Georgia Tech Research Institute

# T&E of Autonomous Systems: Challenges and Opportunities of Evaluating Trust

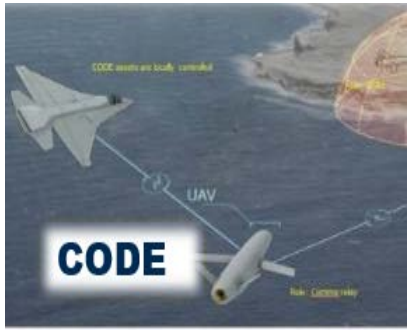
**Don Strausberger**

**Georgia Tech Research Institute**

**March 2017**

# Ongoing Autonomy Science & Technology Programs (1 of 3)

Georgia Tech Research Institute

| Domains | Ongoing S&T Efforts | Description and Reference |
|---|---|---|
| Air |  Loyal Wingman | **For both demonstrations, the unmanned aircraft is expected to safely operate in a cooperative/team configuration with a 5th generation manned aircraft from a fixed airfield, accomplish mission objectives, and return …shall be able to operate untethered from the ground without full-time direction from the manned aircraft.** RFI-AFRL-RQKH-2015-003.pdf |
| |   | **Using collaborative autonomy, CODE-enabled unmanned aircraft will find targets and engage them as appropriate under established rules of engagement, leverage nearby CODE-equipped systems with minimal supervision, and adapt to dynamic situations such as attrition of friendly forces or the emergence of unanticipated threats.** http://www.darpa.mil/program/collaborative-operations-in-denied-environment |

# Ongoing Autonomy Science & Technology Programs (2 of 3)

Georgia Tech Research Institute

| Domains | Ongoing S&T Efforts | Description and Reference |
|---|---|---|
| Sea |   | Advance unmanned maritime system autonomy to enable independently deploying systems capable of missions spanning thousands of kilometers of range and months of endurance under a sparse remote supervisory control model. This includes autonomous compliance with maritime laws and conventions for safe navigation, autonomous system management for operational reliability, and autonomous interactions with an intelligent adversary.<br>http://www.darpa.mil/program/anti-submarine-warfare-continuous-trail-unmanned-vessel |
| |   | The objective of the shipboard autonomous firefighting robot (SAFFiR) is to develop human-centric, autonomous systems for fire safety and damage control.<br>… provide shipboard situation awareness and damage control capabilities through … advanced human-robot interaction technologies … to enable teaming with human firefighters<br>https://www.onr.navy.mil/en/Media-Center/Fact-Sheets/Shipboard-Robot-Saffir.aspx |

# Ongoing Autonomy Science & Technology Programs (3 of 3)

| Domains | Ongoing S&T Efforts | Description and Reference |
|---|---|---|
| Land |  AMAS | **The Autonomous Mobility Applique System (AMAS) provides a low-cost/low-risk, kit-based solution to retrofit active safety, semi-autonomy and autonomy capability onto any vehicle in the military's logistics fleet. AMAS offers Driver Warning/Driver Assist functionality, Leader-Follower convoy operations, Waypoint following capabilities and provides growth to fully autonomous operations.** http://www.lockheedmartin.com/us/products/amas1/mfc-amas-overview.html |

# The T&E Challenge

**How am I as a Test Professional going to generate sufficient evidence through Test & Evaluation to assure the commander, operators, and others that observe or interact with the platform of the following:**

## Examples


**The Commander's newly deployed fleet of 40 fully autonomous surface vessels (actively engaged in 90 day continuous anti-submarine warfare missions) will not cause damage to their own fleet or be the center of an international incident as a result of at sea collisions.**


**The behaviors of the autonomous 5th Generation fighter "Wingman" will maintain air collision avoidance protocol, execute evasive maneuvers as necessary, and establish the requisite trust in the lead pilot to successfully execute the mission.**


**Those that encounter the deployed leader-follower convoy observe the vehicle's behavior as "appropriate" in all instances, so the observers can carry out their tasks as if the vehicles were fully manned.**

# Autonomy Testing – Tackling the Elephant coming to the range near you!

Georgia Tech Research Institute

## Purpose

- Drive thought/discussion in the Test Community towards innovative solutions to evaluate trust in the context of autonomous system T&E

## Agenda

- Autonomy Definitions & Terminology
- When/why is autonomy testing different?
- Human-Autonomy relationships and trust implications
- Trust Calibration
- Two key challenges
- Conclusion

# Autonomy Definition*

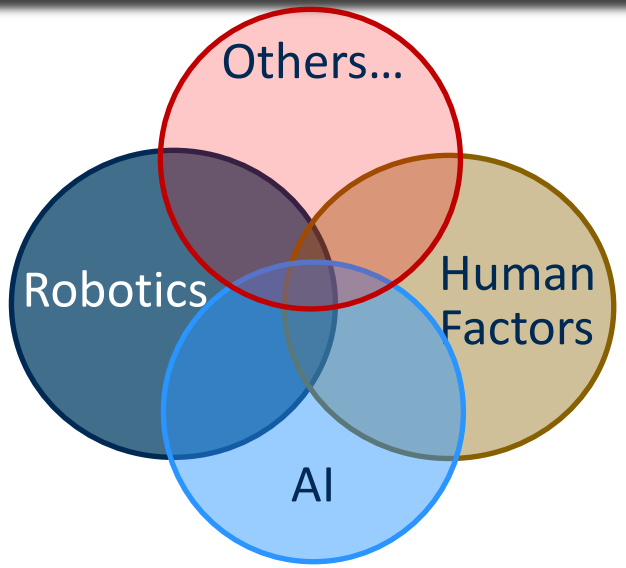
- **Automation**
  - The system functions with no/little human operator involvement; however, the system performance is **limited to the specific actions** it has been designed to do. Typically these are **well-defined tasks** that have **predetermined responses**.
  - **Rule-based responses**
- **Autonomy**
  - The system has a set of **intelligence-based capabilities** that allow it to **respond** to situations that were **not pre-programmed** or anticipated prior to deployment. Autonomous systems have a degree of **self-government** and self-directed behavior.
  - **Decision-based responses**

*As defined in the DoD Autonomy Community of Interest (COI) Test and Evaluation Verification and Validation Working Group Technology Investment Strategy 2015-2018

# Terminology

Georgia Tech Research Institute

**Different disciplines have developed different terminology in closely related areas**




**Terminology does not map one-to-one across disciplines, however they can be associated with the OODA loop to provide a common reference.**

**Goal-based**
**Rule Based**

**Perception**
**Comprehension**
**Projection**


**Beliefs**
**Desires**
**Intent**

**Reflective**
**Deliberative**
**Reactive**

**Situational Awareness**
**Situational Understanding**
**Situational Assessment**

# Why is Autonomy Testing Different?

Georgia Tech Research Institute

## Human

> Memory

> Projection

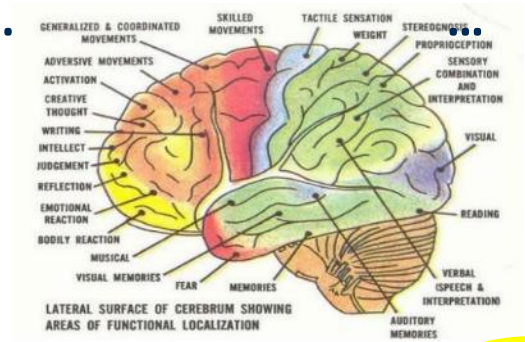
> Sensing

> Multi Sensor Interpretation

> Judgement

> Skilled Movements

...




...

> Human Values

> Empathy

> Fight or Flight

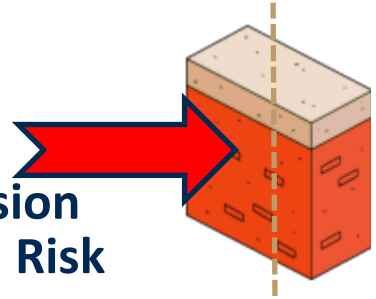
...

> Trust

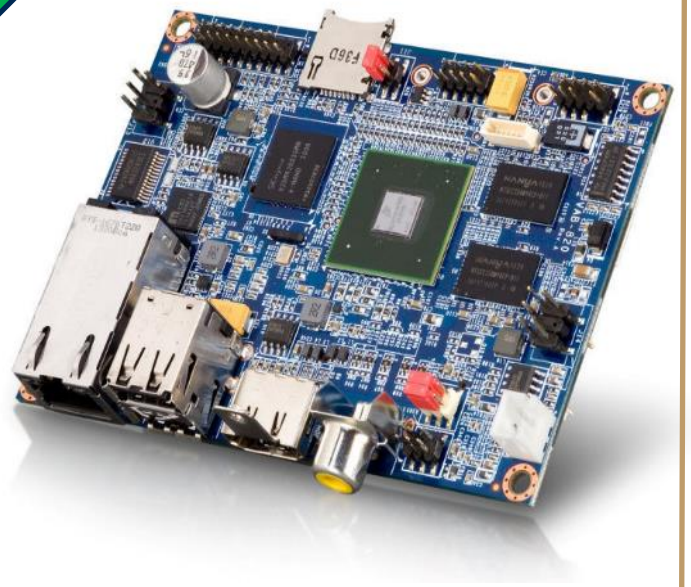
> Compassion

> Cost and Risk

...



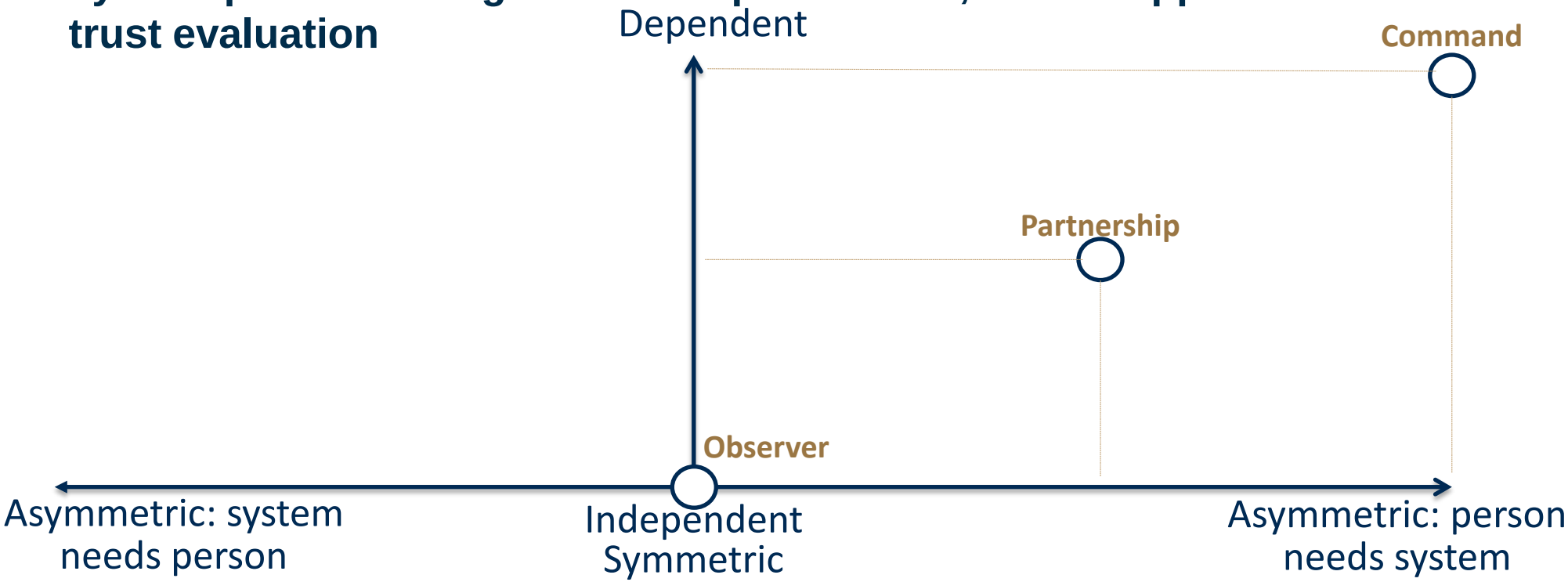
## Machine



**Several of the core conditions that guide a human's decision making on a daily basis are unaccounted for in today's autonomy. This fundamental aspect of autonomy significantly impacts the "acceptance criteria" of autonomy, and indirectly (but significantly) impacts T&E.**

# Human/Autonomy Relationships and Trust Implications

- **Human/Autonomy (H/A) relationships can be described/decomposed via ……**
  - **Interdependence: how much the human relies on the autonomy**
  - **Symmetry: the ratio of how much the human relies on autonomy, and (conversely) how much autonomy relies on the human**
- **Describing/decomposing a human's relationship with the autonomous system provides insight to the importance of, and an approach toward trust evaluation**



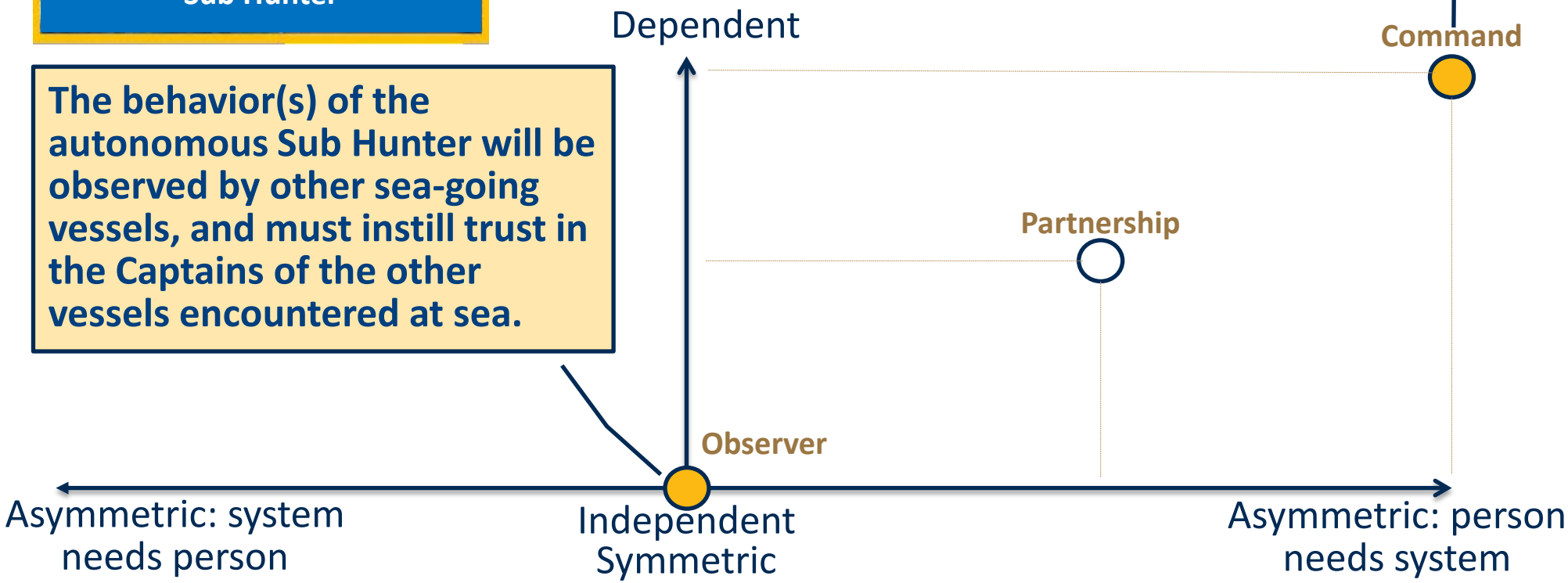
Georgia Tech Research Institute

# Example 1

Sub Hunter

**The Commander must trust the Sub Hunter understands the mission intent and is effectively executing.**

**The behavior(s) of the autonomous Sub Hunter will be observed by other sea-going vessels, and must instill trust in the Captains of the other vessels encountered at sea.**

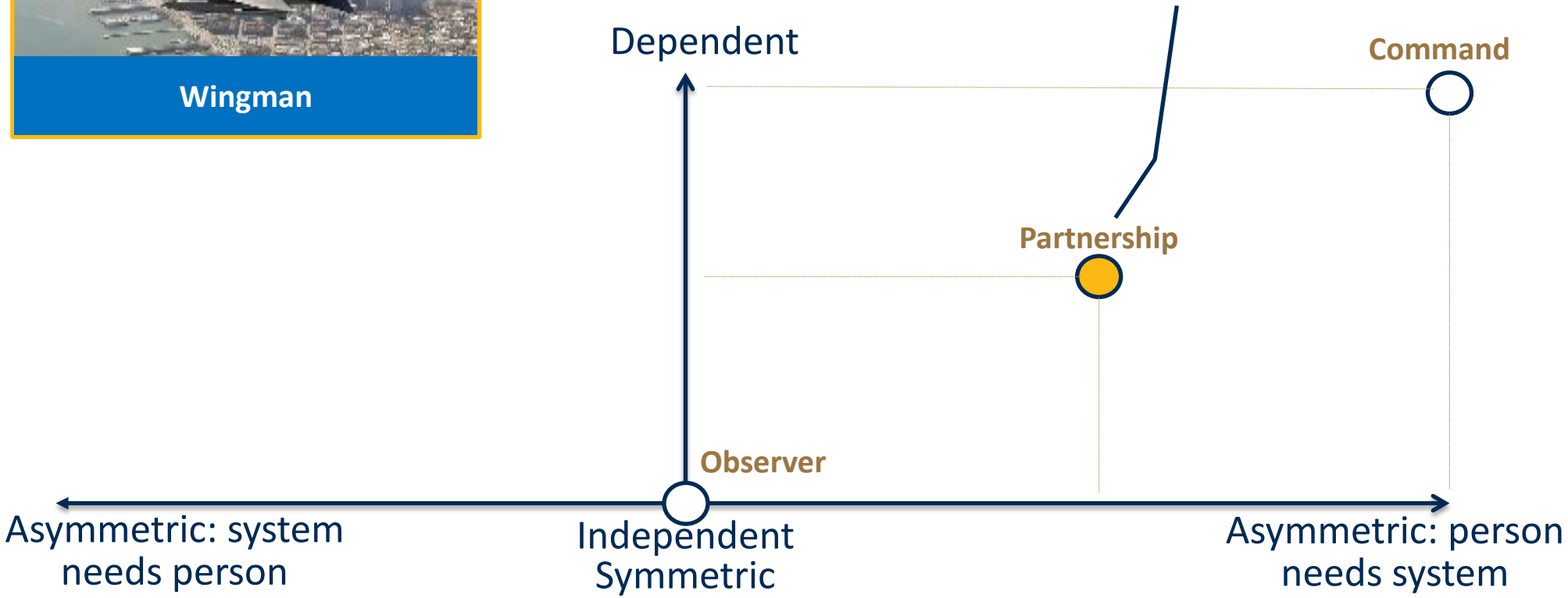


Georgia Tech Research Institute

# Example 2

Wingman

**The manned fighter is closely coupled to the autonomous wingman to execute the mission.**



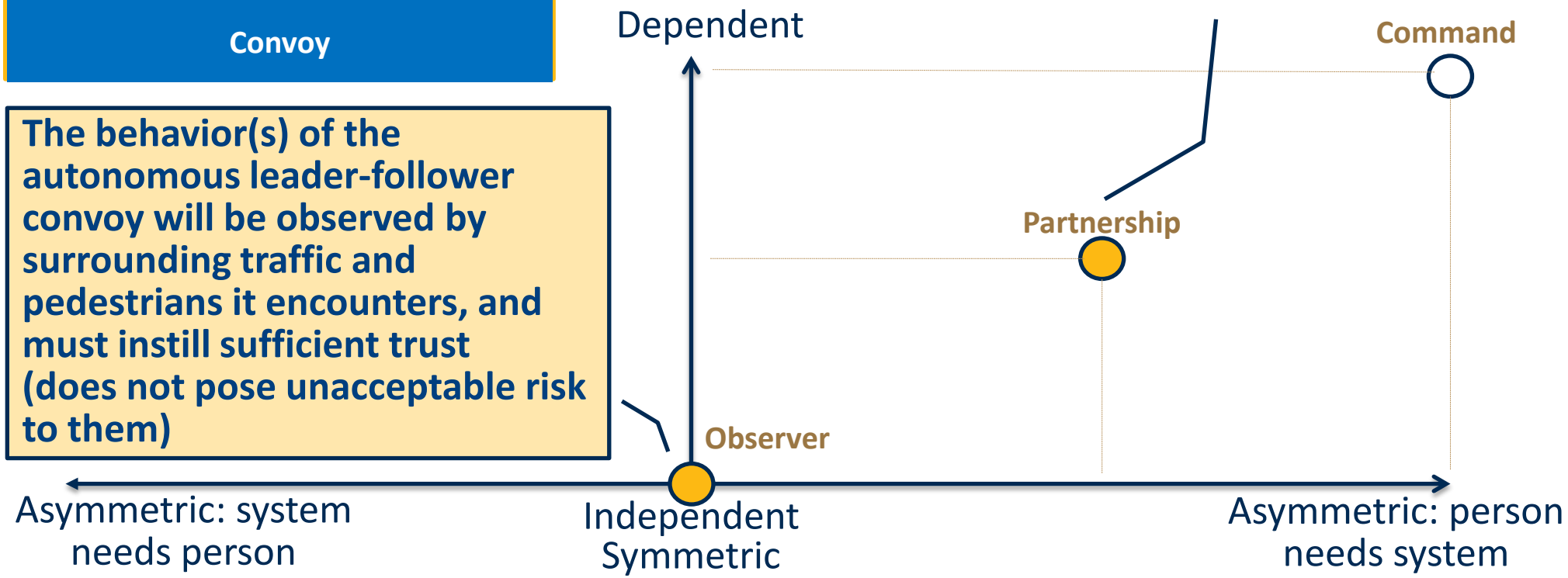
Georgia Tech Research Institute

# Example 3

Convoy

**The crew of the autonomous leader-follower is closely coupled to the performance of the autonomous follower vehicles to execute the mission.**

**The behavior(s) of the autonomous leader-follower convoy will be observed by surrounding traffic and pedestrians it encounters, and must instill sufficient trust (does not pose unacceptable risk to them)**

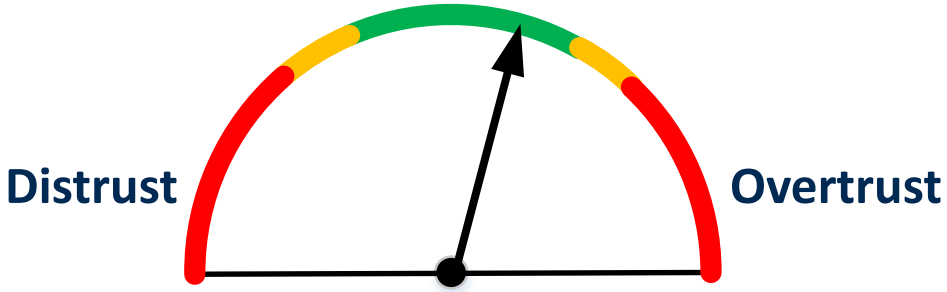


# Key Takeaways

- **Quantifying and clearly identifying human-autonomy relationships will aid T&E planning**
  - **'Peer', 'Partnership', and 'Command' used here as descriptors to convey differences***
  - **Recognizing and identifying when multiple human relationships exist (dependent upon CONOPS) will aid T&E planning**

* Other 'Partnership' descriptive terms: Human-Machine Teaming, Manned-Unmanned Teaming (MUM-T), …

# Trust Calibration




- **What is Trust Calibration?**
  - **Trust calibration refers to the continuum of humans effective utilization of autonomy**
  - **Trust calibration is critically important with broad ramifications for both disuse (Distrust) and misuse (Overtrust)**
  - **Distrust / Overtrust results**
    - **wrong decisions leading to undesired outcomes**
    - **reduced task effectiveness**
    - **…**

Georgia Tech Research Institute

# Key Takeaways

- **'Trust Calibration' is an established vernacular in the study of human/autonomy and can be utilized by the T&E community**
  - **Evaluation of both "Overtrust" and "Undertrust" must be considered as part of T&E planning**
  - **Predisposition of the human subject is a critical consideration in the evaluation of trust**
  - **Experiment structure will be an important facet of evaluating trust during T&E**

# Key Challenge #1

- How do you evaluate trust?
  - Experiment designs?
  - Metrics, measures, observations?
  - Pass Fail Criteria/Measures of effectiveness?

- Future T&E community needs
  - Methodologies and infrastructure to expand developmental test methods and approaches to address trust requirements
  - Expanding skillset and training policies to recognize the necessity of skills and evaluation methods of other disciplines (human factors, psychology, etc.)

# Key Challenge #2

- **Who must solve the T&E of trust problem?**
  - **Developmental Test?**
  - **Operational Test?**
  - **Training?**
  - **All the above?**

- **Future T&E community needs**
  - **An approach (supported by policy) that addresses a division of labor between DT&E, OT&E, and legacy training to effectively address the importance of, and challenges to, evaluating trust in autonomous systems**

# Conclusion

Georgia Tech Research Institute

## Evaluating trust will be a critical element of the T&E of autonomous systems

- Quantifying and clearly identifying human-autonomy relationships will aid T&E planning
- 'Trust Calibration' is an established vernacular in the study of human/autonomy and can be utilized by the T&E community
- Solutions to how you evaluate trust, and who evaluates trust (DT&E, OT&E, Training) may drive new T&E methodology, policy, and infrastructure needs

# Questions/Comments?

- **Contact Information**
  - **donald.strausberger@gtri.gatech.edu**
  - **pete.crump@gtri.gatech.edu**