



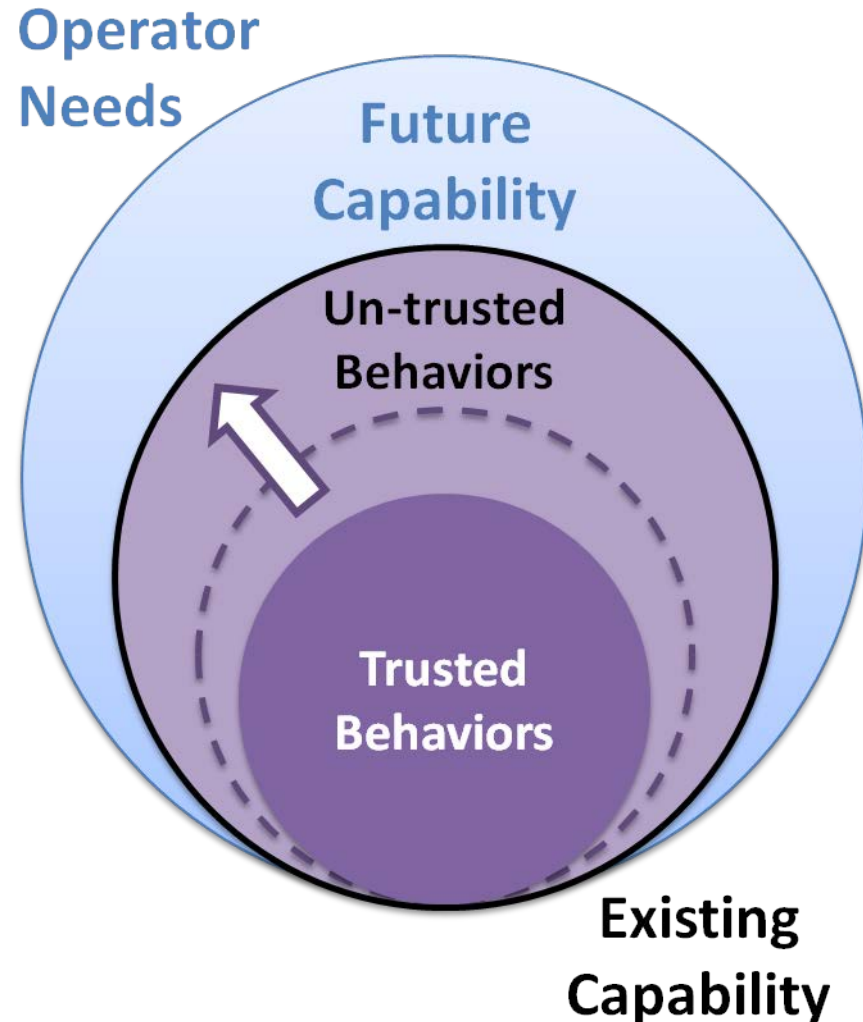
# **A Heuristic-Based Framework for Assessing Operator Trust in Autonomous Systems**

**Kim Jackson, Emily Vincent,  
Eric Jones, Zahar Prasov**

**NDIA Human Systems Conference  
February 10-11, 2015**

# Project Goal

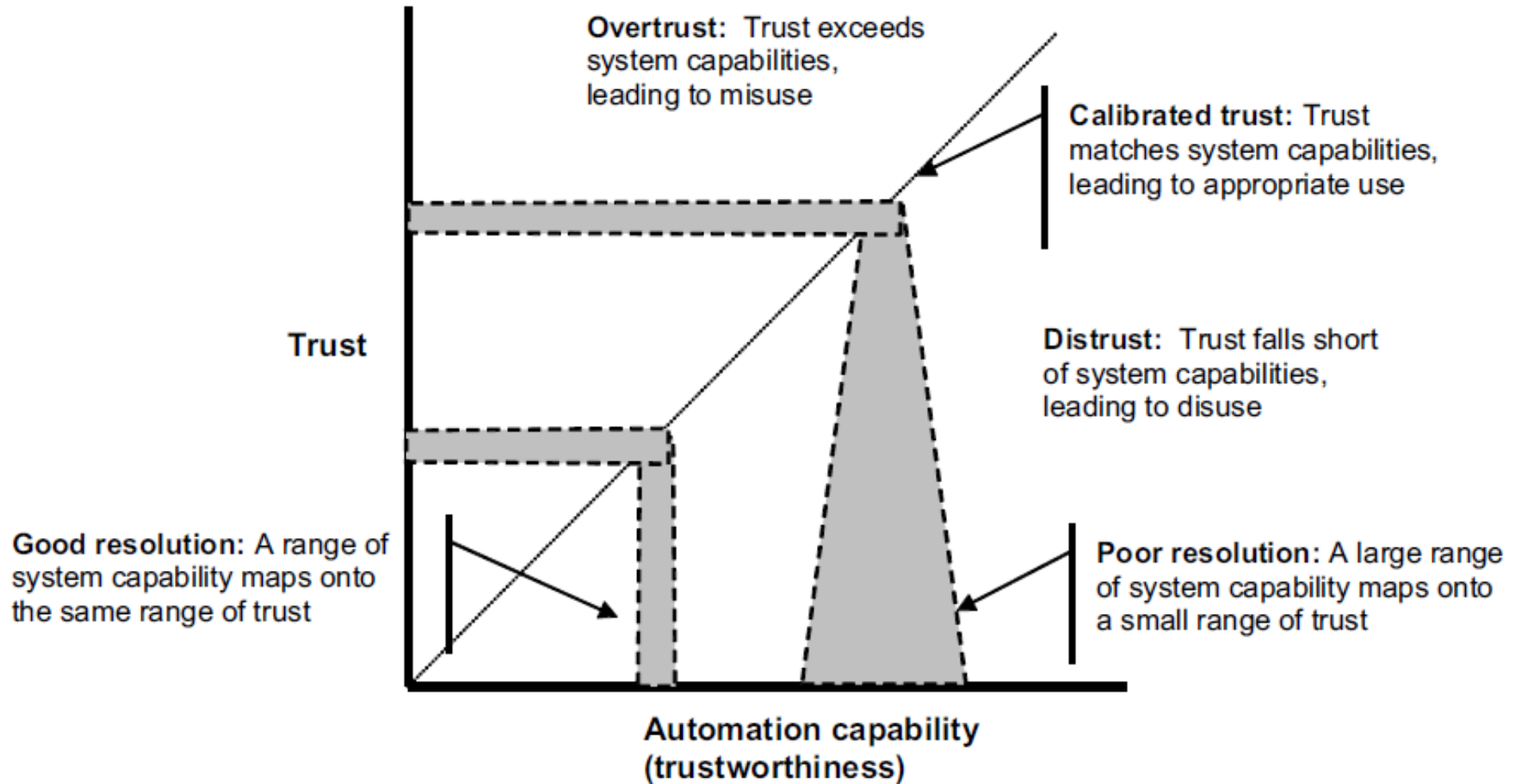
Develop actionable guidelines to determine how to design and evaluate autonomous systems that will support appropriate levels of operator trust



# What do we mean by “trust”?

- Trust (working definition): “*adoption of and reliance on the system*”
  - **Adoption:** Is the operator willing to use the system?
  - **Reliance:** Is the operator willing to rely on the system for mission-critical tasks?
- Trust needs to be appropriate for the given task.

# Mismatched expectations lead to a lack of trust



Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46(1), 50-80

# How can we build trusted (or trustable) systems?

## Needs:

- Identify unstated & unmet operator needs for trusted autonomous systems
- Understand what we know (academically) about how trust is built, measured, and understood

## Tools:

- Literature review
- Human-Centered Engineering operator study

# Data Collection

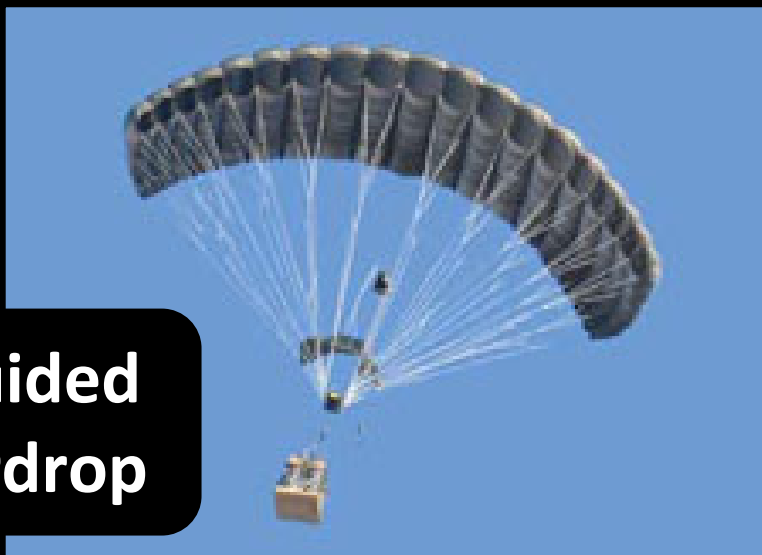
**Tigershark**



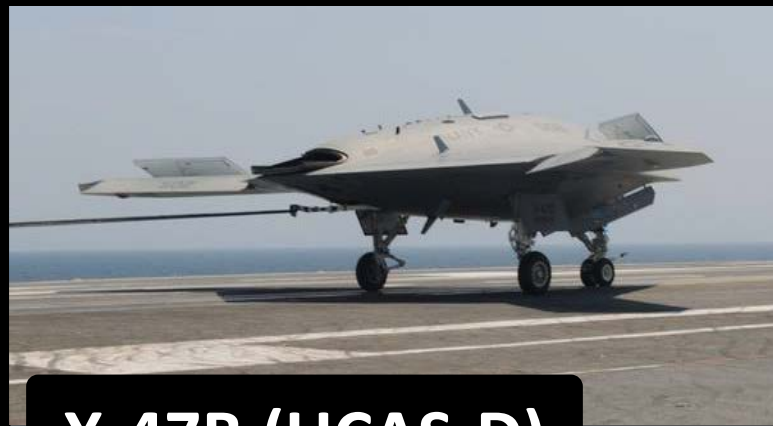
**Fire Scout**



**Guided  
Airdrop**



**X-47B (UCAS-D)**



# Synthesis



# Heuristics for Trusted Autonomy

**Visibility of Current System Behavior**

**Visibility of Probable System Behavior**

**Awareness of Latencies and Delays**

**Visibility of System Capabilities & Limitations**

**Transparency of Failure**

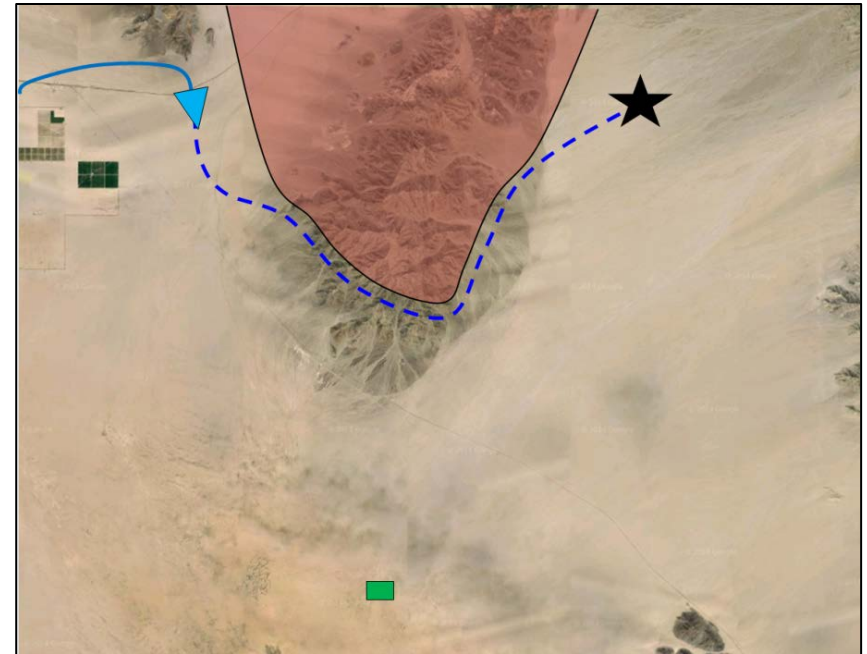
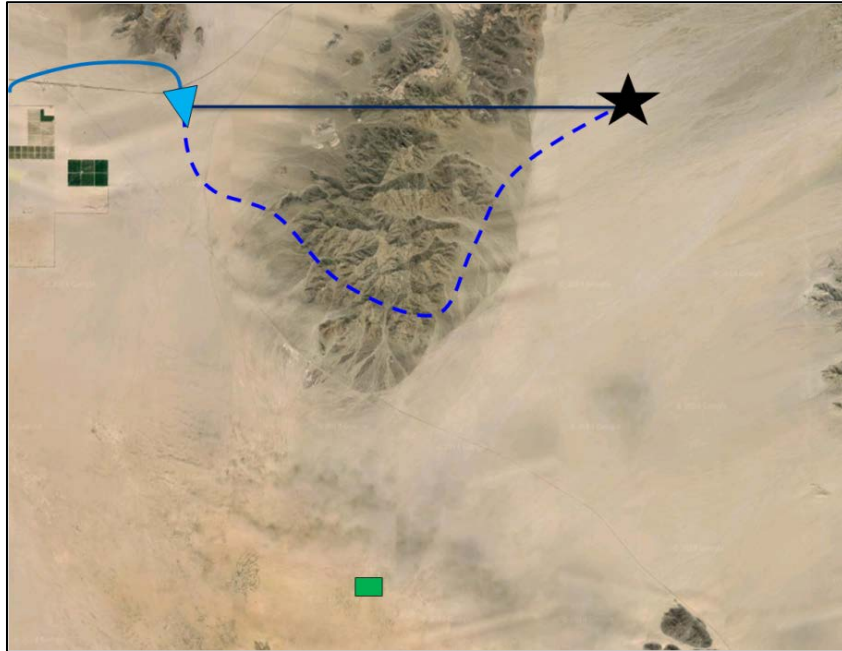
**Fit with Users and Operations**

**Accessibility of System Rationale**

- Display information about the system's decision making process in a language familiar to the user
- Include, where appropriate
  - information about the decision making algorithm
  - Alternative behaviors
  - The system's situational awareness of the environment
  - Levels of confidence
  - Algorithm performance metrics



# Example: Accessibility of System Rationale



*Planning  
constraints hidden*

*Terrain and  
Line-of-sight  
constraints shown*

# Heuristics as tools for design or evaluation

## Heuristic Evaluation

Expert evaluation method to determine whether a system adheres to each heuristic. Produces actionable information for any deficiencies along with associated severity ratings.



\*Nielsen, J. (1994). *Usability engineering*

# Standardized Scoring Examples

## Frequency

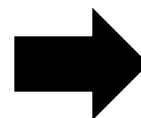
- 1 Occurs rarely - less than once per mission
- 2 Occurs once or twice per mission
- 3 Occurs once or twice per hour
- 4 Occurs many times an hour

## Impact

...

## Persistence

...



## Severity

Persistence = 1

Impact	4	1	2	3	3
	3	1	2	2	3
	2	1	1	2	2
	1	1	1	1	1
		1	2	3	4

Frequency

Persistence = 2

Impact	4	2	3	4	4
	3	2	3	3	4
	2	1	2	3	3
	1	1	1	2	2
		1	2	3	4

Frequency

...

# Next Steps

- Validate heuristics
- Refine and validate evaluation method
- Investigate extensibility to other domains (beyond autonomous vehicle operations)
- Investigate applicability to other “user groups”

# A Heuristic-Based Framework for Assessing Operator Trust in Autonomous Systems

Kim Jackson

Draper Laboratory

617-258-1543

[kjackson@draper.com](mailto:kjackson@draper.com)