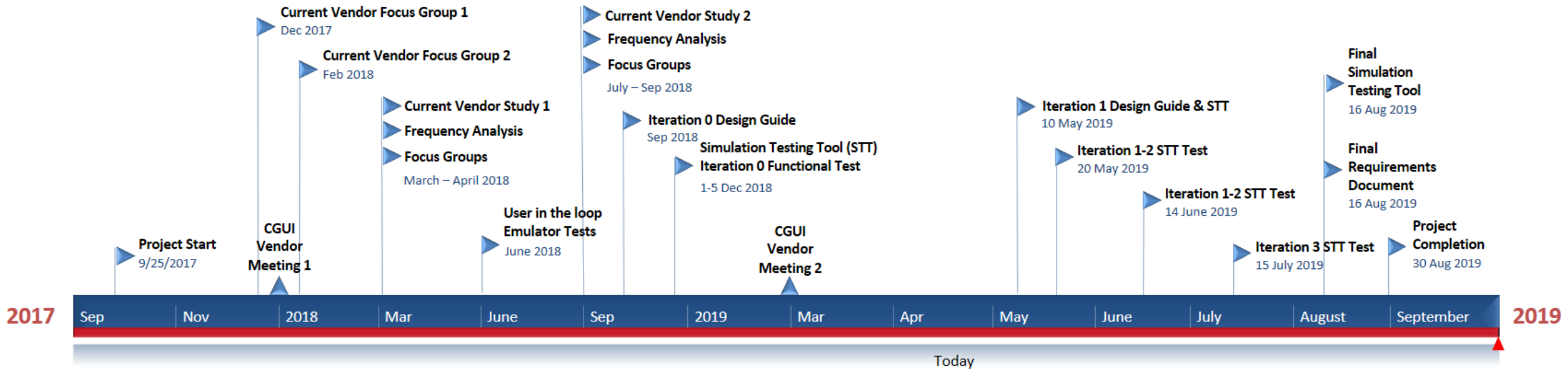
Identify the Project Timeline

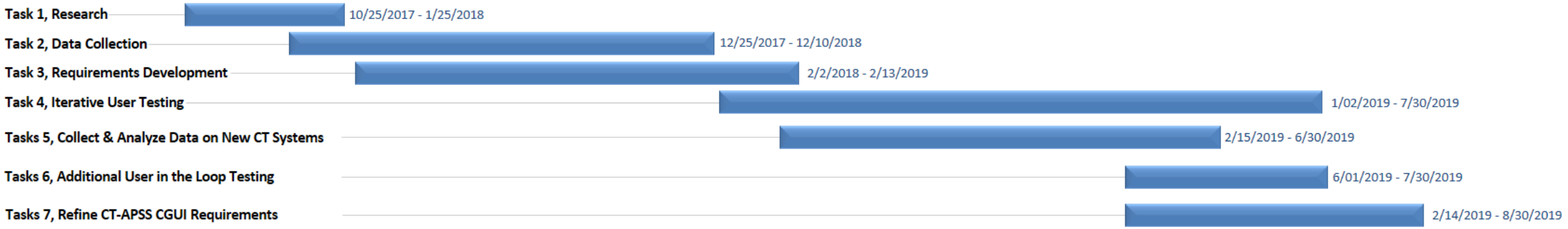
- Review the time available to complete the project
  - How much time is afforded to each step
  - Identify level of flexibility for each step and the iteration cycle

CGUI Project Timeline

- 2 Year Effort
- 7 Tasks
  - Review, Develop, Revise, and Develop Requirements and additional Avenues to Explore

Step 1: Review





Step 1: Review

Review

- The review step encompasses all activities that provide the necessary breadth and depth on the topic of interest to develop testable hypotheses and a test plan
  - Provides guidance to the relevant topic (what works, what doesn't)
  - Identifies unknowns and points of concern to address
  - Identifies possible solutions

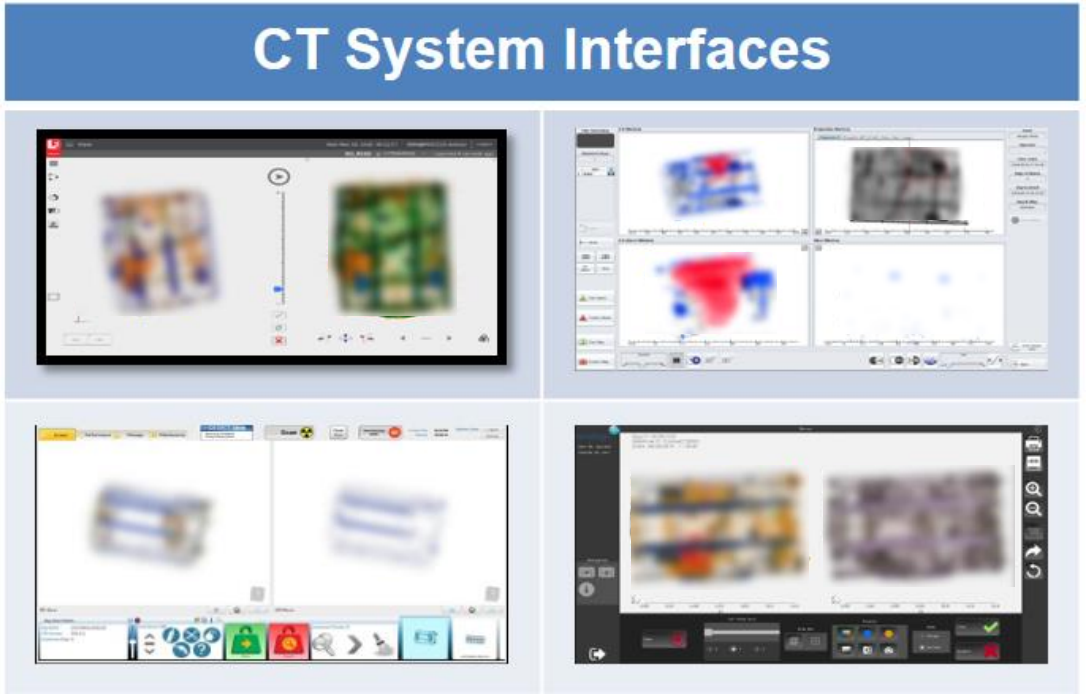
Multimethod Approach

- Begin with examination of available archival information
- Consider active review and feedback methods

Step 1: Review – CT-APSS CGUI

AT-CT System Review

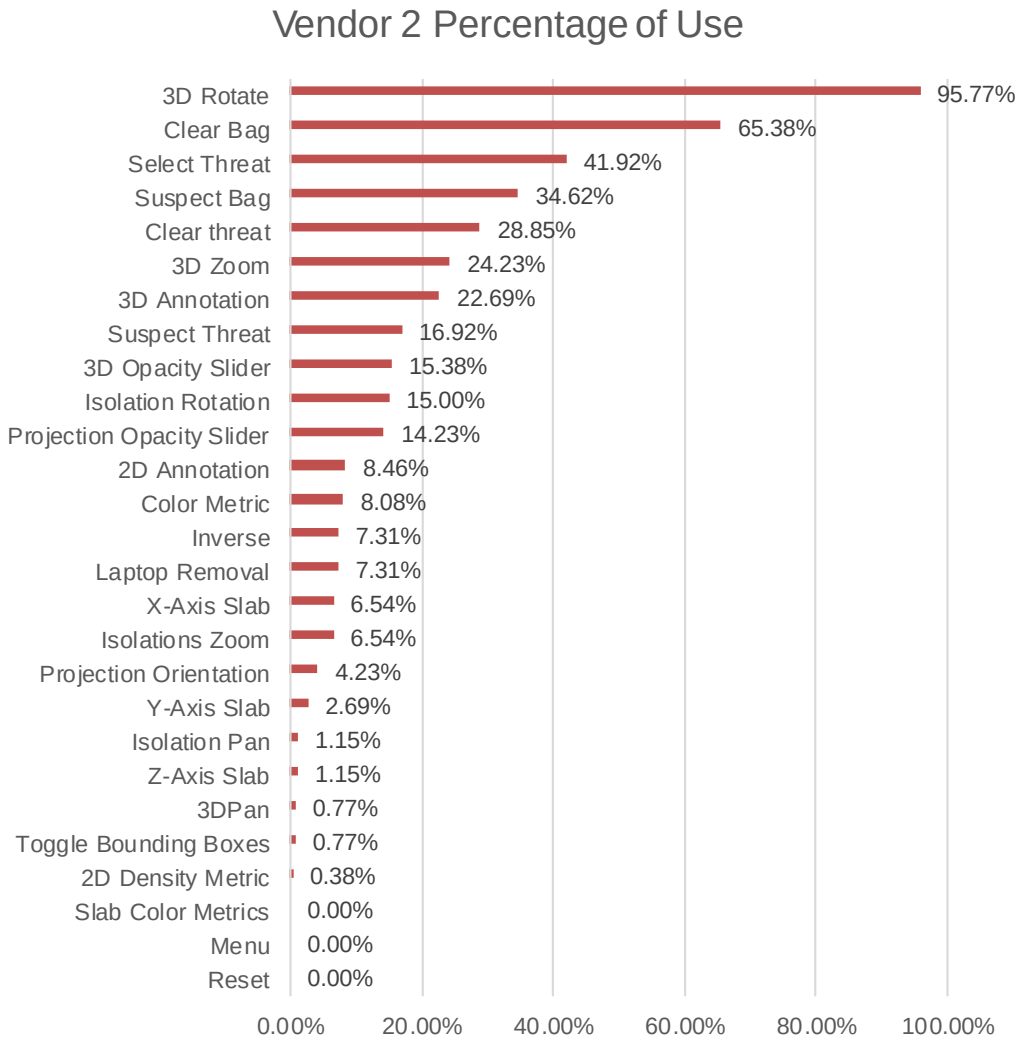
- Began with review of the literature
  - Review of EDS CGUI documentation and Prototypes
  - Review of current AT system interfaces
  - Review of the current CT system interfaces
- Conduct Usability Assessments with AT-CT system operators
  - Observe CT system operations at airports
  - Conduct Focus Groups
  - Conduct Frequency Study of AT-CT tools and functions
  - Conduct Icon Usability Surveys
- Conducted Usability Testing with Emulators
  - Focus Groups identified general perceptions of each system and pros/cons from training and implementation to daily use
  - Surveys provided specific feedback on usability (e.g., tools and design)



Step 1: Review – CT-APSS CGUI

Results

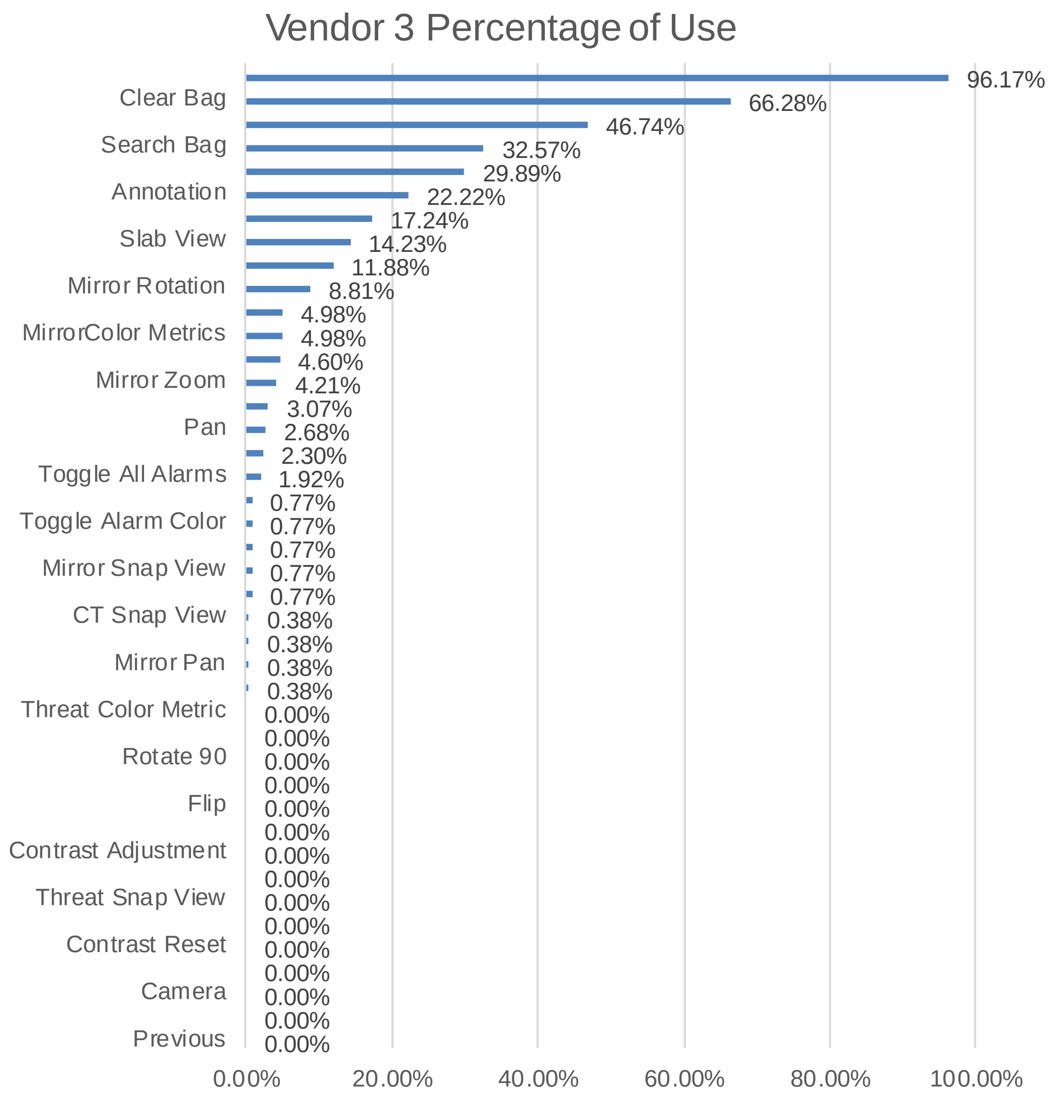
- Frequency of Use was calculated for multiple systems
  - Vendor 2 vs. Vendor 3



Step 1: Review – CT-APSS CGUI

Results

- Frequency of Use was calculated for multiple systems
  - Vendor 2 vs. Vendor 3



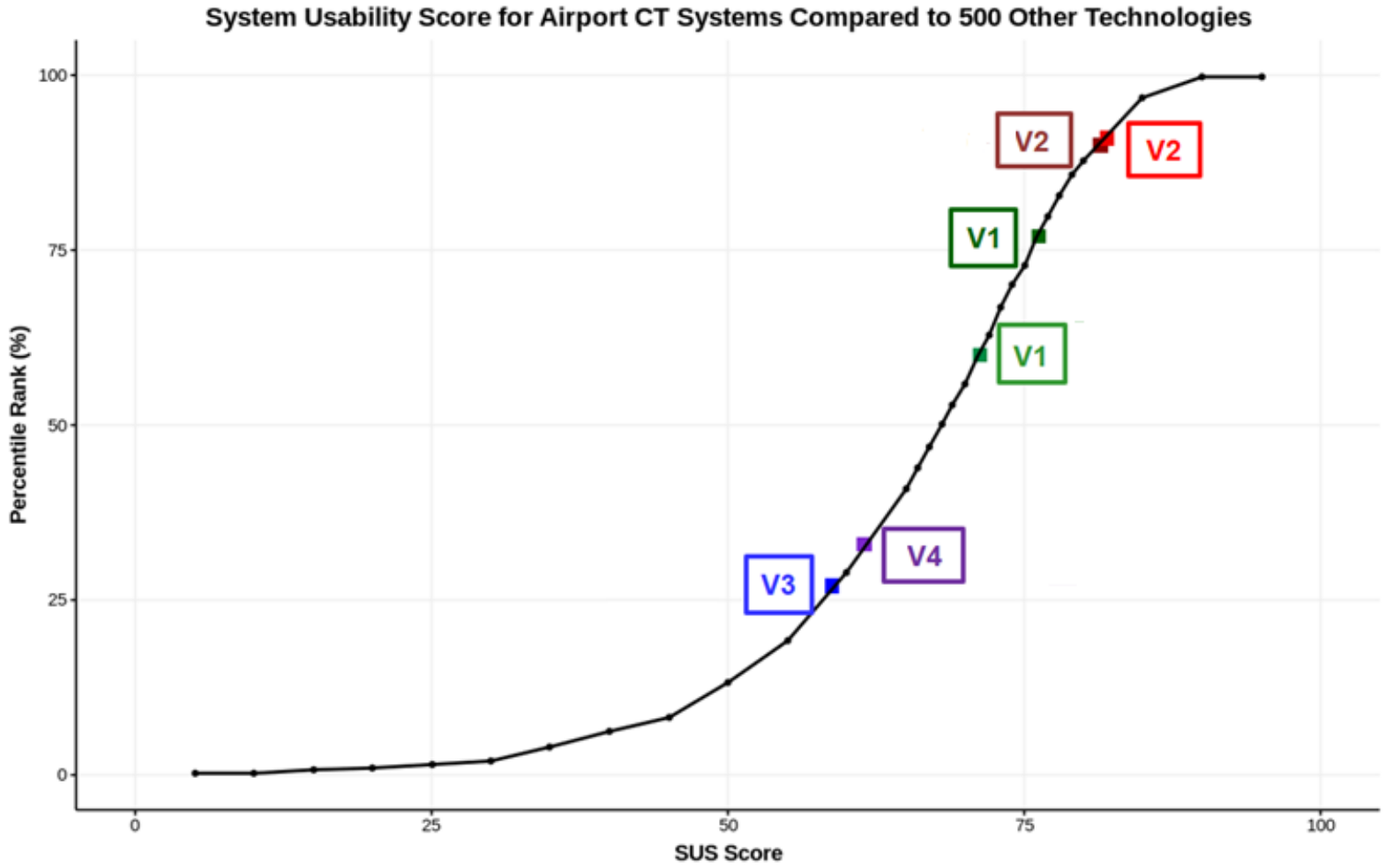
Transportation Security Administration

RCA REQUIREMENTS & CAPABILITIES ANALYSIS

Step 1: Review – CT-APSS CGUI

Results

- Usability assessment of different vendor CT systems. Rating from best usability to worst is:
  - Vendor 1 (V1)
  - Vendor 2 (V2)
  - Vendor 3 (V3)
  - Vendor 4 (V4)



Step 1: Review – CT-APSS CGUI

Results

- Comprehension Scores
  - Total number of icons were calculated for each system and the comprehension for each icon was calculated
    - The number of icons, icons that had low comprehension, and proportion of low comprehension icons were calculated by system
  - Comprehension rates for each system are ordered from best to worst.
    - Any system-specific icons for non-screening functions were removed from contention before tabulation (*).
- Unexpected findings
  - Vendor 4's system was rated as one of the least preferred and usable interfaces but had the best user comprehension of system icons overall.

Vendor	Total Icons	Low Comprehension	Proportion Low Comp*
Vendor 4	42	17	0.405
Vendor 1	24	15	0.625
Vendor 3	32	24	0.750
Vendor 2	21	16	0.762

Step 3: Prototype & Development

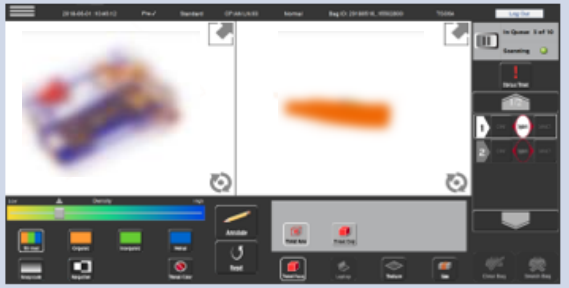
Prototype and Development

- Key Elements lead to Specific, Testable questions for the current project
  - Supports development of a test plan and, when relevant, a testable prototype (for product-based testing)
  - Often can identify additional testable questions for future consideration early
- Team-based dynamic supports faster development
  - All team members develop prototypes/testable questions from content available and team narrows down focus together

CGUI Prototype

- Team used Key Elements identified from testing to develop a prototype interface that all individuals commented on and revised over several cycles until group consensus was reached
- Specific questions about the prototype were developed by the team and verified by HPB to incorporate into the test plan

CGUI (Initial Iteration)



Step 4: Test

Testing Considerations

- Specific testable questions are identified from the Key Elements and Prototype Development
  - Methods to test questions are chosen
    - Surveys
      - Empirically-validated
      - Measure development & validation
    - Performance data collection
    - Additional focus group/qualitative data collection
    - Observational data collection
  - Methods dictate analytic procedures planned on completion of testing
    - Qualitative analysis
      - Frequency data
      - Sentiment analysis
    - Quantitative analysis
      - Experimental structure
        - Within-subjects vs. Between-subjects
      - Parametric and non-parametric procedures given design
        - Does the data meet necessary quality and assumptions for chosen plan

Step 4: Test

Testing Considerations

- A basic test plan structure for experimental testing is completed given the decisions made about test methods and analyses chosen
  - Template structure includes the following considerations:
    - Participants
    - Equipment
    - Time and Logistics
    - Procedure and Methods
    - Analytic Plan

CGUI Test

- Chose a basic experimental design that capitalized on primarily qualitative data due to initial prototype constraints for performance data collection
  - Quantitative Data
    - Performance data indicated useful tools and unnecessary tools for CT
    - Performance data indicated tools requiring improvement based on use timing data
  - Qualitative: usability assessment, survey feedback, verbal feedback for sentiment (positive and negative) coding

Step 4: Test

CGUI Test

- Test Structure
  - Pre-test questionnaire session
    - Demographic survey
  - Testing Session
    - Training on Prototype system
    - Four 150-minute functional testing sessions
      - Four participants (TSOs) per session over 4 sessions for 16 total participants
  - Post-test questionnaire session
    - System Usability Scale
      - About current CT system in use
      - About Prototype system
    - Alternatives Survey
      - Developed and structured survey measure according to basic psychometric principles
      - Assessed preferences for alternatives to interface, icons, and tools in design and use
    - Open-ended feedback option

Step 4: Test

Analyze and Process Results

- Follow the Analytic Plan that matches the Test Structure
  - Examine the data and ensure assumptions are met
    - Consider alternative analyses if original plan does not work (e.g., nonparametric vs parametric procedures
  - Interpret the results and consolidate for feedback

CGUI Results Overview

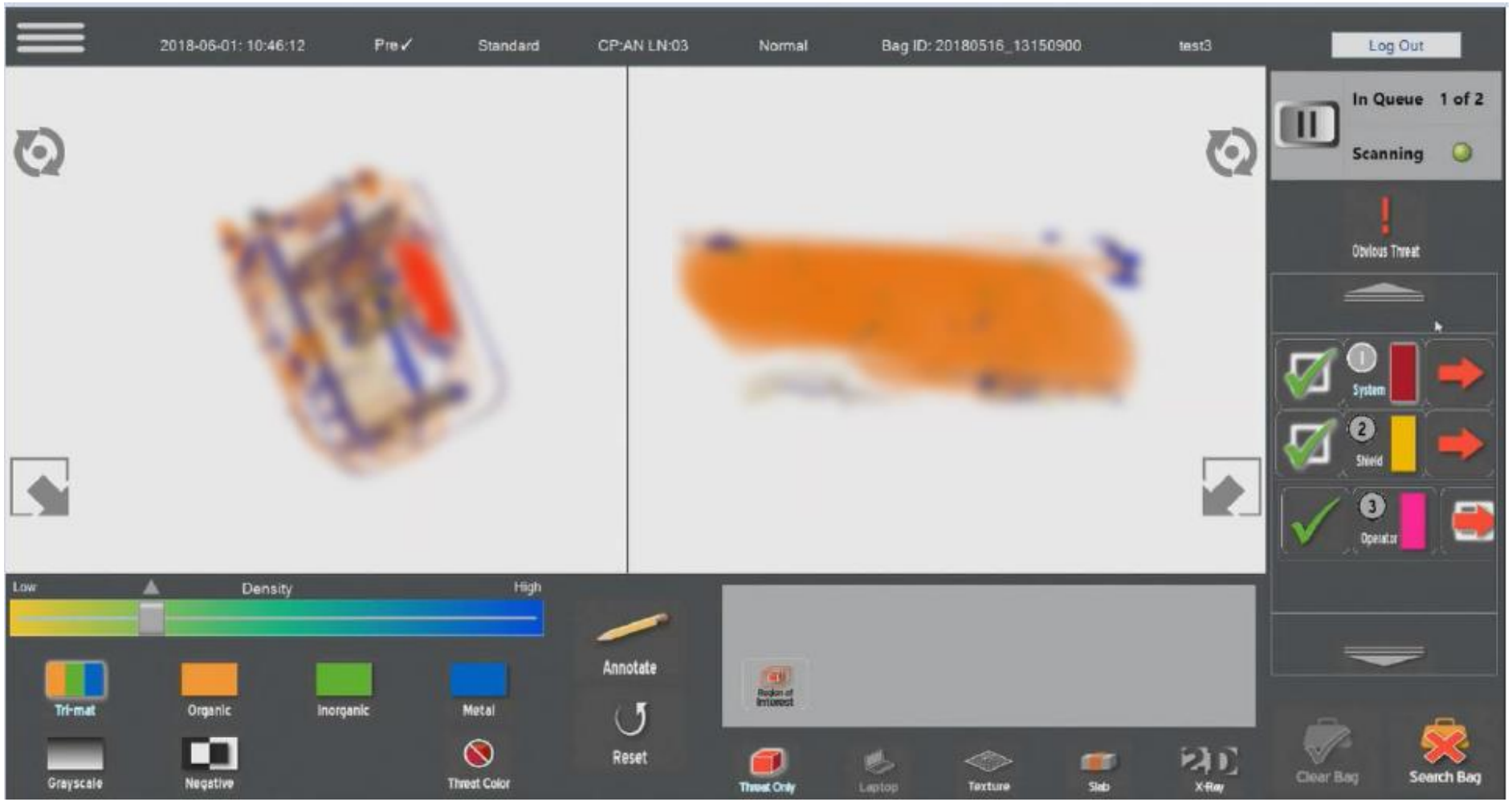
- Quantitative Data
  - Logistic constraints limited interpretation of data
  - Data indicated limited interpretable feedback available from initial prototype test
- Qualitative Data
  - Provided general system feedback
    - System Usability Scale
  - Identified additional improvements to the design as well as the rationale for performance data trends
    - Alternatives survey
      - Frequency analysis
    - Open-ended feedback
      - Sentiment analysis

Step 4: Test

CGUI Results Overview

- Quantitative Data
  - Logistic constraints limited interpretation of data due to USG furlough
  - Data indicated limited interpretable feedback available from initial prototype test
    - Test airport did not have functioning Checkpoint CT for some time, leaving most users unfamiliar with the system and error-prone for reasons that could not be attributed to factors other than the design (e.g., lack of experience)
    - Qualitative data provided detailed rationale for user perspective on interaction with machine and error sources – served as more reliable outcomes for first iteration test

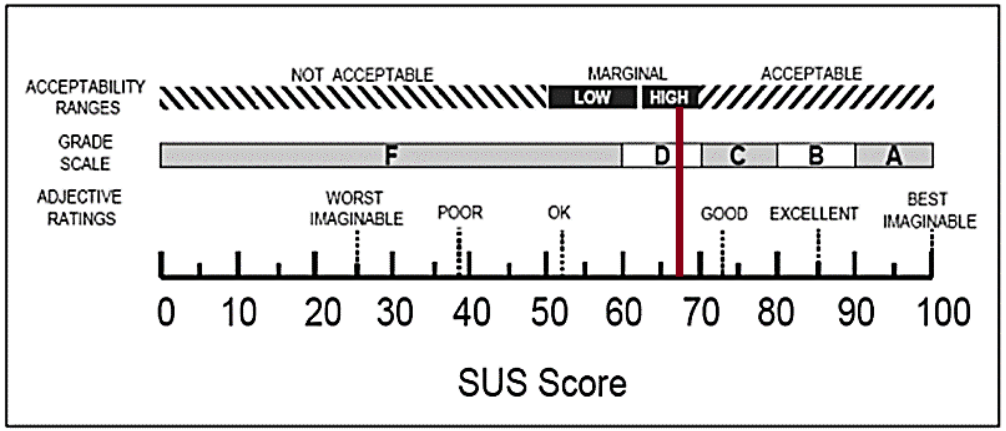
Step 4: Test

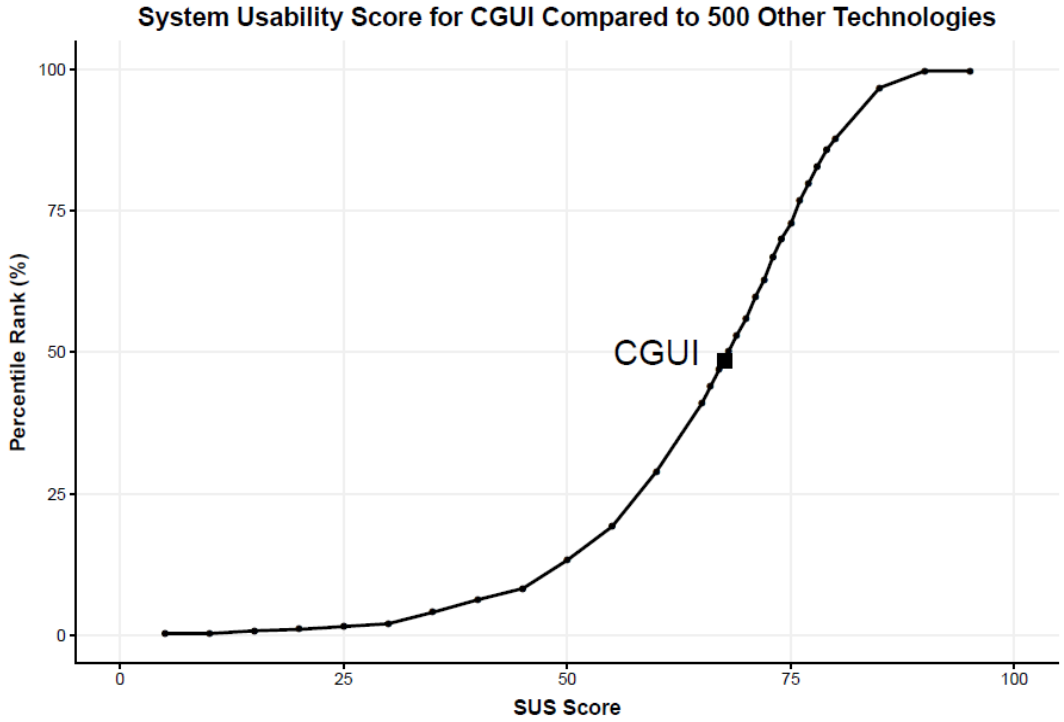

2018-06-01: 10:46:12
Pre✓
Standard
CP:AN LN:03
Normal
Bag ID: 20180516_13150900
test3
Log Out
In Queue 1 of 2
Scanning
Obvious Threat
System
Shield
Operator
Low
Density
High
Tri-mat
Organic
Inorganic
Metal
Grayscale
Negative
Threat Color
Annotate
Reset
Threat Only
Laptop
Texture
Slab
X-Ray
Clear Bag
Search Bag

Step 4: Test

CGUI Results Overview

- Qualitative Data
  - Provided general system feedback
    - System Usability Scale

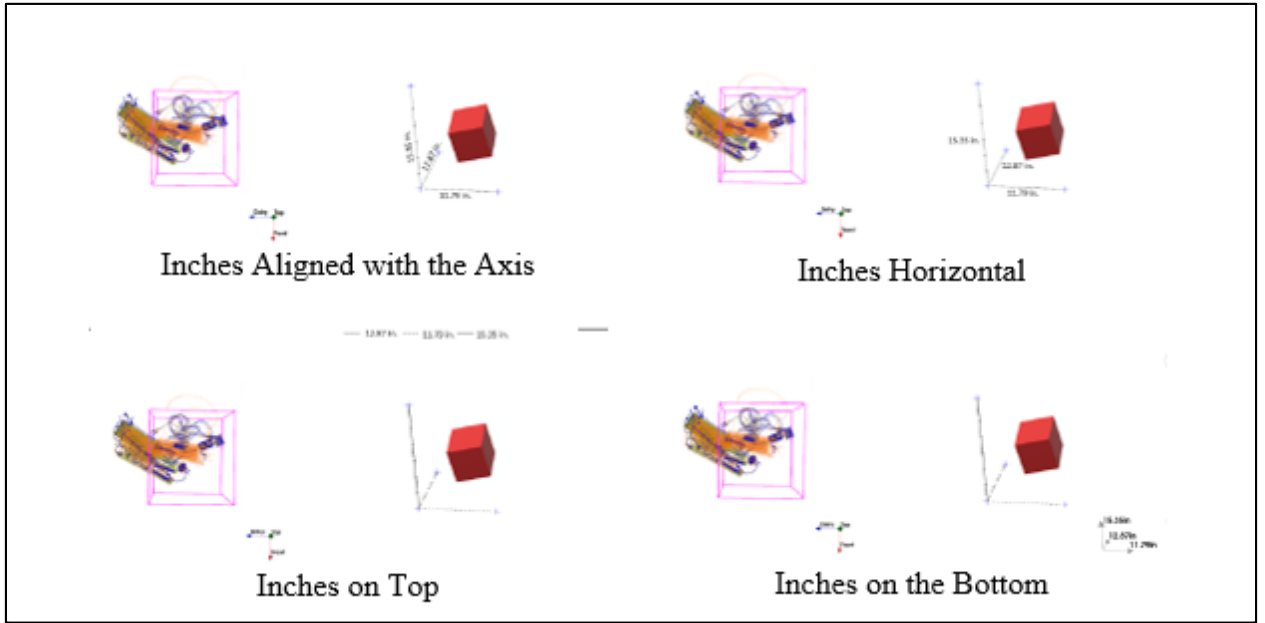




Step 4: Test

CGUI Results Overview

- Qualitative Data
  - Identified additional improvements to the design as well as the rationale for performance data trends
    - Alternatives survey
      - Frequency analysis

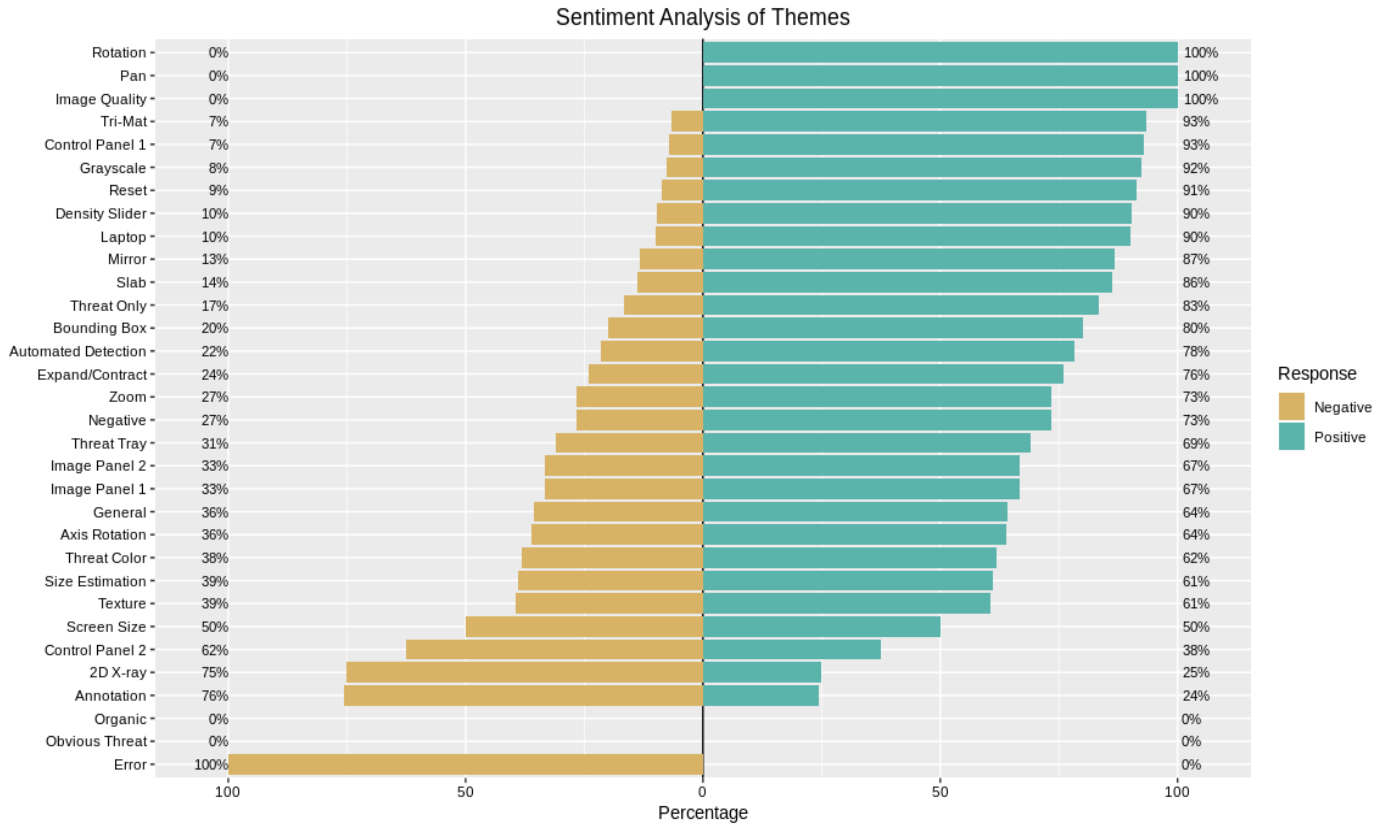


Annotation Label Questions				
	Rank			
	1	2	3	4
Inches Aligned with Axis	33.33%	46.67%	0.00%	20.00%
Inches Horizontal	21.43%	28.57%	42.86%	7.14%
Inches on Top	33.33%	0.00%	33.33%	33.34%
Inches on Bottom	21.43%	21.43%	21.43%	35.71%

Step 4: Test

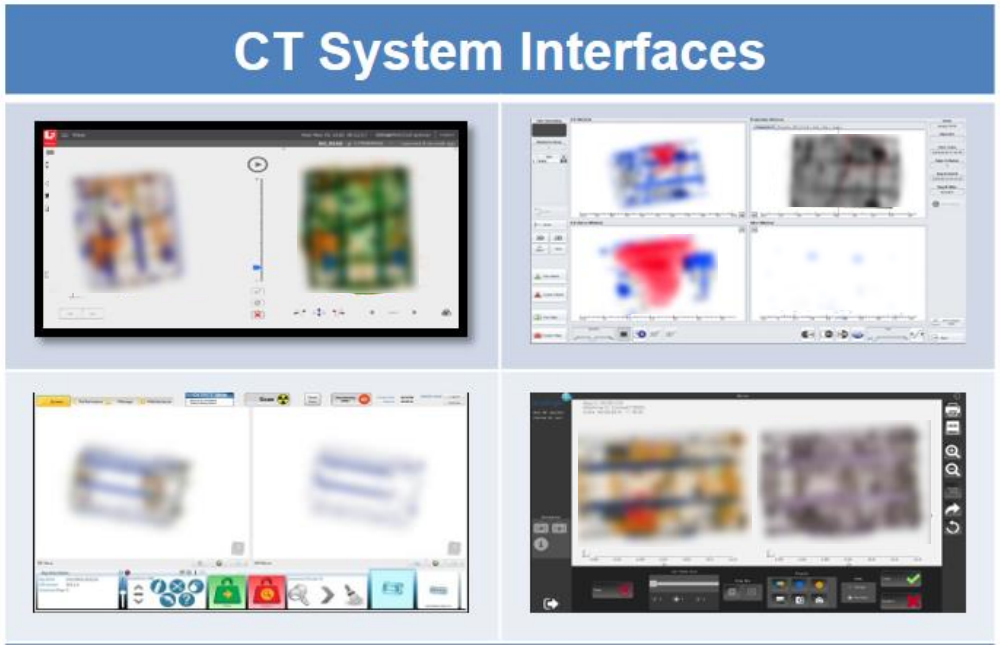
CGUI Results Overview

- Qualitative Data
  - Identified additional improvements to the design as well as the rationale for performance data trends
    - Open-ended feedback
      - Sentiment analysis

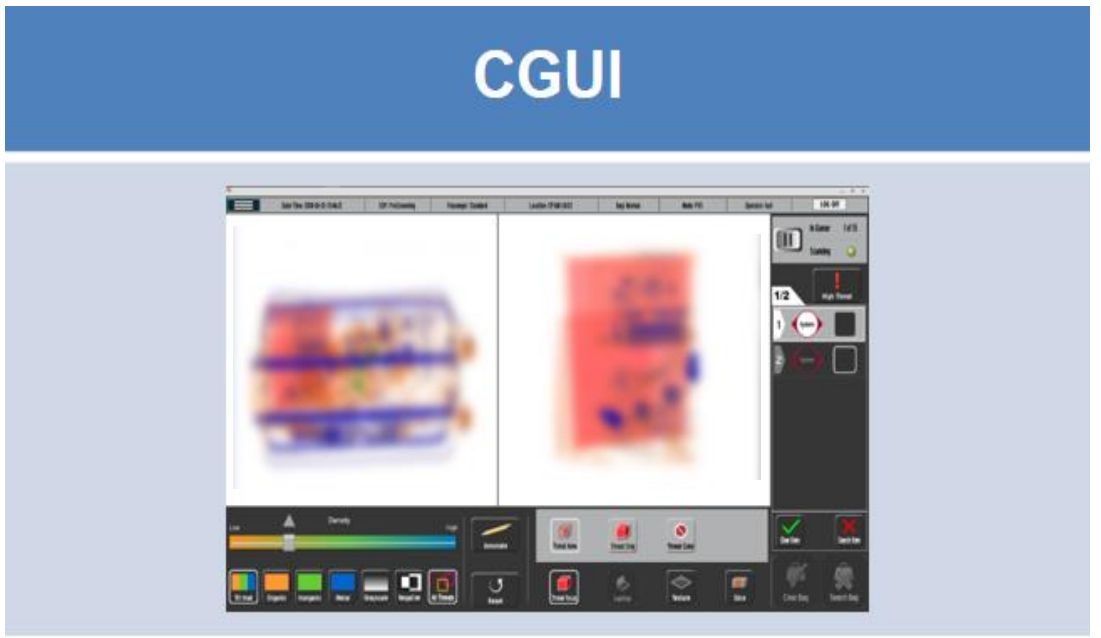


Step 5: Revise and Incorporate New Questions

Revise
&
Incorporate New Questions



Next Testing Cycle
Informed by
Revisions and Questions



Step 4-5: CGUI Iterations

CGUI Test Process for Additional Iterations	

Parallel Experimental Design	• Quantitative Data:
	  • Performance data
	• Qualitative Data:
	  • Usability assessment, survey feedback, verbal feedback for sentiment (+/–) coding

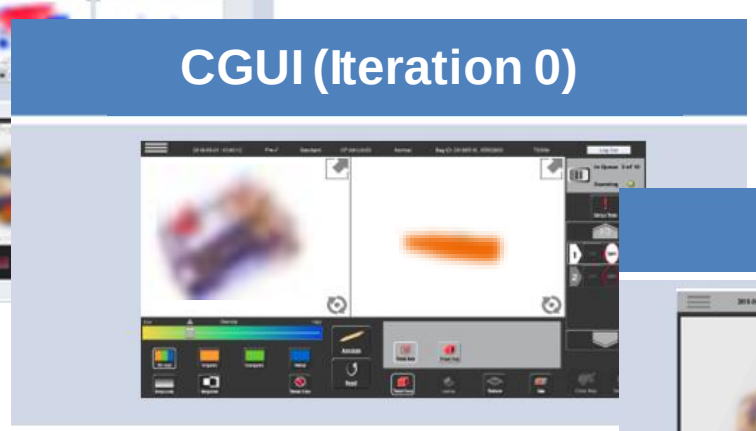
Parallel Test Structure	• Pre-test questionnaire session
	  • Demographic survey
	• Testing Session
	  • Training on Prototype system
	  • Four 150-minute functional testing sessions
	• Post-test questionnaire session
	  • System Usability Scale for current CT system in use & Prototype
	  • Alternatives Survey
	  • Developed and structured updates survey measures according to basic psychometric principles
	  • Assessed preferences for interface, icons, and tools in design and use
	  • Open-ended feedback option

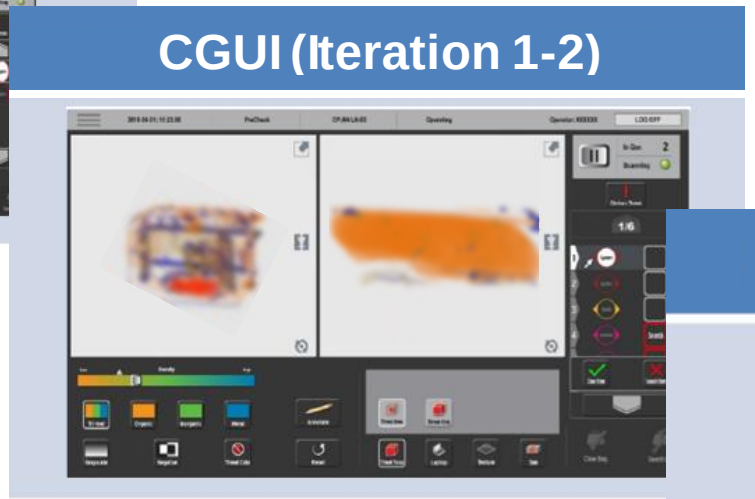
Step 4-5: CGUI Iterations

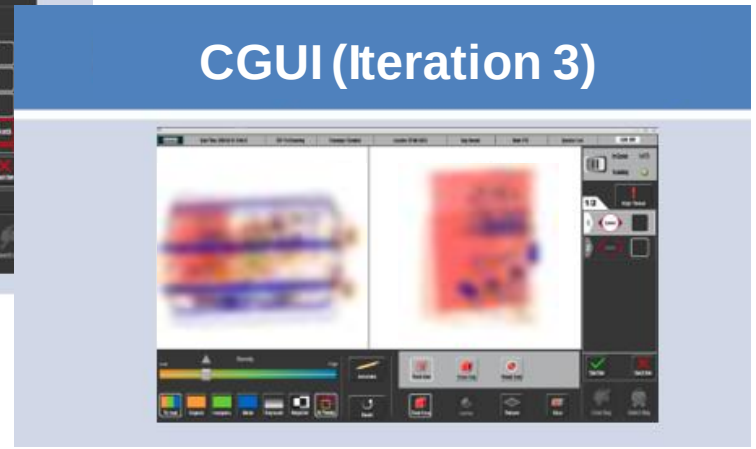
We collected quantitative and qualitative data from 6 airports
Data from over 90 TSOs helped to shape 3 iterations of the CGUI

Airport	Date	Iteration	Airport CT Vendor
1	December 4-5	Iteration 0	Vendor 1
2	May 20-23	Iteration 1-2	Vendor 2
3	June 3-5	Iteration 1-2	Vendor 3
4	June 3-5	Iteration 1-2	Vendors 1 & 4
5	June 10-12	Iteration 1-2	Vendor 1
6	July 16-17	Iteration 3	Vendors 1 & 4

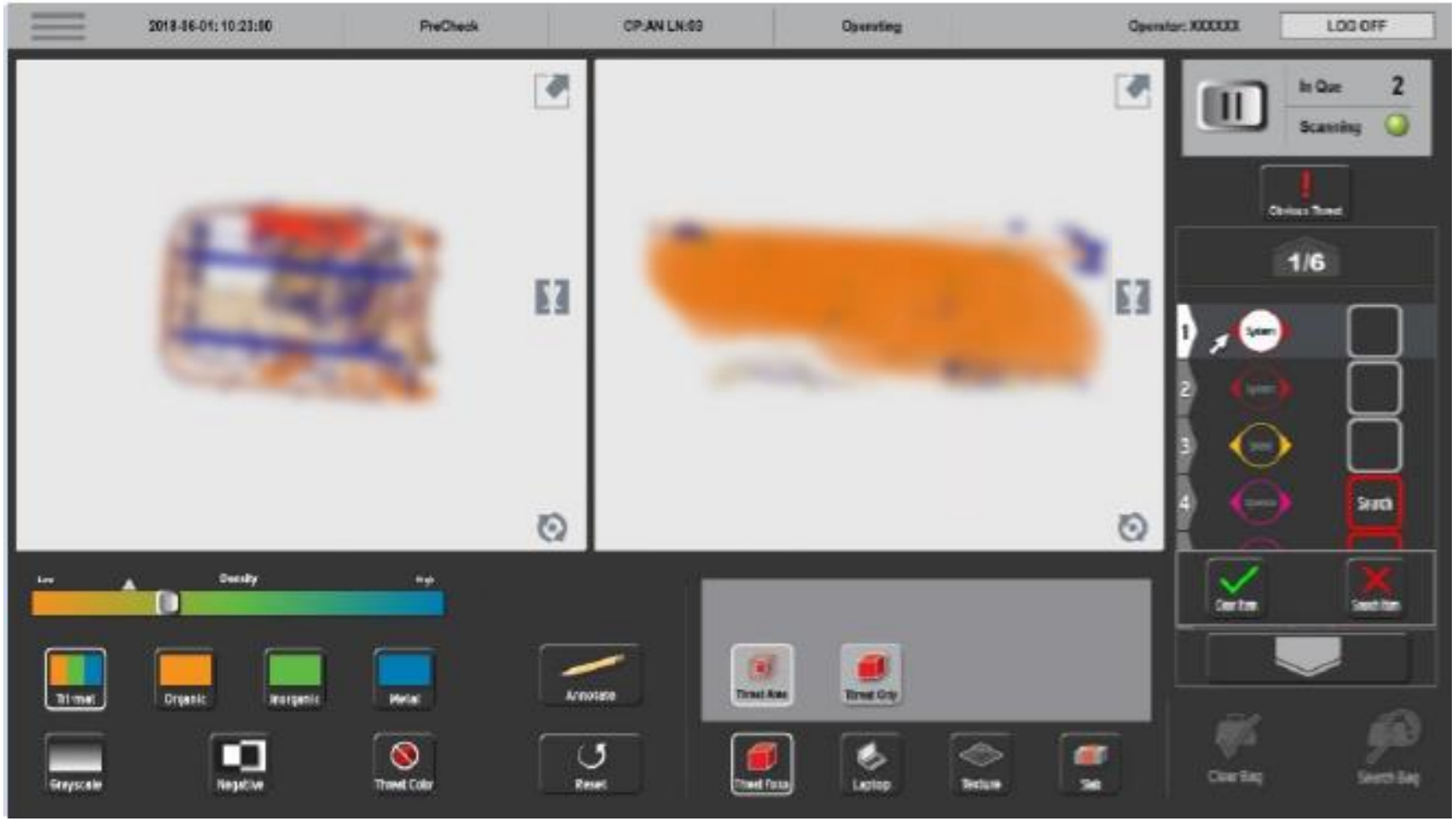
Steps 4-5: Repeat and Refine



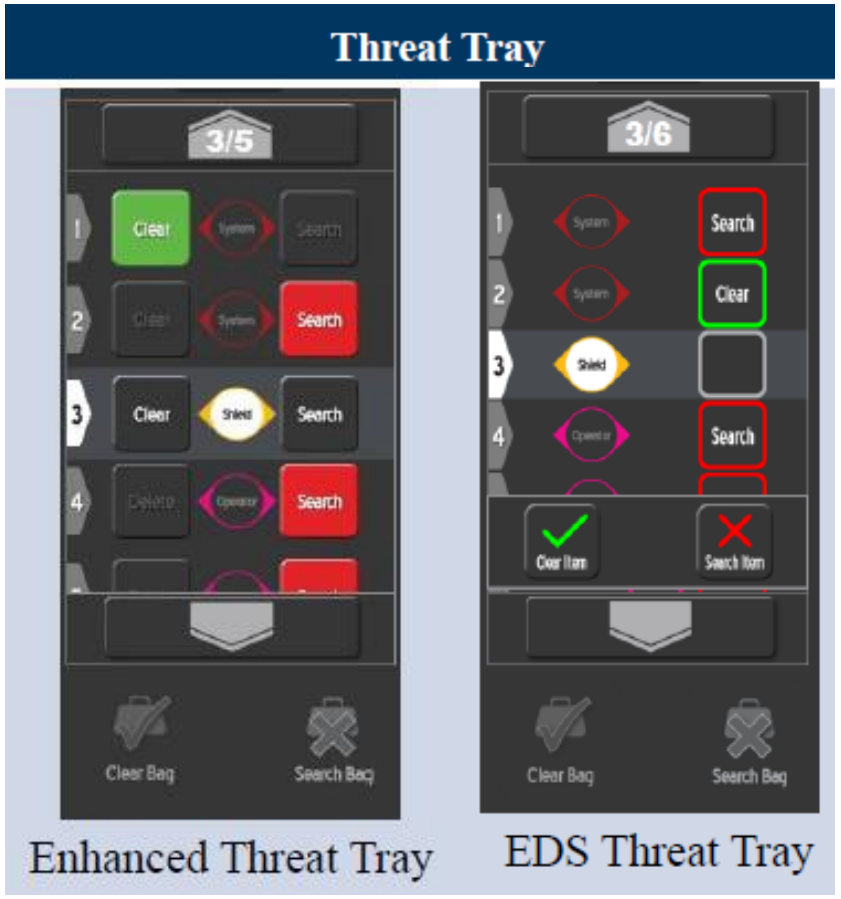


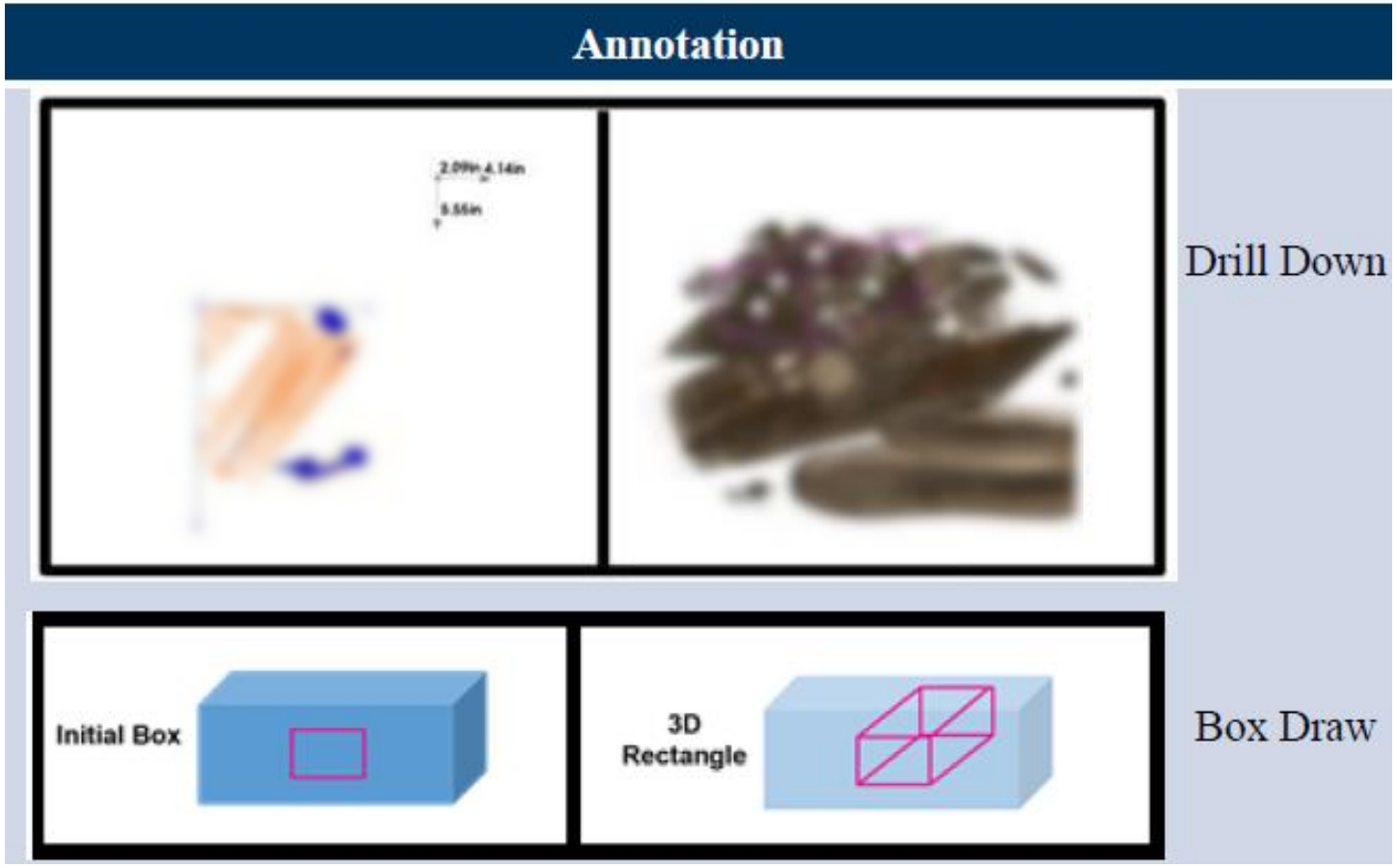


Step 4-5: CGUI Iteration1-2



Step 4-5: CGUI Iteration1-2

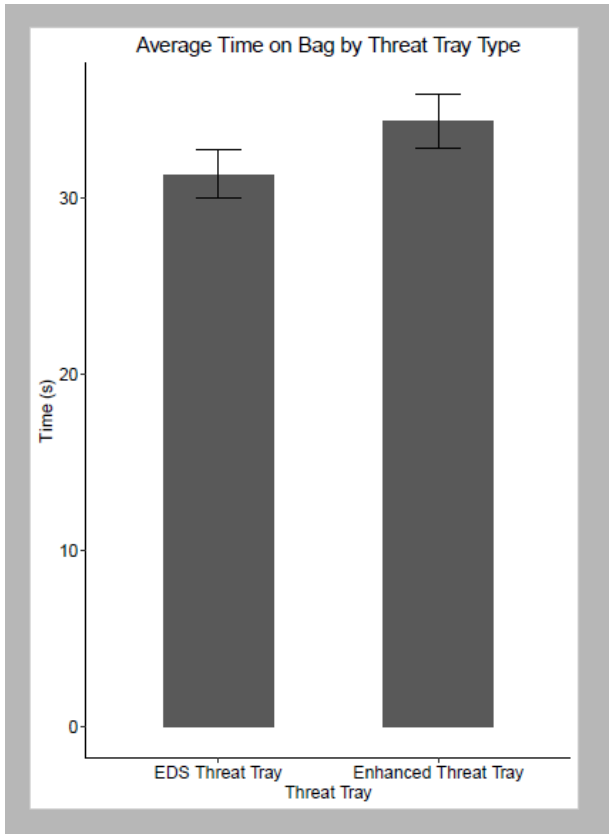


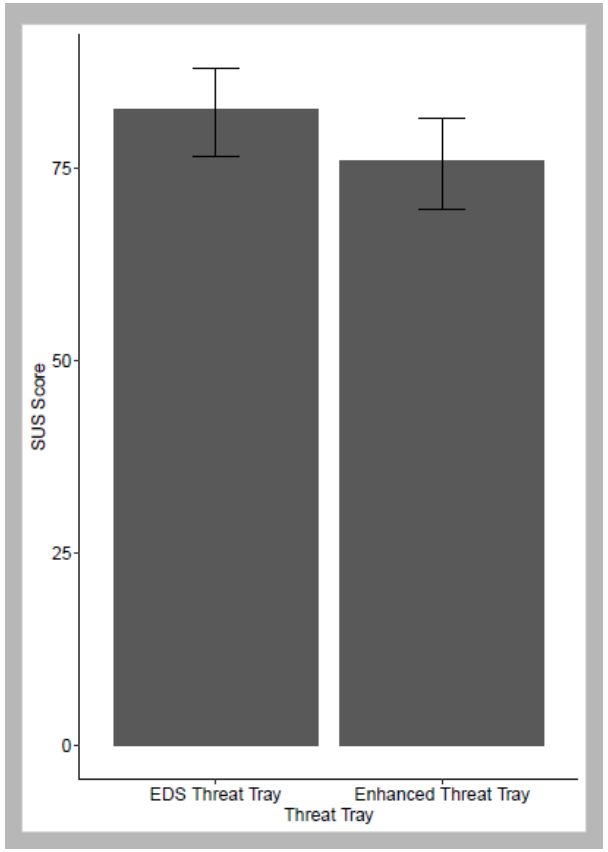


Step 4-5: CGUI Iteration 1-2

Iteration 1-2 Results Overview

- Quantitative Data
  - No significant differences in threat discrimination for threat tray conditions
  - Determined that the EDS threat tray offered the greatest opportunity to increase image panel real estate
    - The performance statistics (d' & A') indicate the following Threat Tray and Annotation combinations:
      - 1st: EDS and Drill Down
      - 2nd: Enhanced and Drill Down
      - 3rd: Enhanced and Box Draw
      - 4th: EDS and Box Draw

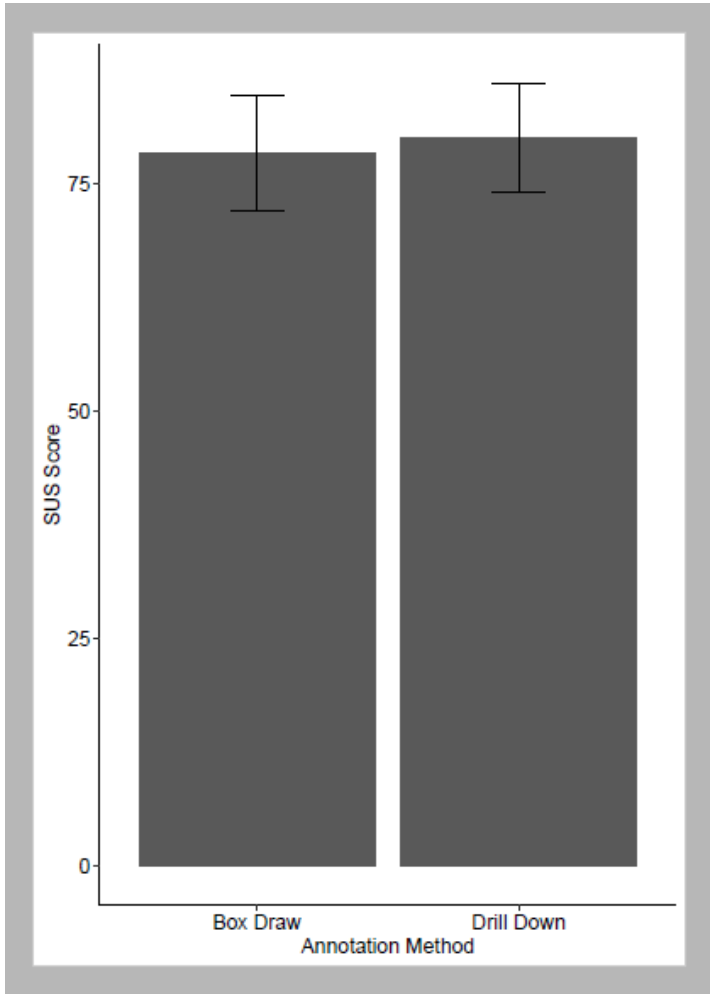




Step 4-5: CGUI Iteration 1-2

Iteration 1-2 Results Overview

- Quantitative Data
  - No significant differences in threat discrimination for annotation conditions
  - Drill Down annotation method had greatest opportunity to become the fastest method given advancement in object estimation algorithms
    - The performance statistics (d' & A') indicate the following Threat Tray and Annotation combinations:
      - 1st: EDS and Drill Down
      - 2nd: Enhanced and Drill Down
      - 3rd: Enhanced and Box Draw
      - 4th: EDS and Box Draw

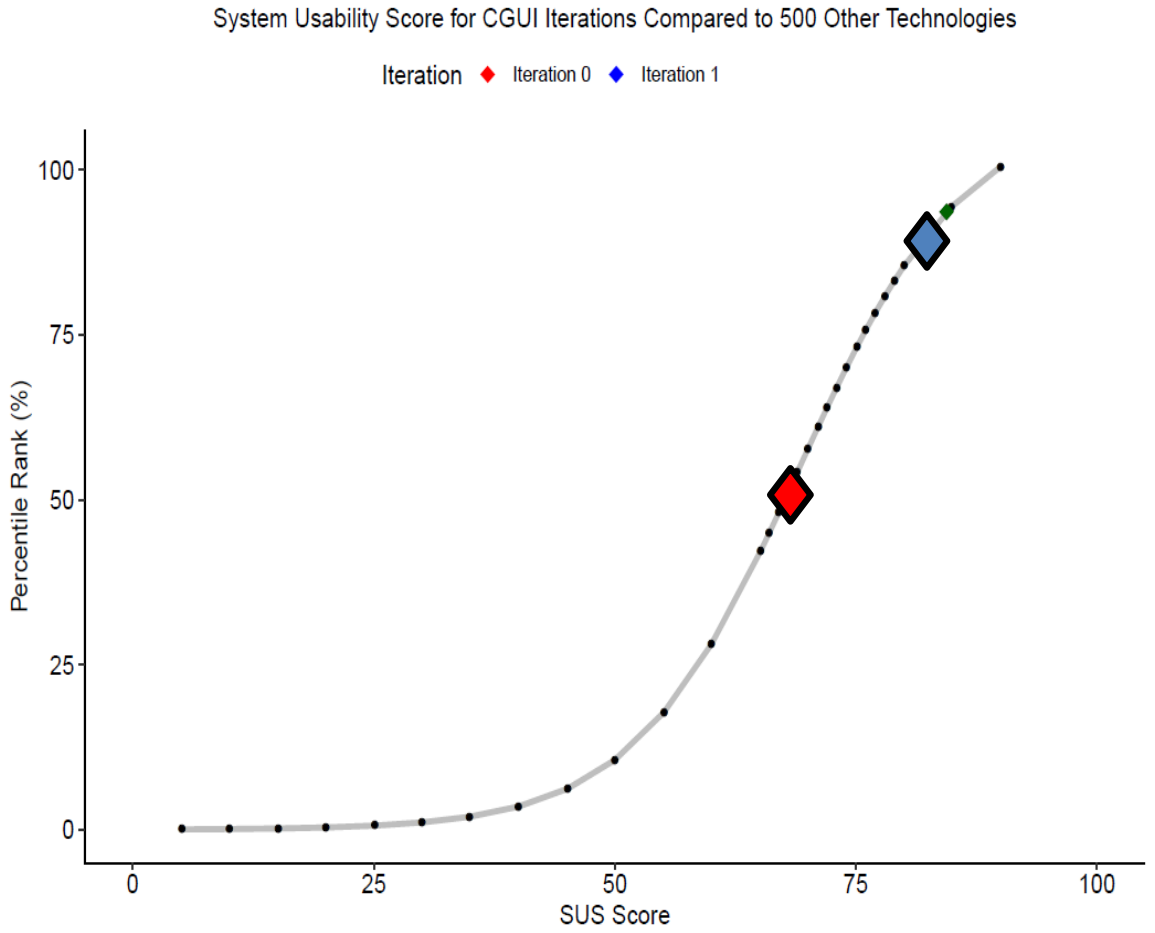


Step 4-5: CGUI Iteration 1-2

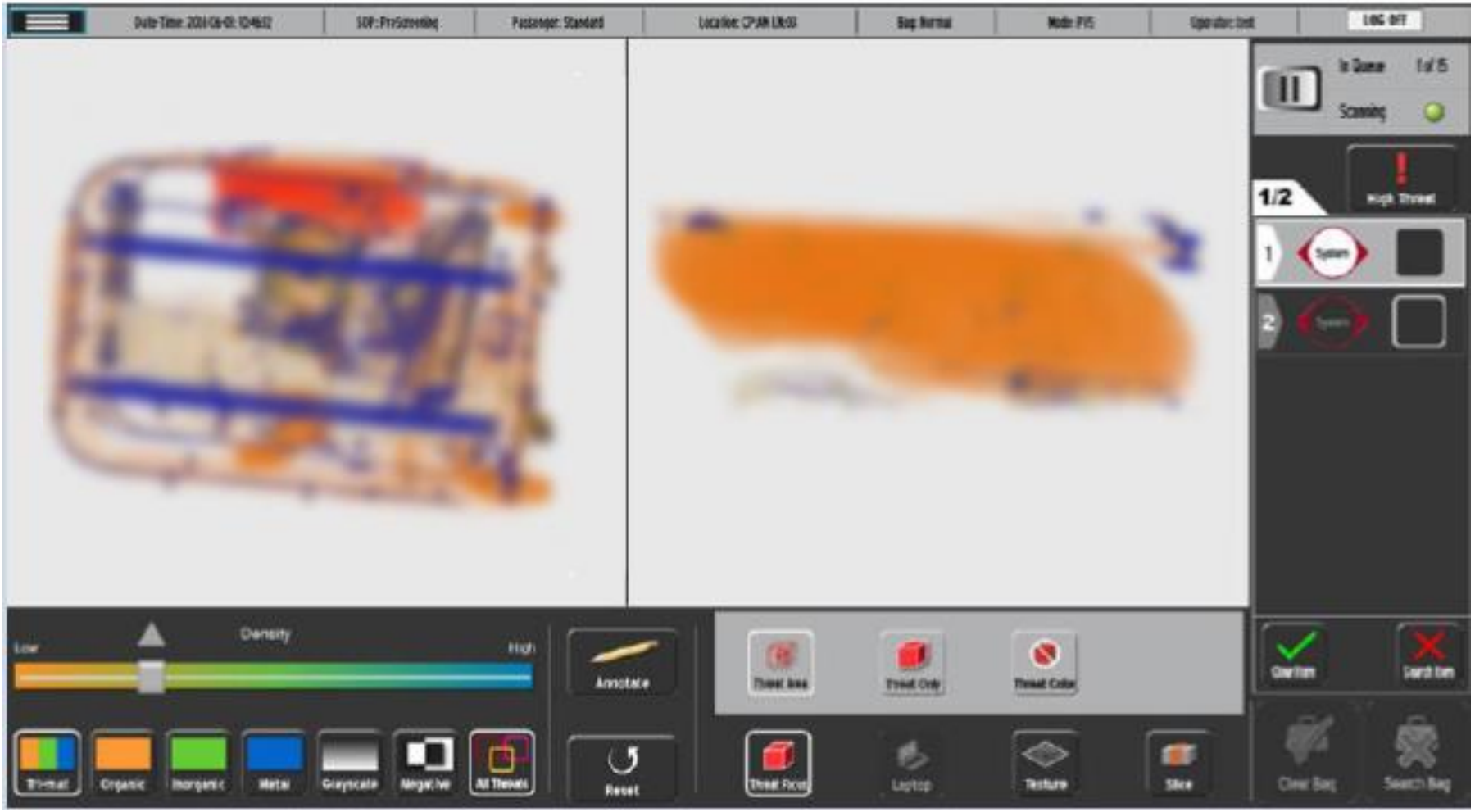
Iteration 1-2 Results Overview

- Qualitative Data
  - Each additional iteration increased the perceived usability of the system.
  - TSOs felt that the CGUI was a highly usable, easy to learn system by the last iteration.
  - Several specific desires were noted from other qualitative results

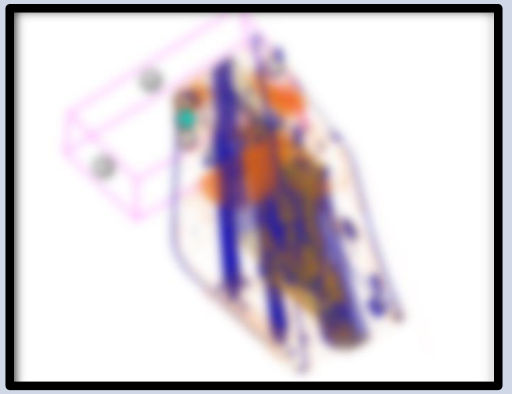
Iteration	SUS Score	Percentile	Grade	Adjective
Iteration 0	67.66	50.16	D	OK
Iteration 1-2	82.35	90.16	B	Good



Step 4-5: CGUI Iteration 3



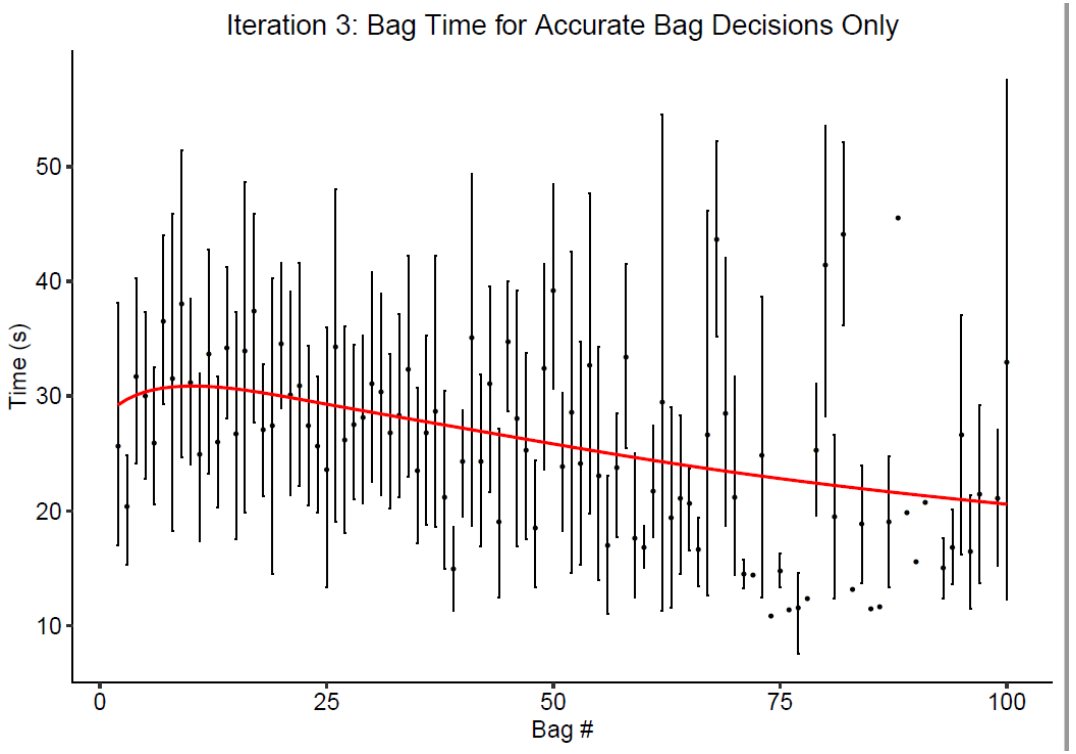
Step 4-5: CGUI Iteration 3

Modality	Threat Tray	Annotation
 Mouse  Touchscreen	 EDS Threat Tray	 Hybrid

Step 4-5: CGUI Iteration 3

Iteration 3 Results Overview

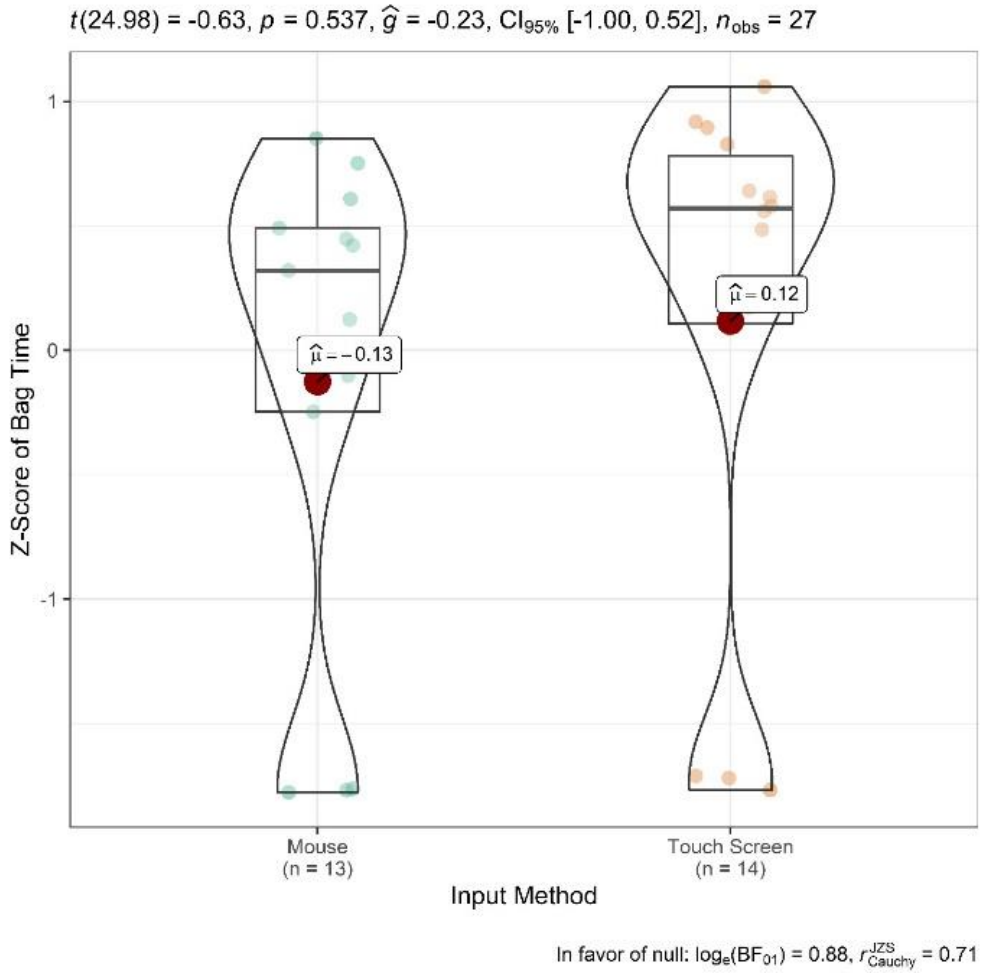
Category	Measure	Results
System Learning	Timing	↑ accuracy w/in 1 hour of training*
	Accuracy	↑ speed w/in 1 hour of training*
System Modality	Timing	Touchscreen > Mouse*
	Accuracy	Touchscreen ≈ Mouse
Iteration	Timing	Iteration 1-2 < Iteration 3
	Accuracy	Iteration 1-2 ≈ Iteration 3



Step 4-5: CGUI Iteration 3

Iteration 3 Results Overview

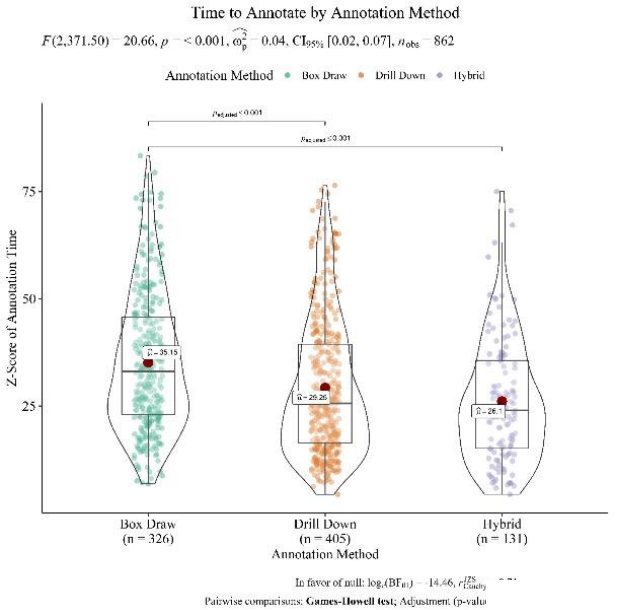
Category	Measure	Results
System Learning	Timing	↑ accuracy w/in 1 hour of training*
	Accuracy	↑ speed w/in 1 hour of training*
System Modality	Timing	Touchscreen > Mouse*
	Accuracy	Touchscreen ≈ Mouse
Iteration	Timing	Iteration 1-2 < Iteration 3
	Accuracy	Iteration 1-2 ≈ Iteration 3

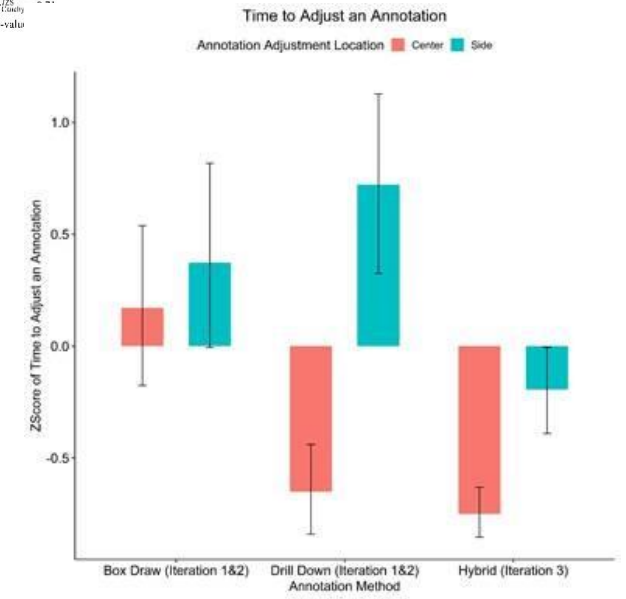


Step 4-5: CGUI Iteration 3

Iteration 3 Results Overview

Category	Measure	Results
System Learning	Timing	↑ accuracy w/in 1 hour of training*
	Accuracy	↑ speed w/in 1 hour of training*
System Modality	Timing	Touchscreen > Mouse*
	Accuracy	Touchscreen ≈ Mouse
Iteration	Timing	Iteration 1-2 < Iteration 3
	Accuracy	Iteration 1-2 ≈ Iteration 3

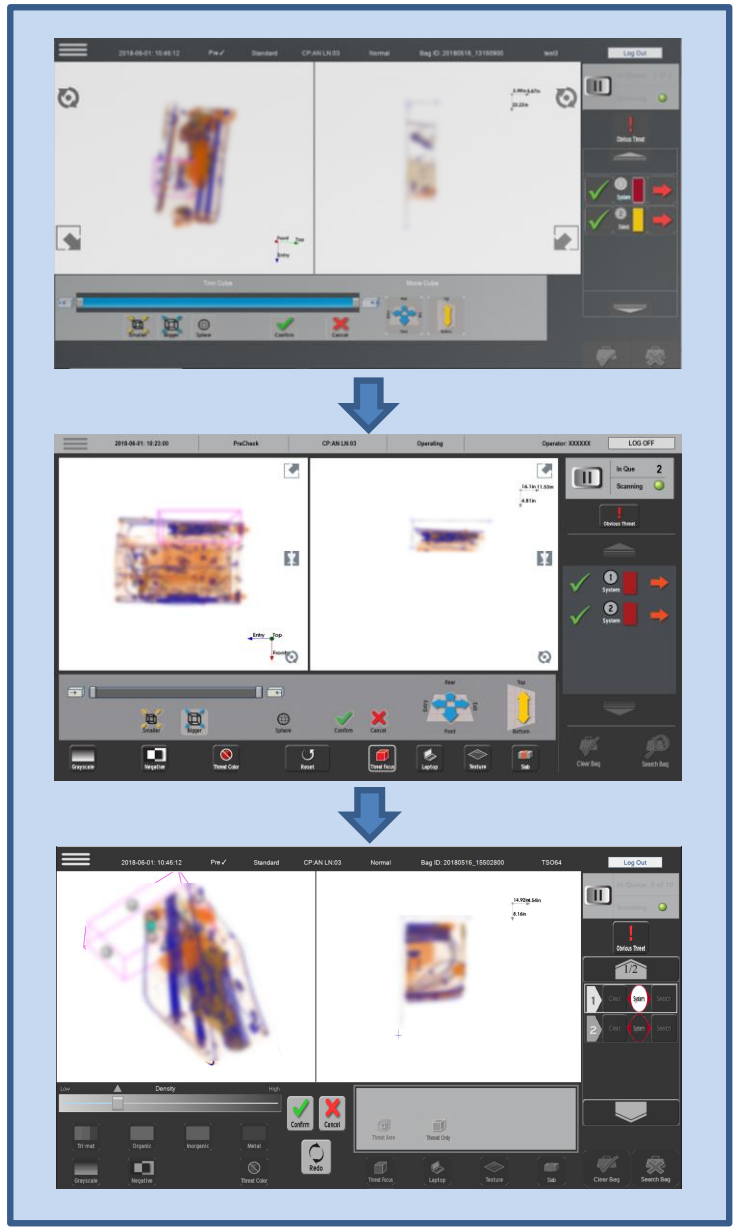




Step 4-5: CGUI Iteration 3

Unexpected Lessons - Developing 3D Annotation

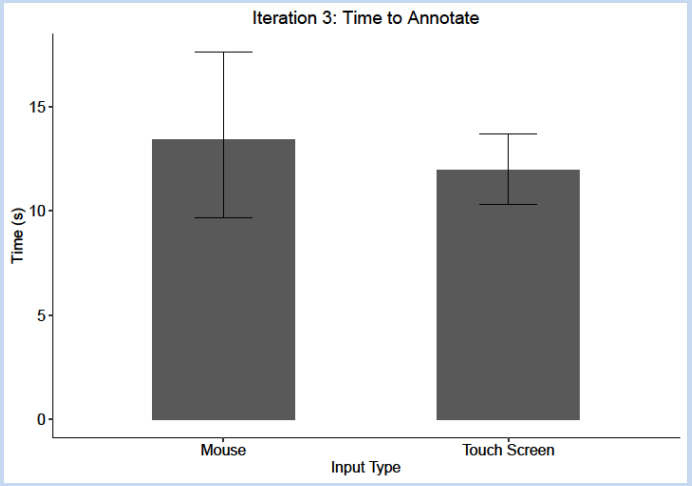
- Developing Annotation from a 2D to 3D process and creating an efficient and easily learnable process proved challenging
  - Significant changes to each iteration of Annotation
  - Qualitative and Quantitative data were vital to inform the necessary changes, interpret performance issues, and identify methods of improvement
  - Process:
    - Started with click-and-drag placement of a box that relied on tool-based adjustments
    - Shifted to testing options:
      - Click-and-drag placement (Box Draw) vs. Automatic placement on screen click (Drill Down)
      - Tool-based vs. Touchpoint-based adjustments
    - Refined the options that performed best and were preferred – Drill Down + Touchpoints

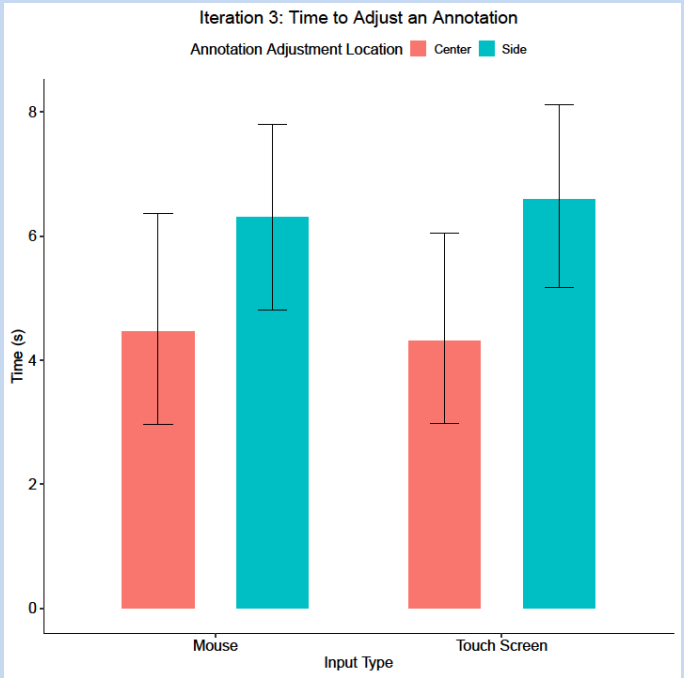


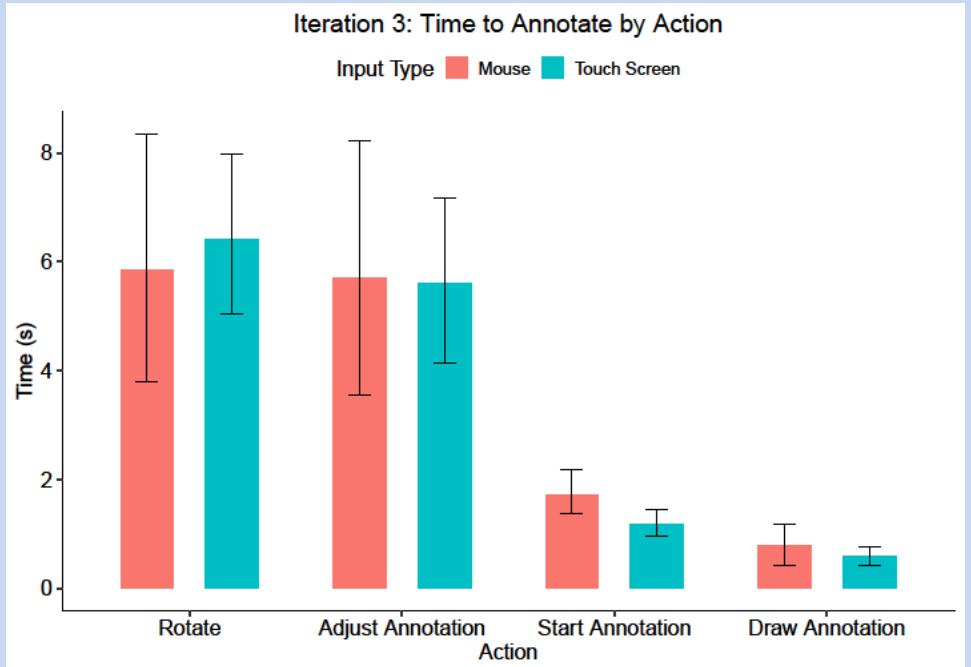
Step 4-5: CGUI Iteration 3

Unexpected Lessons - Developing 3D Annotation

Different levels of detail examining how users prefer and use different tools and elements of the interface provide insight toward the best ways to alter the design (see: left to right - Broad to Detailed examination of Time to perform Annotation and its separate components)



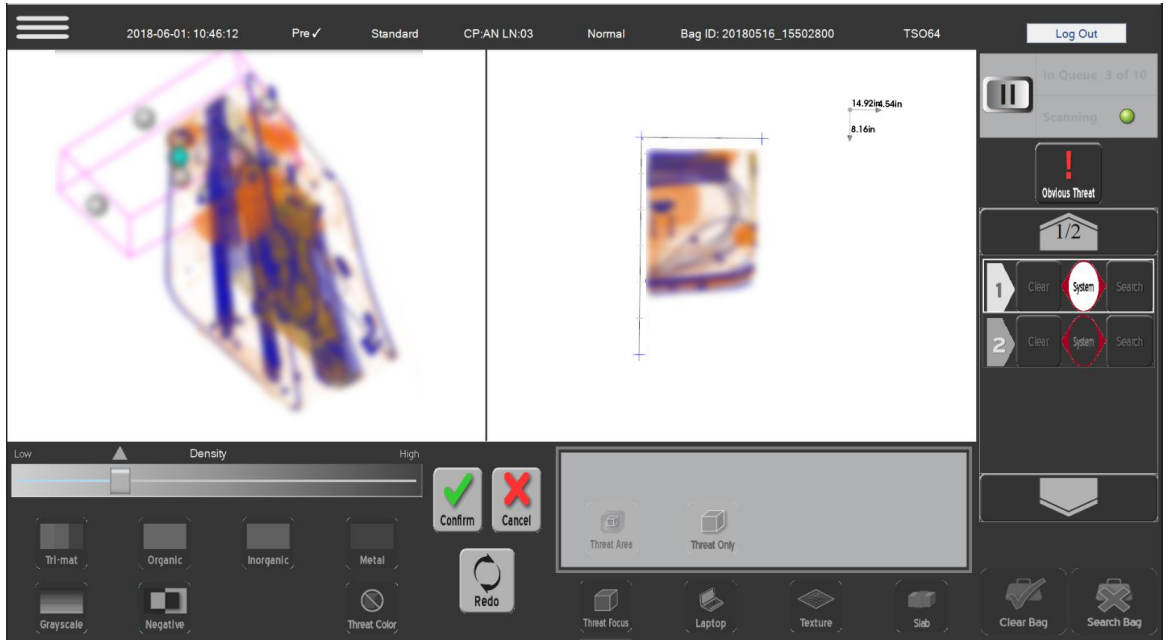




Step 4-5: CGUI Iteration 3

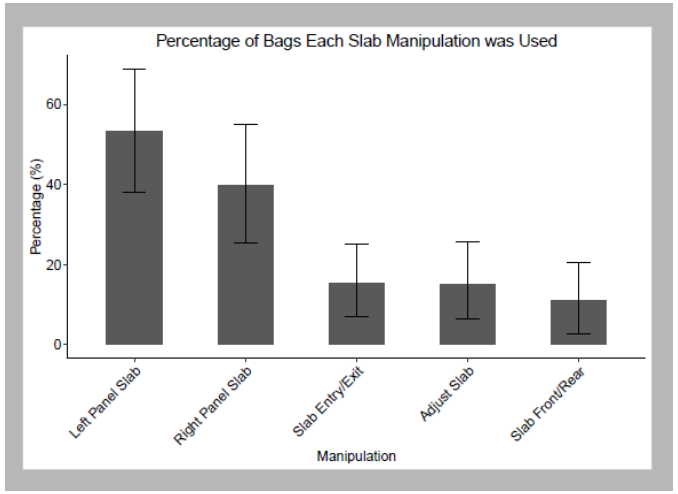
Unexpected Lessons - Developing 3D Annotation

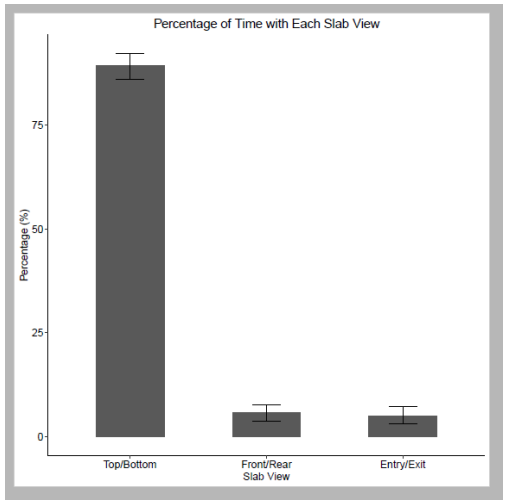
- Ultimately developed the following:
  - A 3D box that is automatically placed where the TSO clicks on the screen
    - Improved algorithms can enhance placement in the future and require less adjustment
  - Touchpoint-based
    - Touchpoints have minimum size for easy selection to manipulate
    - Touchpoints to adjust edges and move box are uniquely different for ease of use
  - Verified with both Mouse and Touchscreen modalities

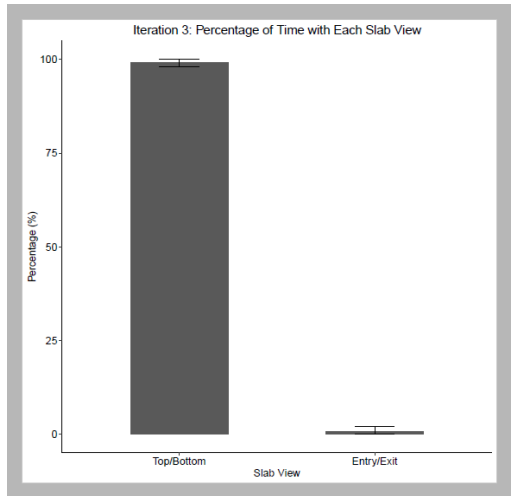


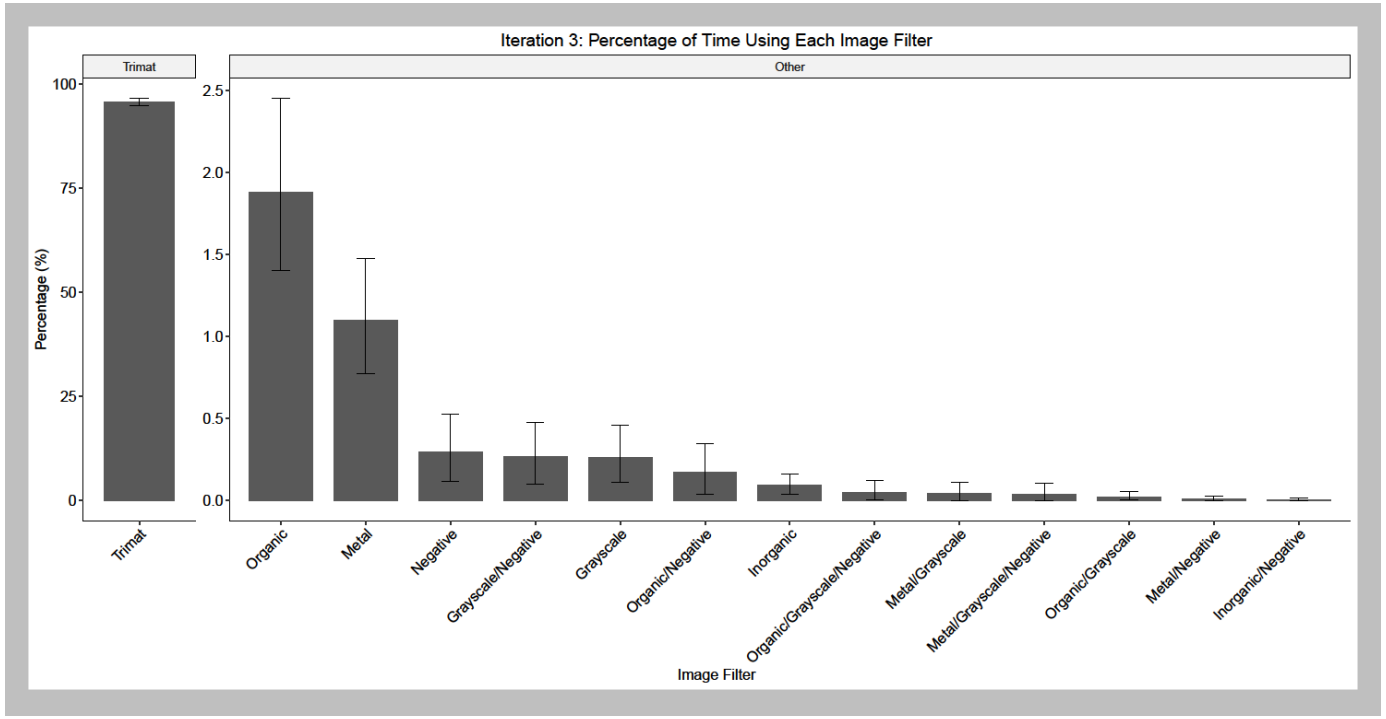
Step 4-5: CGUI Iteration 3

Iteration 3 Results Overview







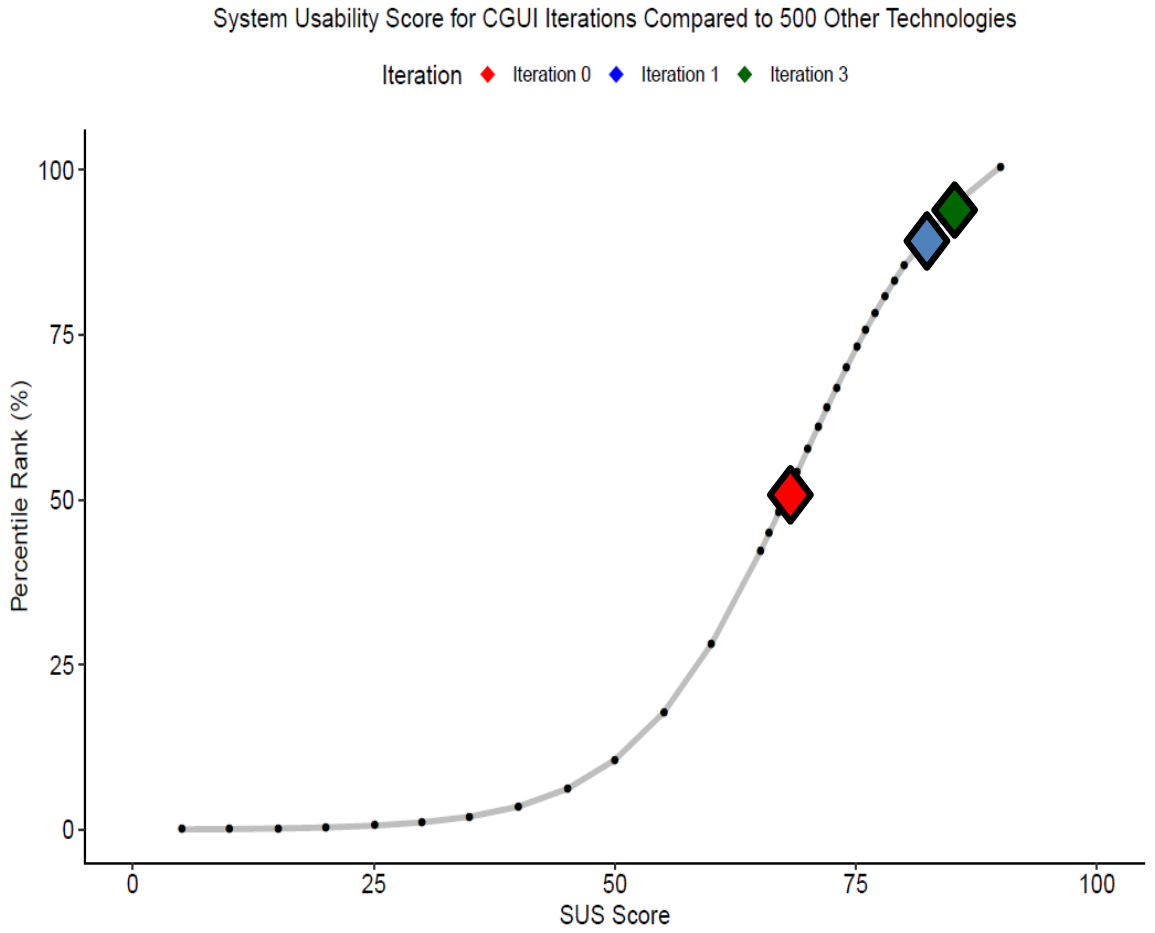


Step 4-5: CGUI Iterations

CGUI Results Summary

- Qualitative Data
  - Each additional iteration increased the perceived usability of the system.
  - TSOs felt that the CGUI was a highly usable, easy to learn system by the last iteration.

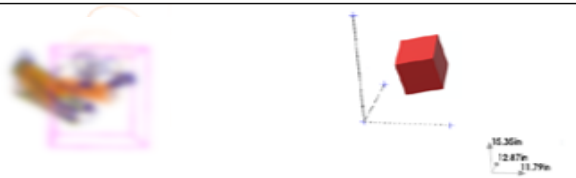
Iteration	SUS Score	Percentile	Grade	Adjective
Iteration 0	67.66	50.16	D	OK
Iteration 1-2	82.35	90.16	B	Good
Iteration 3	84.40	93.63	B	Good



Step 4-5: CGUI Iterations

CGUI Results Summary

- Qualitative Data
  - Identified additional improvements to the design as well as the rationale for performance data trends
    - Alternatives survey
      - Frequency analysis provided insight toward why some performance effects were found
      - Analysis further identified:
        - Improvements to the design
        - Filters/Tools that were preferred, necessary, and/or unnecessary
        - Methods to improve current tools (e.g., annotation)

Annotation Measurements	Description
	Inches On Axes, Inches Aligned with Axes
	Inches On Axes, Inches Horizontal to Axes
	Inches Not On Axes, Inches On Bottom
	Inches Not On Axes, Inches On Top

Annotation Measurement Display	Rank	% Preference
Inches on Axes, Inches Aligned with Axes	1	49.58%
Inches on Axes, Inches Horizontal to Axes	2	27.27%
Inches not on Axes, Inches on Bottom	3	13.23%
Inches not on Axes, Inches on Top	4	9.92%

Step 4-5: CGUI Iterations

CGUI Results Summary

- Qualitative Data
  - Identified additional improvements to the design as well as the rationale for performance data trends
    - Open-ended feedback
      - Sentiment analysis did not provide much additional insight beyond the performance and alternatives survey feedback

Step 4-5: CGUI Iterations

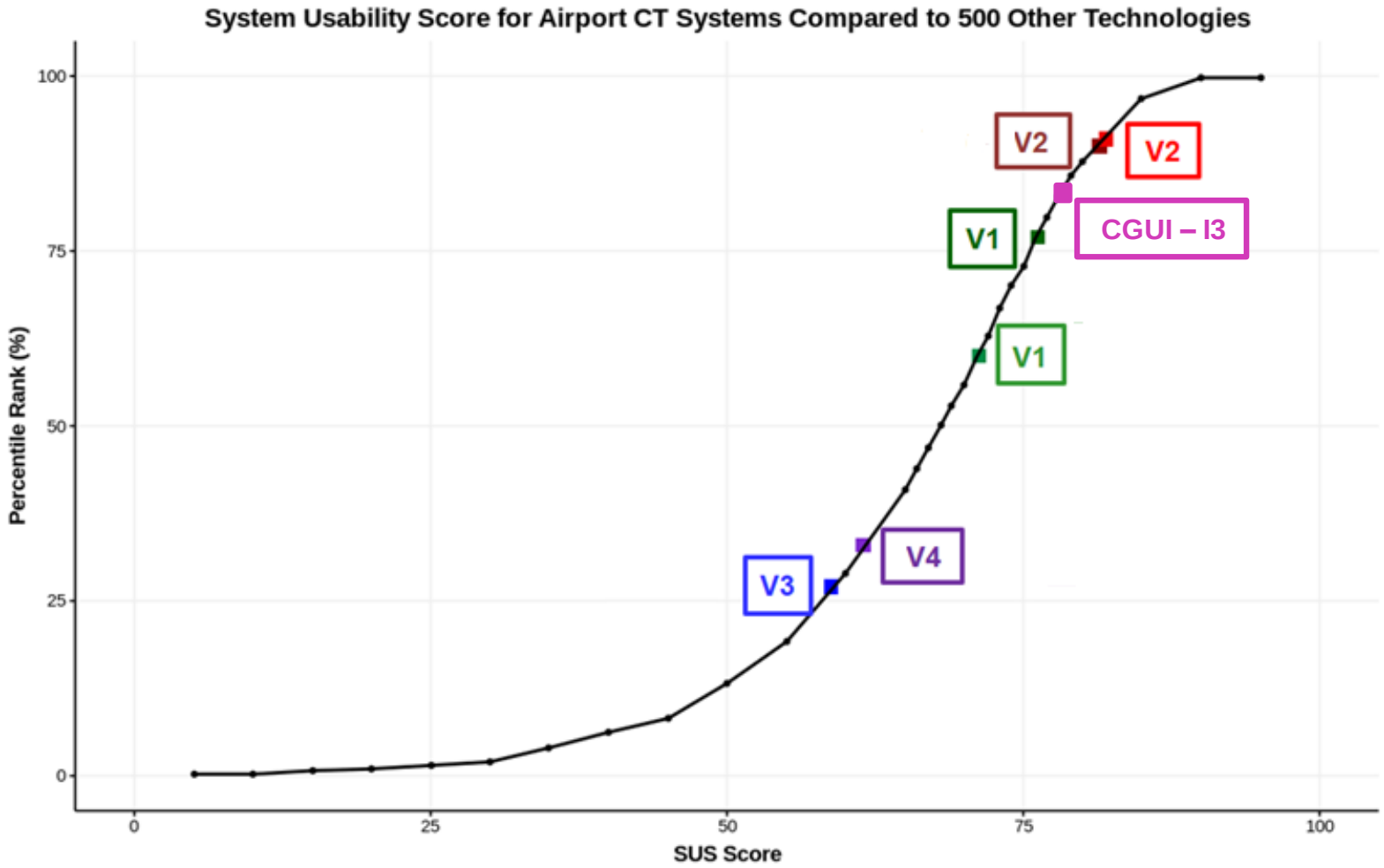
CGUI Results Overview

- Combining the quantitative and qualitative results, the following inferences were gathered:
  - The combination of survey and experimental data suggest that TSOs overestimate the frequency of use of several different requested features including:
    - Individual Density Sliders (~0% of bags)
    - Negative (< 5% of bags)
    - Grayscale (< 5% of bags)
    - Slab or Slice (< 15% of bags)
  - TSOs like seeing all the threats at the same time with the option to turn off this feature.

Step 4-5: CGUI Iterations

CGUI Results Overview

- The average CGUI score compares well with current vendor systems



Step 6: Develop Recommendations

Development of Recommendations follows satisfactory results after Steps 1-5 are complete

- Benefits
  - Recommendations are data-driven and typically more reliable
    - Level of reliability of recommendations can be quantified based on the experimental structure
      - Validating quantitative findings against qualitative findings help provide valuable insight.
        - Performance and preference trade-offs are noted
        - The rationale for some performance results are explained by qualitative input
    - Strengths and limitations of recommendations are more clearly communicated

  - Specific and broad recommendations can be identified
    - Implementable and broad solutions requiring further testing are possible to develop from this method
    - Recommendations on how to conduct future tests are easily developed and included under the same structure

Step 6: Develop Recommendations

CGUI Recommendations

- Specific
  - Specific prototype for initial implementation was developed
  - Required tools for all systems regardless of design were identified

- Broad
  - Certain metrics can serve TSA broadly
    - Measures like the System Usability Scale (SUS) can help establish standards for development and procurement
  - General training considerations were identified
    - Development of a structured training
    - How to support knowledge shift required of TSOs from one system to another
  - Adaptability of system to future projects
    - Developed initial prototype versions of how CT interface would react in a scenario where screening protocols shift (e.g., PreCheck vs Standard lanes)
    - Discussed considerations for system updates given Deep Learning advancements

Step 6: Develop Recommendations