
PSCR 2021
THE DIGITAL EXPERIENCE
#PSCR2021 • PSCR.GOV
NIST
PSCR

AUGMENTED REALITY USABILITY EVALUATION FRAMEWORK FOR PUBLIC SAFETY

NIST

#PSCR2021

PSCR

AR USABILITY TEAM

Yee-Yin Choong
Human Factors Scientist

Kurtis Goad
Human Factors Scientist

Kevin Mangold
Computer Scientist

Samia Benjida
Foreign Guest Researcher

Visualization and Usability Group — Information Technology Laboratory

NIST

#PSCR2021

PSCR

DISCLAIMER

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately.

Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

*** Please note, unless mentioned in reference to a NIST Publication, all information and data presented is preliminary/in-progress and subject to change**

AGENDA

- Project Overview
- AR Usability Evaluation Framework for PSCR
- Framework Walkthrough – Fire Fighting scenario
- Usability Test Console
- CONCLUSION

PROJECT OVERVIEW

- Augmented Reality has the potential to benefit first responders
- Usability consists of three main aspects:
  - Effectiveness
  - Efficiency
  - User satisfaction
- This framework will provide guidance on:
  - What types of data should be collected
  - How that data can be analyzed

WHAT IS AUGMENTED REALITY (AR)?

- “an interactive experience of a real-world environment where the objects that reside in the real world are enhanced by computer-generated perceptual information” (Wikipedia, March 17, 2021)
- Digital enhancements added to the real world
- Examples:
  - Digitally testing out a new couch before you make a purchase
  - Using a GPS to find directions to a new restaurant

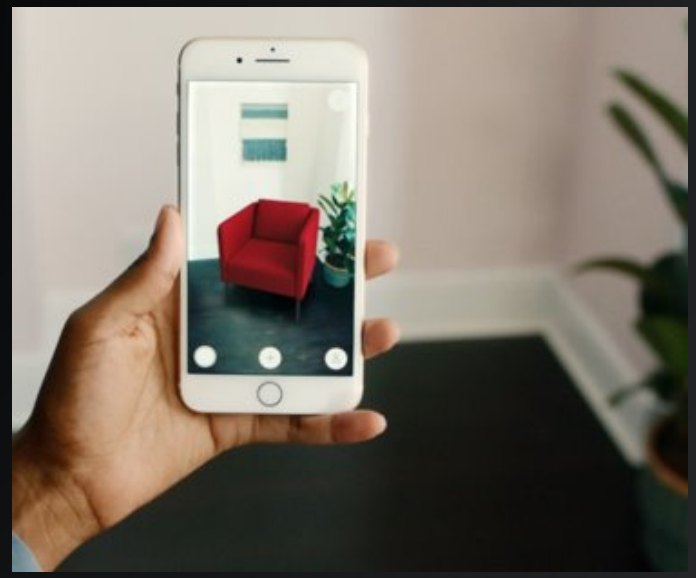


Image Source: https://www.wired.com/story/ikea-place-ar-kit-augmented-reality/

HOW DOES AR WORK?

- Two main experiences:
  - Smartphone video
  - Head-mounted displays (headsets or smart glasses)
- Uses various computer-vision and machine learning algorithms to analyze what the user sees to more realistically enhance what is relayed to the user.

Image 1 Source: https://4experience.co/hololens-2-vs-hololens-1-whats-new/
Image 2 Source: https://www.abiresearch.com/blogs/2020/03/05/augmented-reality-smart-glasses/

HOW CAN AR HELP FIRST RESPONDERS?

- Fire Fighters:
  - Important information could be displayed inside a firefighter's face mask so they don’t have to be fumble around with various devices attached to their person.
- Law Enforcement
  - A law enforcement officer could run a check on a driver's license without leaving the person unattended.
- Emergency Medical Services
  - An EMT can view information about a patient while on the scene when determining a treatment plan.

WHAT IS USABILITY?

EFFECTIVENESS

Accuracy and completeness with which users achieve specified goals

EFFICIENCY

Resources (e.g., time, human effort, money and materials) used in relation to the results achieved

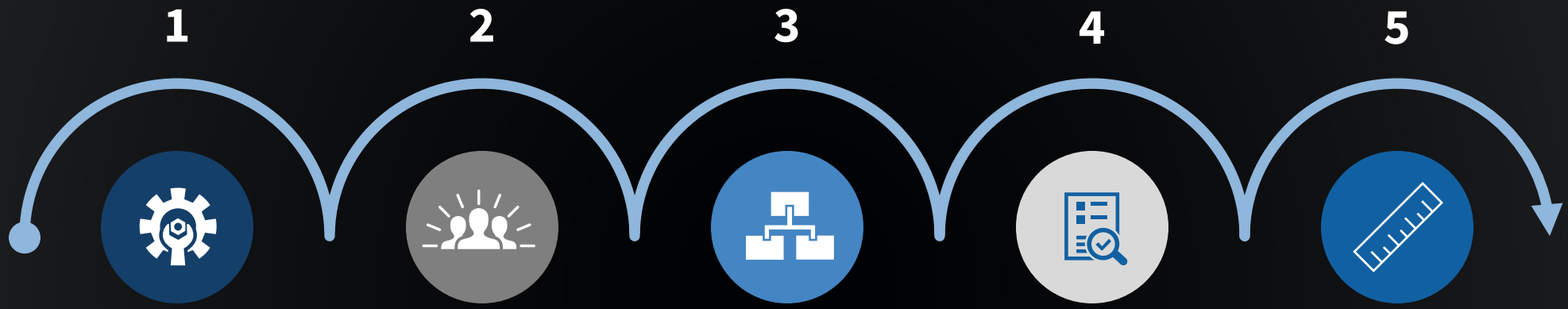
SATISFACTION

The extent to which the user's physical, cognitive and emotional responses that result from the use of a technology meet the user's needs and expectations

Usability is the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [ISO9241-11]

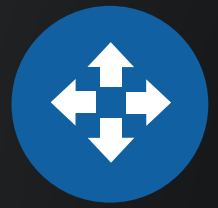
AR USABILITY EVALUATION FRAMEWORK

Assessing the Technology, Not the User



1	2	3	4	5
Evaluation Scope	***Users & Context***	***Scenario & Tasks***	***Metrics Selection***	***Measures***
Select AR functions or features to be evaluated	Identify representative users and the context of use	Develop representative scenario and select tasks for evaluation	Select applicable metrics	Define measures for the selected metrics

TASK TYPES FOR AR USABILITY EVALUATION



ACTION (A)

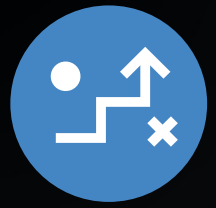
User performs a specific action in order to complete the task or make a change in the system or physical environment.

DETECTION (D)

Targets such as objects, information, and people will be placed in the environment for the user to detect.

COMMUNICATION (C)

User is required to communicate with either another user or the usability test administrator, e.g., audio, gesture, non-audio information transferred through the AR device.



MONITORING (M)

User must track specific states, conditions, objects and people over time and could be required for monitoring multiple targets simultaneously.

METRICS FOR AR USABILITY EVALUATION

- **Actors**
  - *Users* participate in usability evaluation
  - *Test admin* conducts/facilitates the usability evaluation
- **Metrics Types**
  - Performance
  - Behavioral and Physiological
  - Issues-based
  - Self-reported
- **Usability Category**
  - Effectiveness
  - Efficiency
  - Satisfaction

PERFORMANCE METRICS

Usability	Usability Metrics	Description	Data Type	Actor
Effectiveness	Task completion	Whether or not the user completes task intended	Binary success or Levels of success	User Test Admin
	Session completion	Whether or not the user completes usability session	Binary success or levels of success	User Test Admin
	Completeness	Ratio of events completed to total events expected	Ratio	User Test Admin
	Accuracy/Errors	Frequency of user events that do (do not) cause an expected outcome	Counts	User Test Admin
	Spatial Accuracy	Correct physical interactions with real or virtual objects not missing contact with the objects	Counts or ratio	User Test Admin
	Event deviation	Events performed by the user that do not aid in completing the search task	Counts or ratio	User Test Admin

PERFORMANCE METRICS

(continued)

Usability	Usability Metrics	Description	Data Type	Actor
Efficiency	Time-on-Task	Time spent performing a task	Task duration	User
	Time until Event	The time between a predefined stimuli presentation (e.g., visual, auditory) and the start of a user event	Task duration	User
	Time-on-Session	Time spent performing the usability session	Session duration	User
Effectiveness Efficiency	Learnability	Whether user's performance differ (improve or degrade) over time	Multiple trials over time	User

BEHAVIORAL AND PHYSIOLOGICAL METRICS

Usability	Usability Metrics	Description	Data Type	Actor
Effectiveness	Eye Tracking–Scan patterns	The order or pattern in which the user looks at while completing a task	Eye-tracking heat maps	User
	Mental workload	An index of questions answered by the user in order to assess mental workload for a task or the session	NASA-TLX	User
Efficiency	Eye Tracking–Dwell time	Duration of eye gaze directed at a specific target	Duration	User
	Eye Tracking–number of fixations	Frequency of instances of eye gazes directed at a specific location or object	Counts	User
	Communication effort–Speaker turns	Number of turns in conversation between two speakers	Counts	User
	Communication Effort–Words spoken	Number of words spoken by one user	Counts	User
	Communication Effort–Grounding questions asked	Number of questions asked to another user in order to help understand information presented by the system or in the environment of the user	Counts	User

BEHAVIORAL AND PHYSIOLOGICAL METRICS

(continued)

Usability	Usability Metrics	Description	Data Type	Actor
Effectiveness Efficiency	Verbal	User's verbal interactions with test admin during the session	Observations	User Test Admin
	Nonverbal	User's nonverbal information observed by test admin during the session	Observations	User Test Admin
	Facial expressions	User's facial expressions observed by test admin during the session	Observations	User Test Admin

SELF-REPORTED METRICS & ISSUES-BASED METRICS

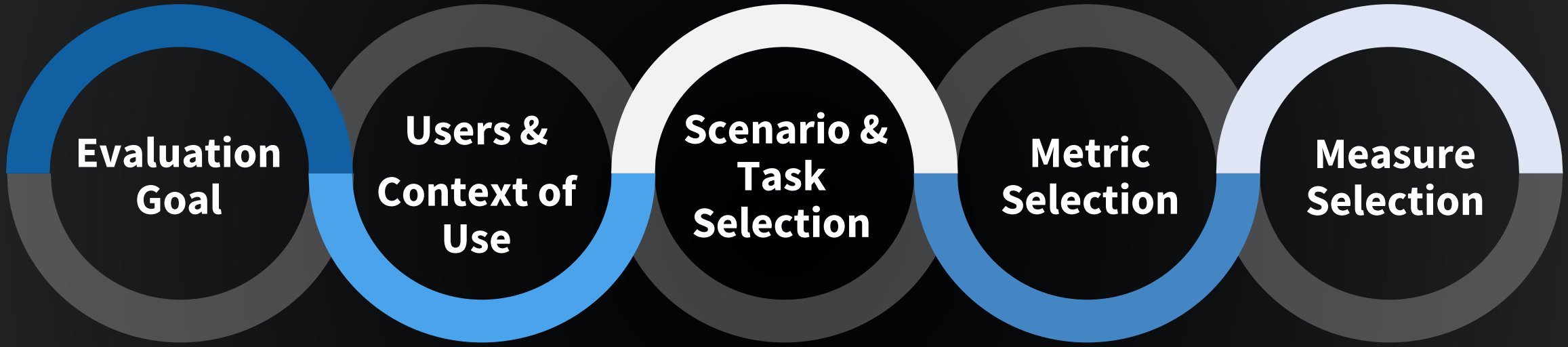
- Self-reported Metrics

Category	Usability Metrics	Description	Data Type	Actor
Satisfaction	Post-task Post-session	• Ease of Use • Task and Content Specific Questions • Perception of Outcomes/Interactions • Comfort • Learnability	Scale ratings and/or open-ended	User
	Pre-Session Expectations	An index of questions answered by the user, before using the system to assess the user's expectations about the system prior to using it	Scale ratings and/or open-ended	User

- Issues-based Metrics

Category	Usability Metrics	Description	Data Type	Actor
Effectiveness ***Efficiency***	Identify issues and severity ratings	Usability issues identified by the Test Admin during the session	Counts and severity ratings	Test Admin

FRAMEWORK STEPS



Select AR functions or features to be evaluated

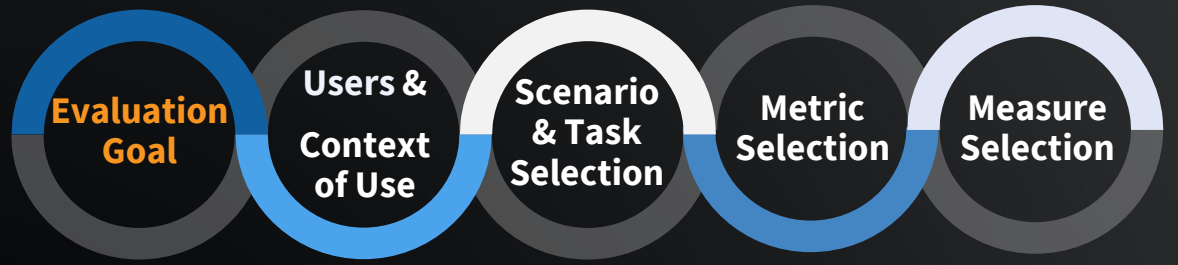
Identify representative users and the context of use

Develop representative scenario and select tasks for evaluation

Select applicable metrics

Define measures for the selected metrics

FRAMEWORK EXAMPLE



HEADS UP DISPLAY

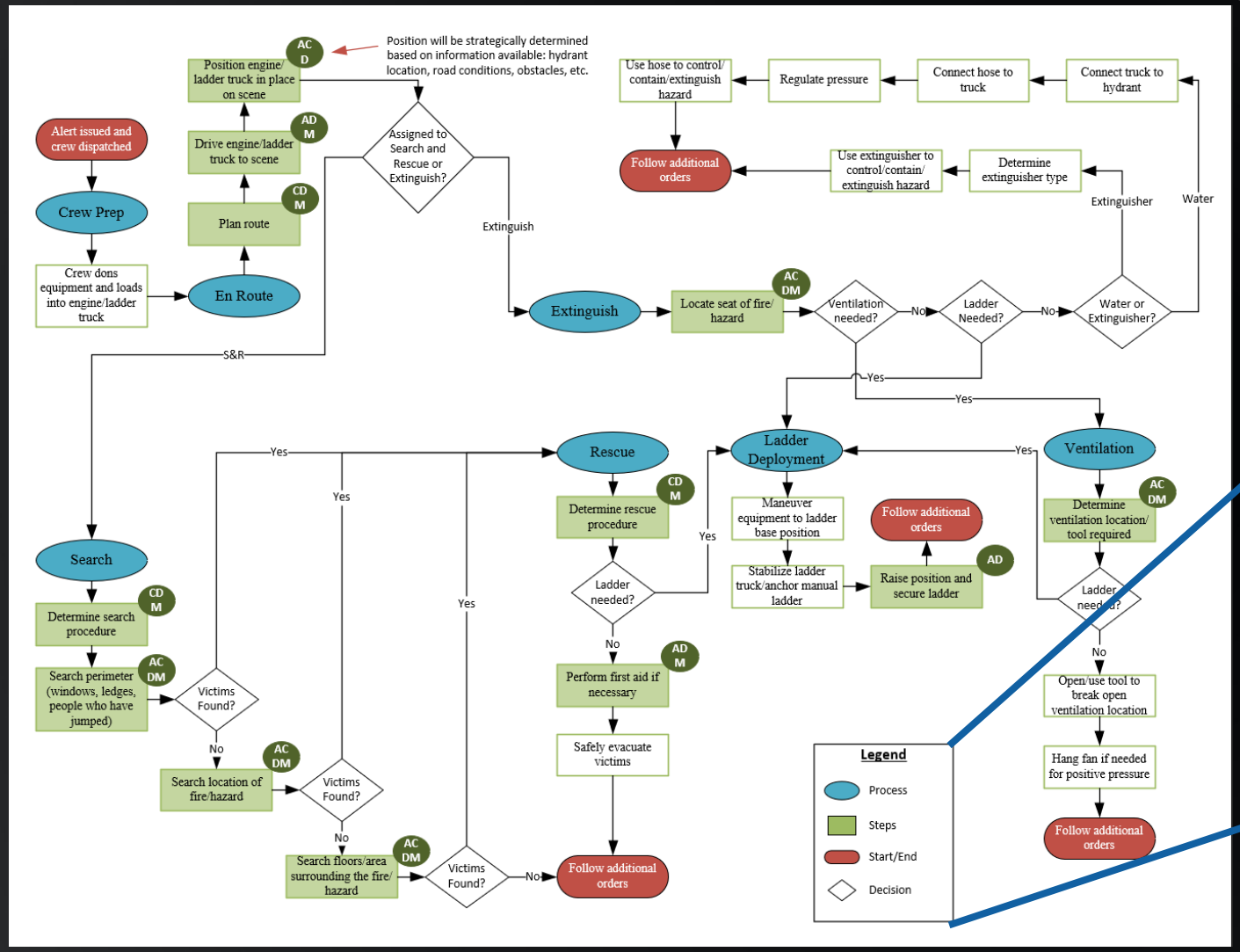
A HUD was designed for firefighters to use when responding to emergencies. The HUD uses AR components to aid in firefighters' missions.

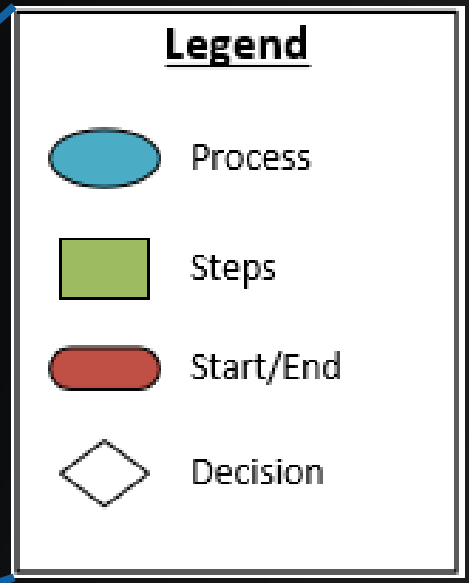
BOOTS ON THE GROUND FIREFIGHTER

A boots on the ground perspective was chosen for evaluation of the HUD.

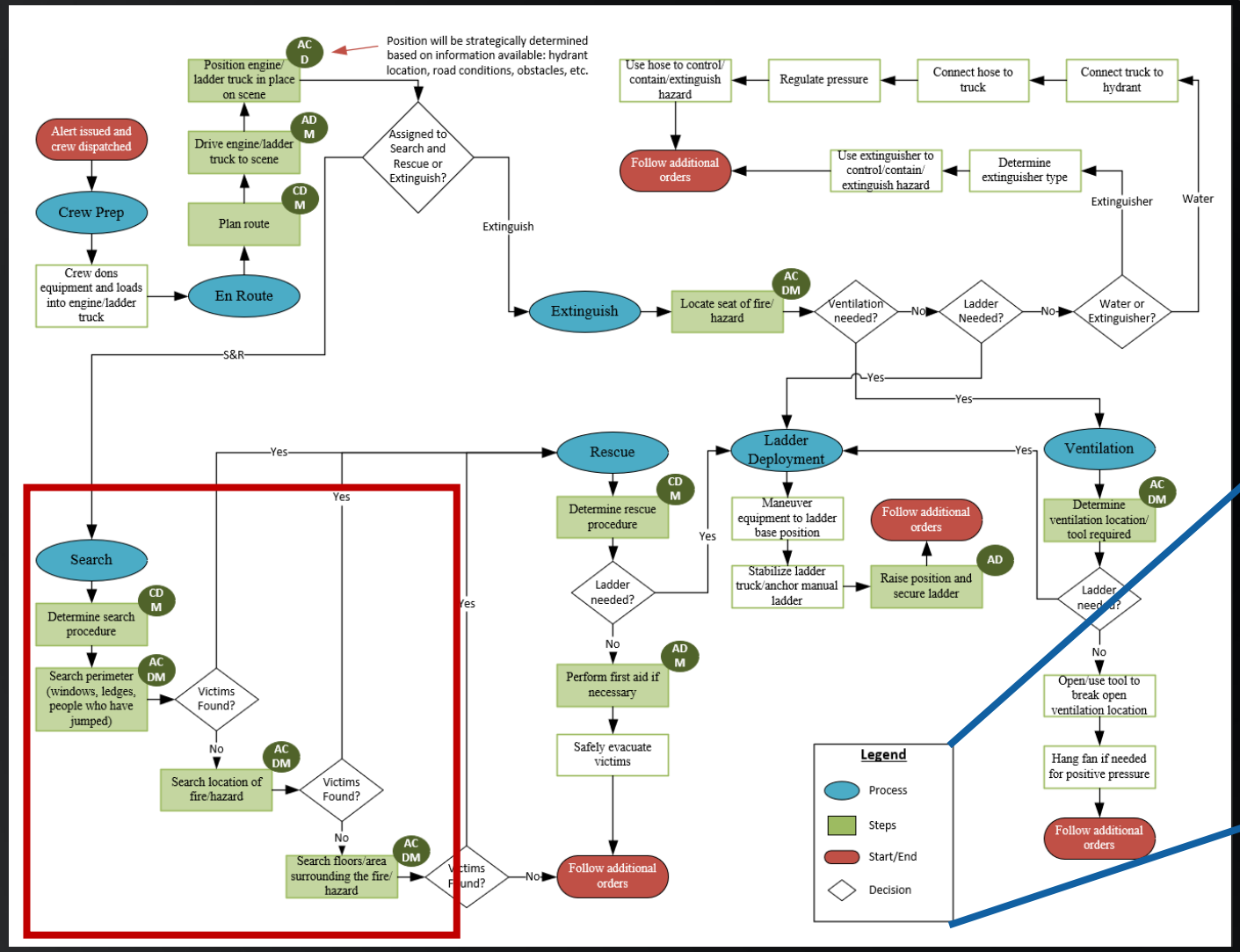
A fire was reported to have started in an apartment in a building. The local fire department was alerted and dispatched, a crew donned their gear and loaded onto a fire engine and a ladder truck, and they drove to the scene following a route planned by the command officer. Upon arrival, smoke can be seen coming from the open 2nd story apartment window of the two-story building. It was reported to the command officer by other building residents that they did not know whether the resident(s) of the apartment on fire was in the building or not.

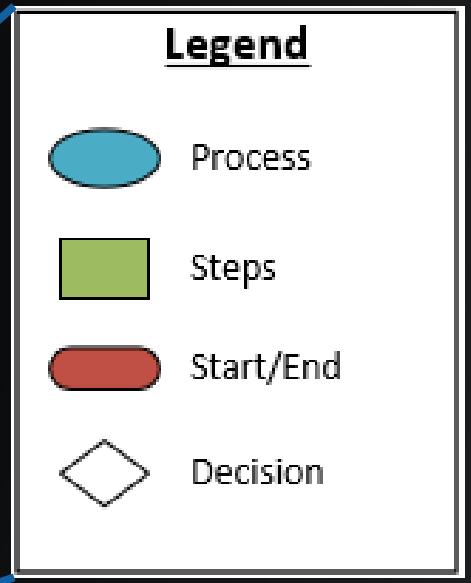
PROCEDURAL ANALYSIS DIAGRAM



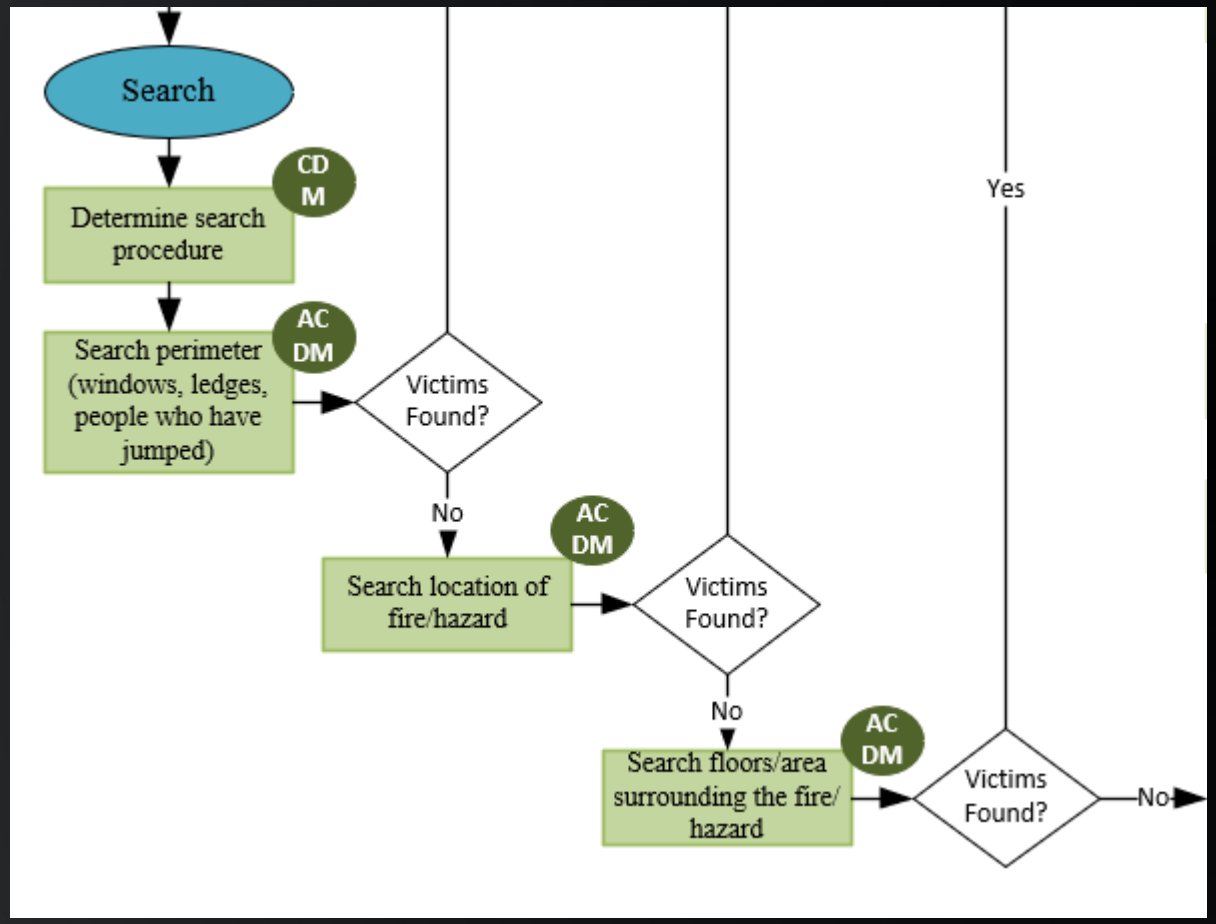


PROCEDURAL ANALYSIS DIAGRAM

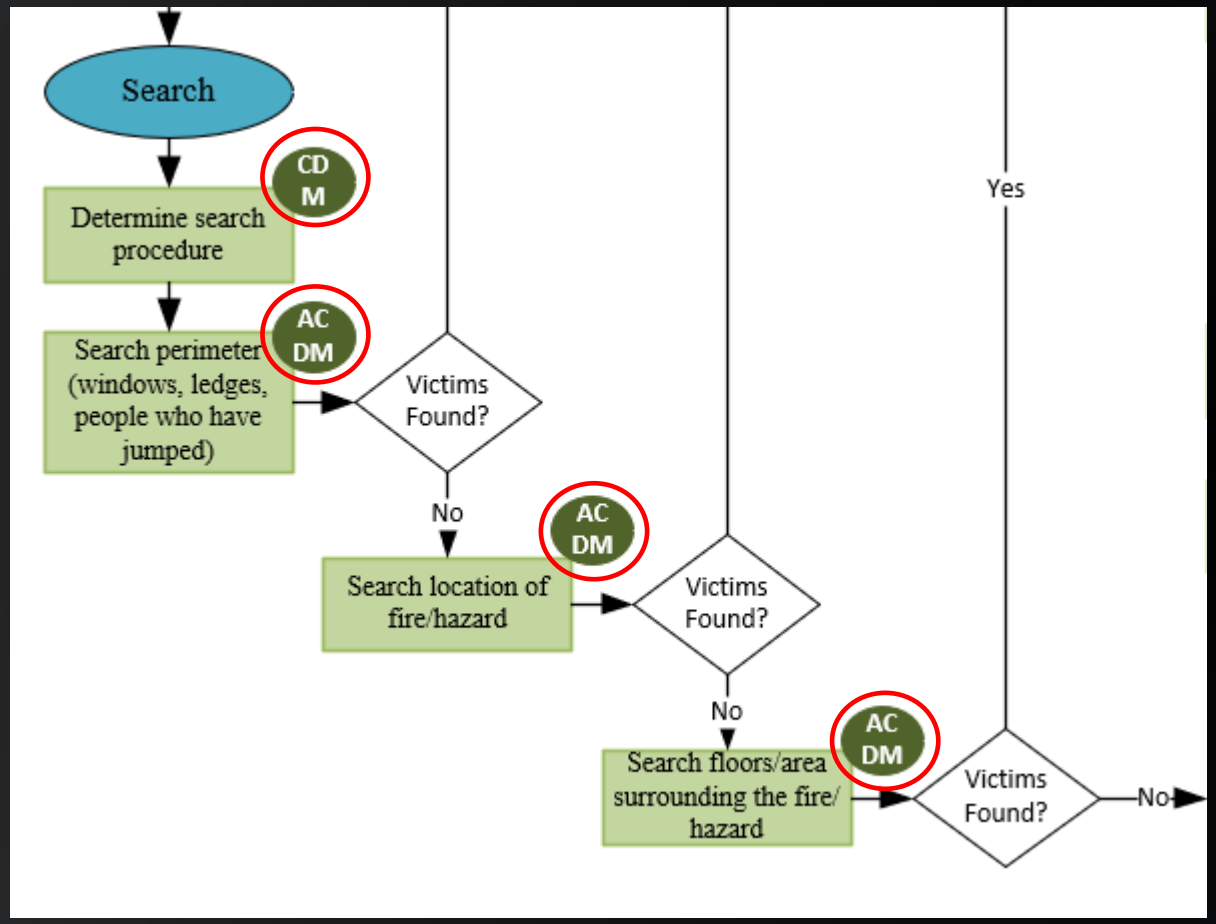




PROCEDURAL ANALYSIS DIAGRAM

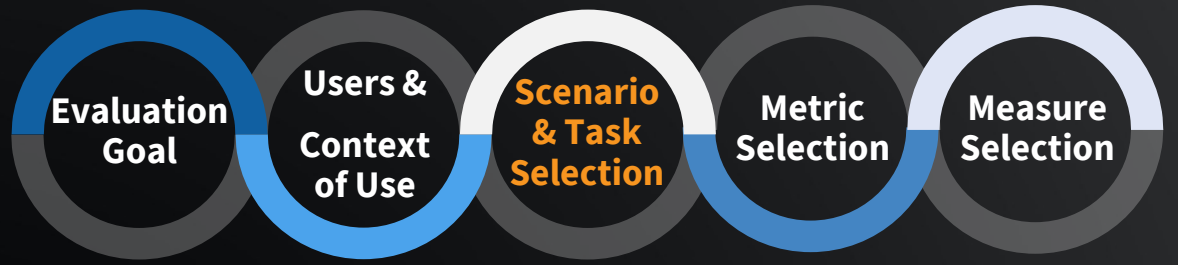


TASK TYPES



- Action
- Communication
- Detection
- Monitor

TASK SELECTION



ACTION

- **Search Location**
- Open/Close Doors
- Move Obstacles

COMMUNICATION

- Report Environmental conditions
- Report Victims
- Report Self Status

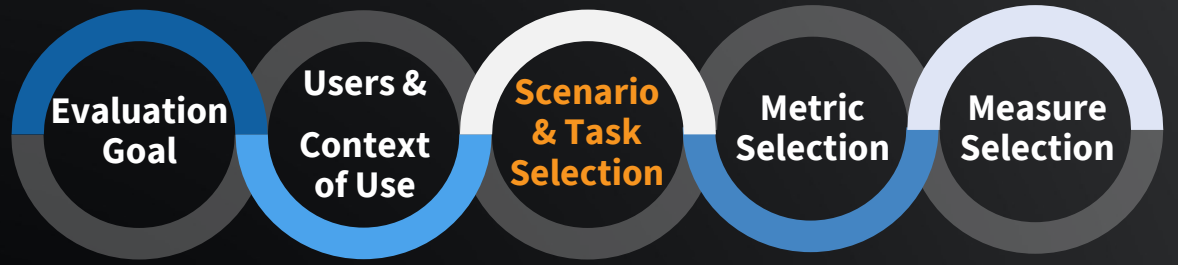
DETECTION

- Detect Hazards
- Detect Victims
- Detect Building Conditions

MONITOR

- Monitor Self Status
- Monitor Building/Hazard Status
- Monitor Victim Status

TASK SELECTION



ACTION

- **Search Location**
- Open/Close Doors
- Move Obstacles

COMMUNICATION

- Report Environmental conditions
- Report Victims
- Report Self Status

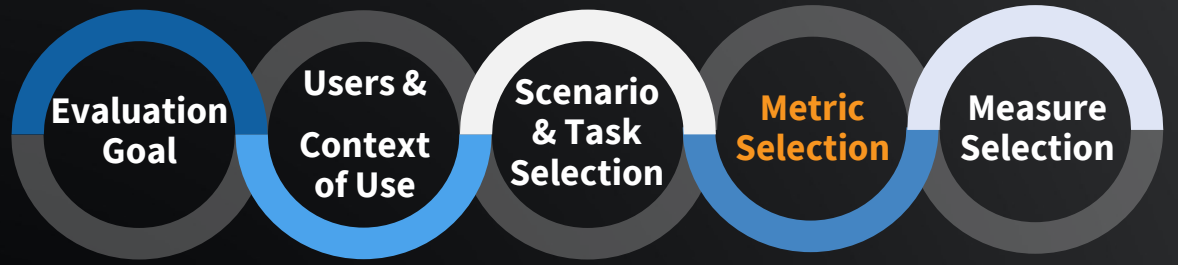
DETECTION

- Detect Hazards
- Detect Victims
- Detect Building Conditions

MONITOR

- Monitor Self Status
- Monitor Building/Hazard Status
- Monitor Victim Status

METRIC SELECTION



EFFECTIVENESS

- Task Completion
- Task Completeness
- Error
- **Eye Tracking (Sequence/Scan Patterns)**
- Event Deviation
- Spatial Accuracy

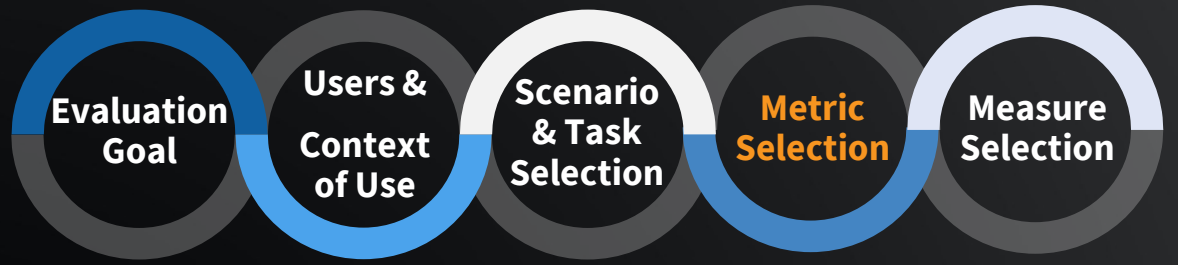
EFFICIENCY

- Time on Task
- Eye Tracking (Dwell Time)
- **Eye Tracking (Number of Fixations)**

SATISFACTION

- Post-Session Survey

METRIC SELECTION



EFFECTIVENESS

- Task Completion
- Task Completeness
- Error
- **Eye Tracking (Sequence/Scan Patterns)**
- Event Deviation
- Spatial Accuracy

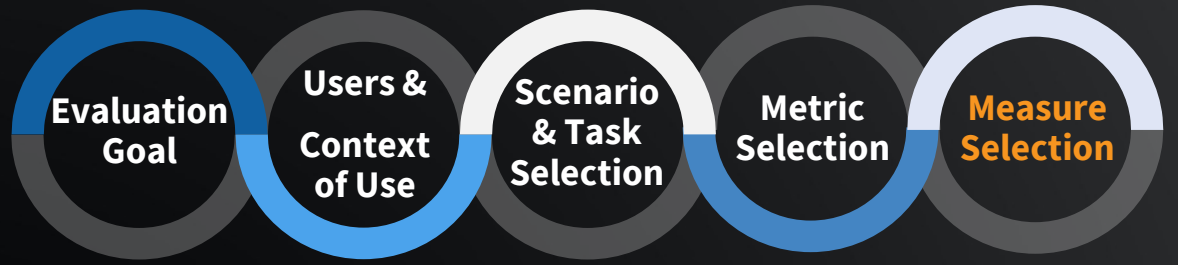
EFFICIENCY

- Time on Task
- Eye Tracking (Dwell Time)
- **Eye Tracking (Number of Fixations)**

SATISFACTION

- Post-Session Survey

MEASURE SELECTION



Eye Tracking (sequence/scan patterns)

The order in which the user looked at objects/locations while completing an action task compared to a pre-determined optimal path

Eye Tracking (number of fixations)

A Frequency count of fixations on objects/locations during a search task

Post-Session Survey

MULTIPLE TASK CONSIDERATION

- **Action**
  - Search Location

MULTIPLE TASK CONSIDERATION

- **Action**
  - Search Location
- **Communication**
  - Report Conditions

MULTIPLE TASK CONSIDERATION

- **Action**
  - Search Location
- **Communication**
  - Report Conditions
- **Detection**
  - Detect Hazards

MULTIPLE TASK CONSIDERATION

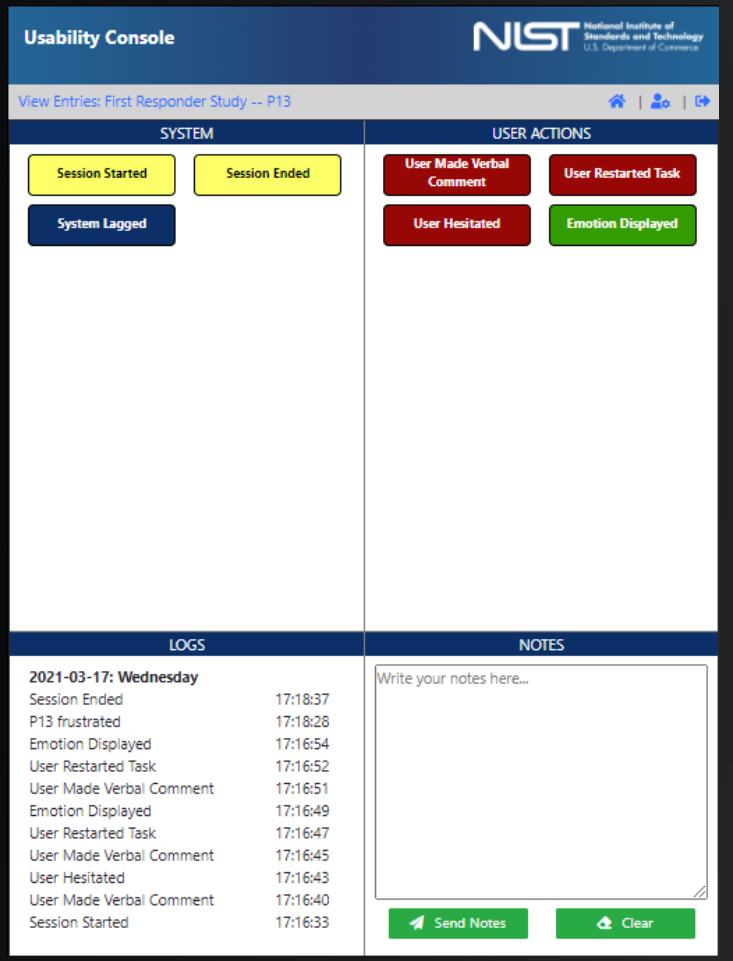
- **Action**
  - Search Location
- **Communication**
  - Report Conditions
- **Detection**
  - Detect Hazards
- **Monitoring**
  - Monitor Self Status

USABILITY CONSOLE

- A web-enabled tool used to aid usability test admins in annotating important events over the course of an evaluation session
  - Can be accessed through a computer or mobile device
- Designed to replace a laptop or paper/pen
- Keeps track of multiple participants and all of the events for a particular evaluation session
- Capable of ingesting live data from external devices

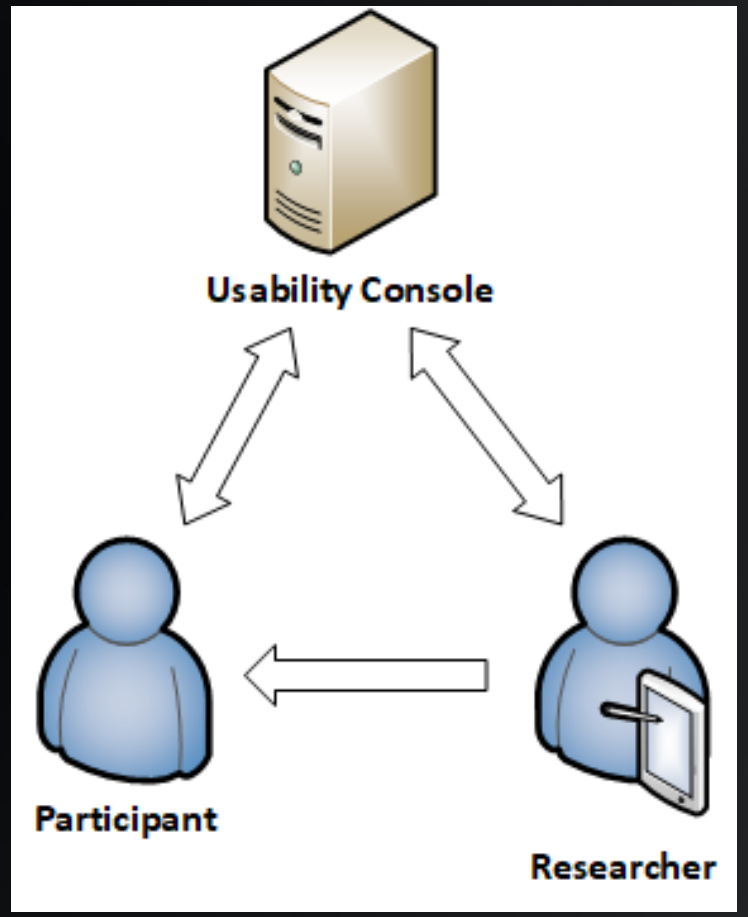
USABILITY CONSOLE BENEFITS

- Buttons can be predefined to allow for quick and easy annotation of highly anticipated events
- Ability to add custom notes or event
- Quick view of recently added events



EXTERNAL DEVICE INTEGRATION

- External devices are able to connect to the console to log events or other relevant data
  - Can be an Augmented Reality headset or other internet-capable devices/sensors



CONCLUSION

AR Usability Evaluation Framework for PSCR

Explicit structures

Consistent terminology and an initial set of usability metrics

Comparability across public safety AR research and development efforts

Sharing of usability evaluation results and help establish design guidelines

THANK YOU

usability@nist.gov

https://www.nist.gov/itl/iad/visualization-and-usability-group
https://www.nist.gov/ctl/pscr/research-portfolios/user-interfaceuser-experience