# HAVE R2D2

23 Sept 2022

# “CALIBRATING” “TRUST” IN “AUTONOMY”

DISTRIBUTION STATEMENT A - Approved for public release; direction is unlimited. 412TW-PA-22288

# Team Members

- **TPS 21B Team**
  - EP 1: Maj. Tim “Scram” Sattler – Test Pilot
  - EP 2: Maj. Brian “OB” Herring – Project Manager / Test Pilot
  - EP 3: Maj. Darin “Shakes” Flynn – Test Pilot
  - EP 4: Capt. Charlie “Rolex” Sackett – RPA Test Pilot
  - EP 5: Maj. Leila “Roulette” Gerencser – Test WSO
  - EP 6: Joey “TOTE” Brown (Civ) – Flight Test Engineer

| EP Number | Rating | Primary Aircraft | Total Hours | Sorties Flown as EP |
|---|---|---|---|---|
| 1 | Pilot | F-22 | 1300 | 2 |
| 2 | Pilot | F-16 | 1600 | 2 |
| 3 | Pilot | B-52 | 2500 | 2 |
| 4 | RPA Pilot | MQ-9 | 1200 | 2 |
| 5 | WSO | B-52 | 1500 | 1 |
| 6 | FTE | F-16, T-38, C-12 | 65 | 1 |

# OPL Staff

- Thomas “MACH” Schnell – Target Pilot
- Lt Col Patrick “DICE” Highland – Meat Servo / Test Pilot
- Colton “Colt” Thompson
- Zeke “CHIP” Gunnink
- Jon “Hitch” Stoltz
- Brad “Roto” Parker
- Matt “Beaker” Cover
- Nate “Bluto” Smith
- Carl “Bender” Richey
- Isaac Eigner-Bybee

# Overview

- Background
- Test Objectives
- Test Setup
- Definitions
- Test Item Description
- Test Results (the cool videos)
- Lessons Learned

# Background

- **Customer – DARPA**
  - Measuring Operational Trust in Tactical Autonomy
    - Measure of how an agent will *aid* in achieving the *pilots goals* in a situation of *uncertainty* and *vulnerability*

- **Operator Performance Lab - 1v1 Engagements (L-29 Delfin)**
  - Test Aircraft: OPL Safety Pilot in FCP / TPS Student Evaluator pilot in RCP
    - SP follows AI steering commanded flight director (simulates AI controlled aircraft)
    - EP time shares between battle management task and monitoring AI agent
  - Artificial Intelligence Agent
    - Developed by EpiSci in the summer of 2020
    - Successfully completed the DARPA Alpha Dogfight Trials
  - Target Aircraft: OPL Safety Pilot / TPS Student
    - Varied levels of aggressiveness (Winning / Equal / Losing)

# Test Objectives

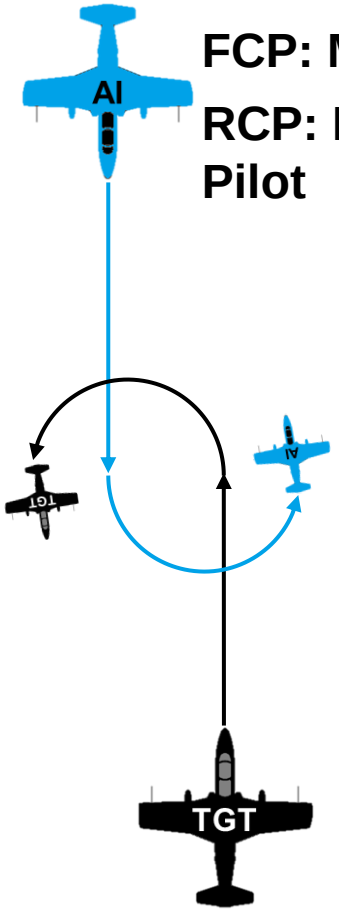
**Demonstrate a Framework for Measuring and Displaying Appropriate Trust with Respect to Own Ship Autonomy**

1. Create Mission Priorities
2. Create a Risk Model
3. Distract the Operator to Measure Trust
4. Create a “Trust-o-Meter”

**Determine Effectiveness**

# Test Setup

## Two-Circle Fights



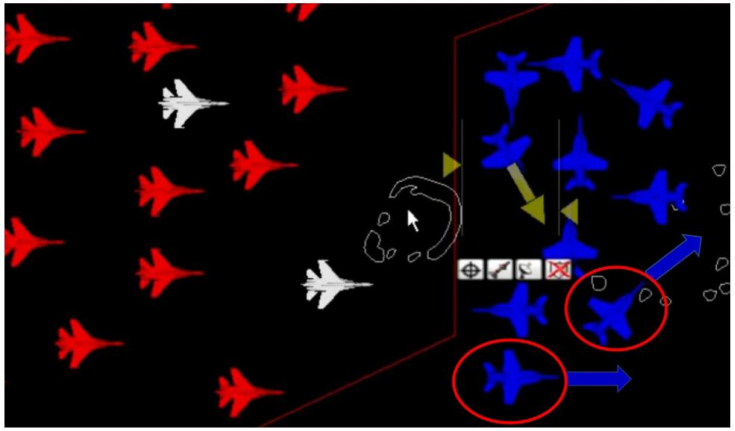

**FCP: Meat Servo**

**RCP: Evaluator Pilot**

**FCP:** AI Steering Commands

**RCP:**
HMD (Trust-o-Meter)

Fancy video game
(Battle Management)



# Test Setup

Helmet Mounted Display

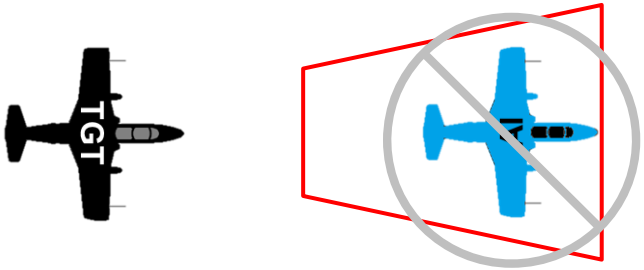
Biometric monitoring

Parachute & Harness

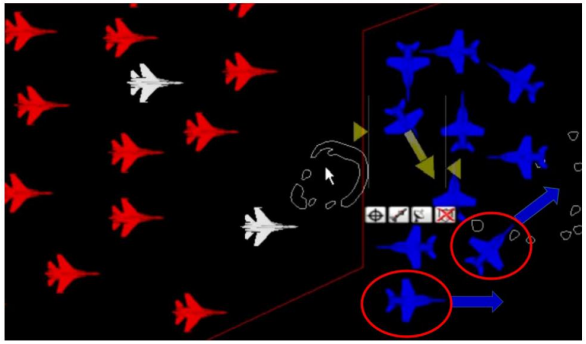
G-Suit

# Mission Priorities

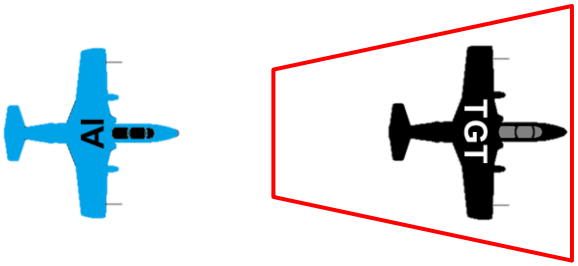
1. Survive




2. Battle Manage

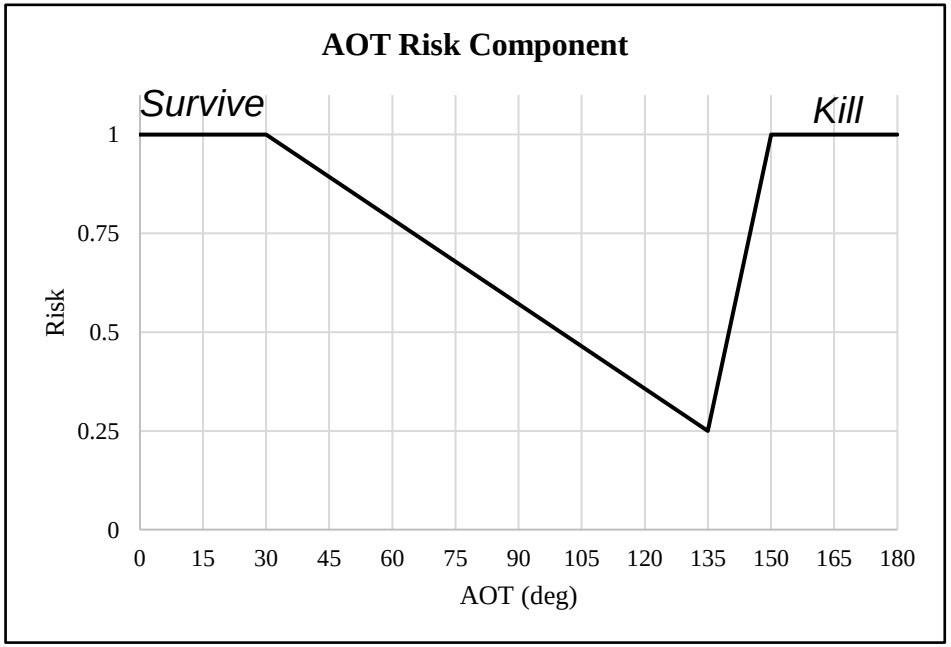


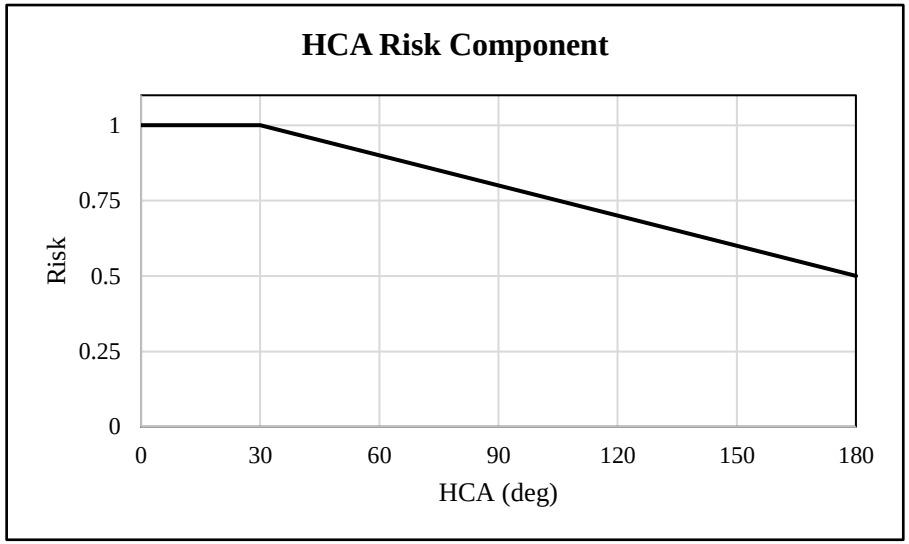
3. Kill

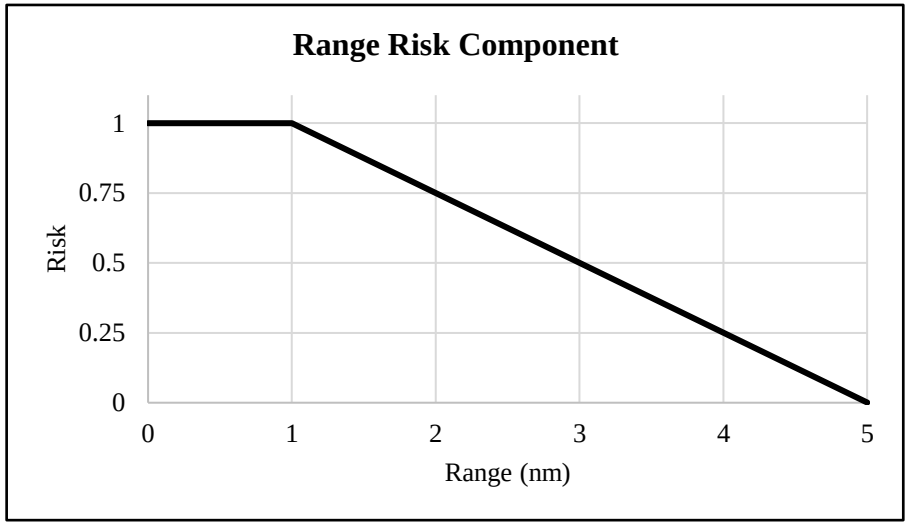


# Risk to Mission Model

- **RTM² = HCA x AOT x Range**
  - Always from test to target
  - Will be a value between 0 and 1



# Trust-O-Meter

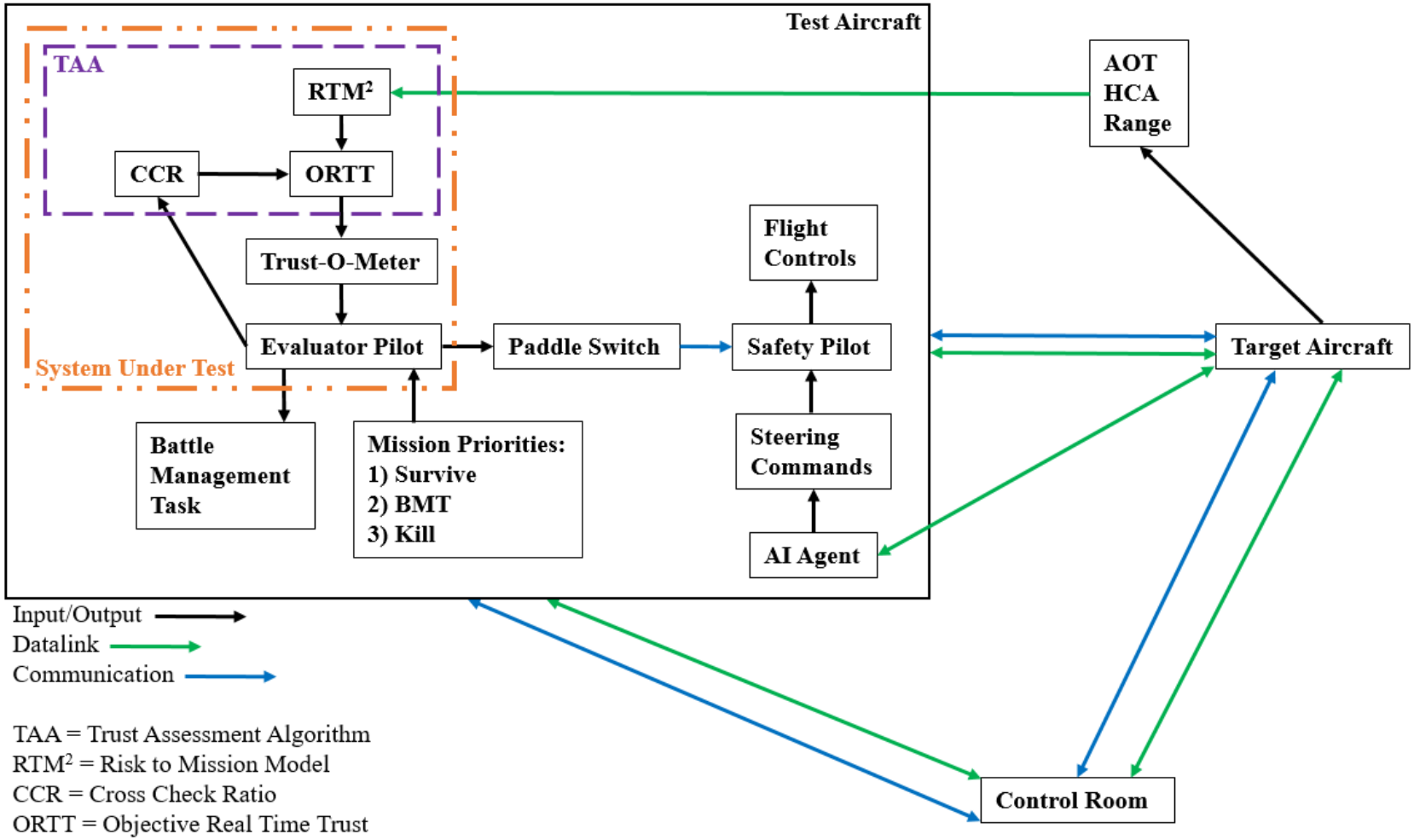
- **Displays output of TAA**
  - TAA output equals ORTT (reminder, ORTT = CCR + $RTM^2$)
  - Value between 0 and 2
  - T-O-M displayed on Operationally Flight Representative Helmet (OFRH)
    - OFRH *based on* F-35 helmet

| ORTT Score | TAA Computed Level of Trust | Trust-O-Meter Symbology Displayed in OFRH |
|---|---|---|
| 0.0 - 0.5 | Under Trust<br>(stop monitoring AI and prioritize battle management task) | ○ ▼ |
| 0.5 - 1.0 | Appropriate Trust<br>(continue battle management task) | ● |
| 1.0 - 1.5 | Appropriate Trust<br>(continue monitoring AI) | |
| 1.5 - 2.0 | Over Trust<br>(stop battle management task and monitor AI) | ▲ ○ |

# Test Item Description

- **System Under Test Terms**
  - Risk to Mission Model ($RTM^2$) (Survive / Battle Manage / Kill)
    - Based on (test to target) range, angle off tail, and heading crossing angle
    - How much risk is the target presenting to the mission
  - Cross Check Ratio (CCR)
    - Moving average of where the EP is looking (Helmet Tracker)
    - Score of 0 = Looking outside // Score of 1 = Looking inside
    - Can be any value between 0 and 1
  - Trust Assessment Algorithm (TAA)
    - Calculates a value that creates the ORTT
  - Objective Real Time Trust (ORTT)
    - ORTT = CCR + $RTM^2$
  - T-O-M visually depicts output of ORTT to EP via helmet symbology

# SUT Block Diagram

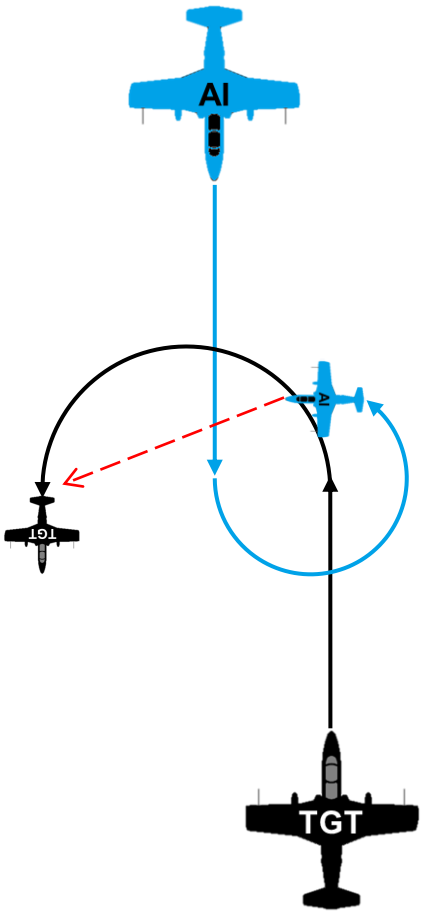



Input/Output
Datalink
Communication

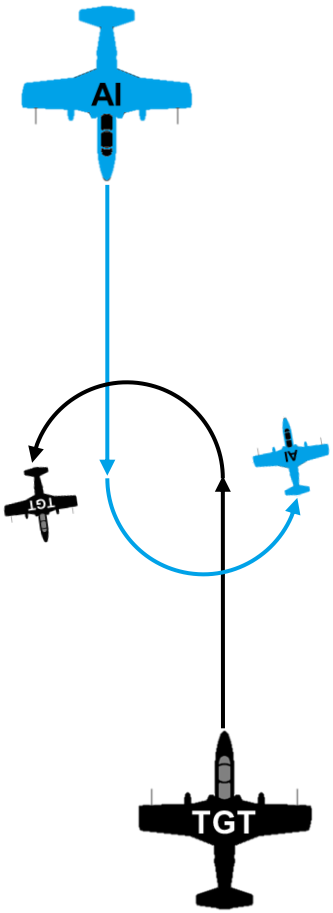
TAA = Trust Assessment Algorithm
RTM² = Risk to Mission Model
CCR = Cross Check Ratio
ORTT = Objective Real Time Trust
AOT = Angle-Off-Tail
HCA = Heading Crossing Angle
BMT = Battle Management Task

# Target Levels

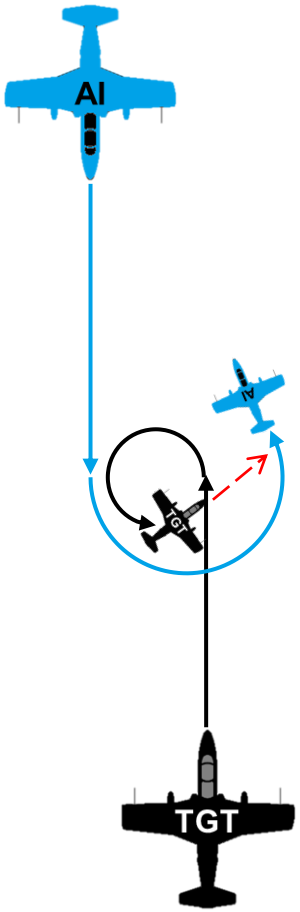
## LVL 1
***AI Win***




## LVL 2
***Neutral***

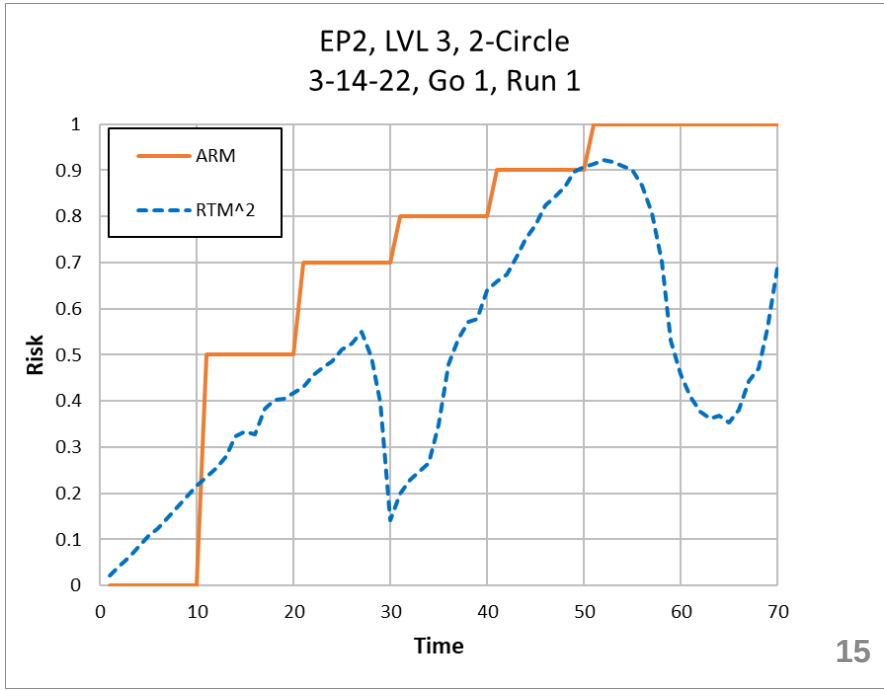



## LVL 3
***TGT Win***



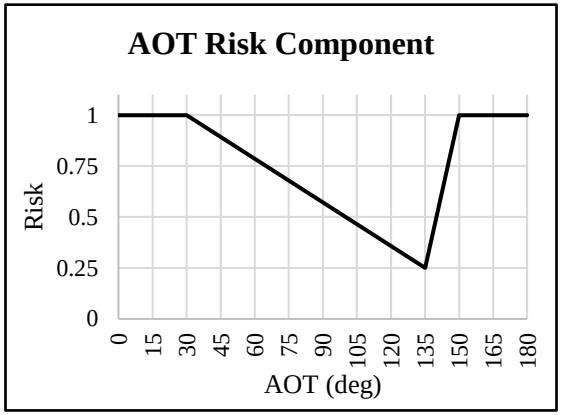
# ARM Assumptions
## (Actual Risk to Mission)

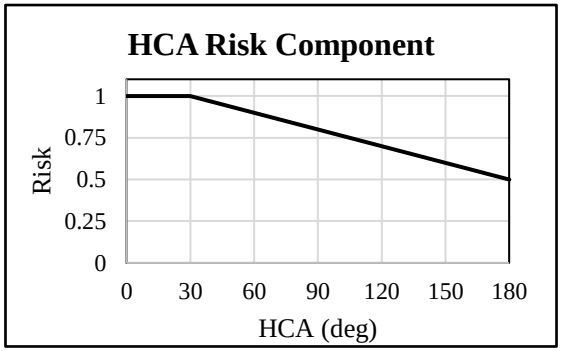
- **AI is directly controlling aircraft**
  - No safety pilot in front seat with survival instincts
- **No Training Rules**
  - No altitude deconfliction / 1,000’ bubble
- **No visibility issues for evaluator pilot in RCP**
  - Visual lookout assumed to be similar to modern fighter aircraft
- **Target and test aircraft have similar turn performance**
  - Same power setting set for both aircraft

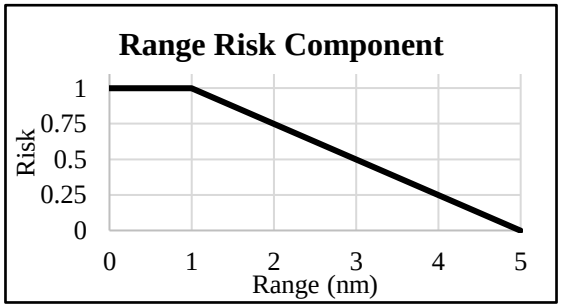


# Risk Model vs ARM

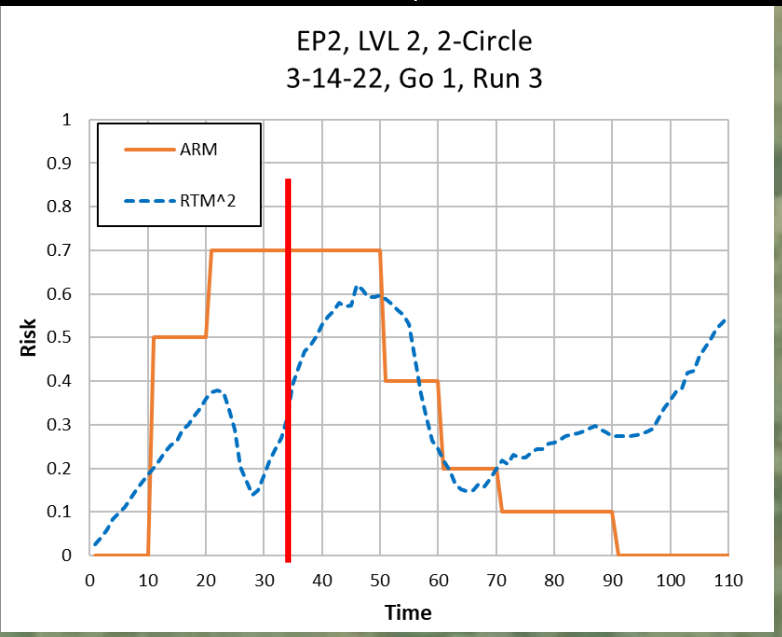
- **Intentionally only utilized three parameters into calculation of risk for simplicity**
  - Heading Crossing Angle (HCA)
  - Angle Off Tail (AOT)
  - Range
- **ARM parameters**
  - RTM$^2$ parameters plus:
    - Line of sight rate (LOSR)
    - Closure
    - Aspect Angle
    - Target vs Test aircraft energy states
    - Nose rotation
    - Pursuit Curve
    - Fight history / trends
    - Planform



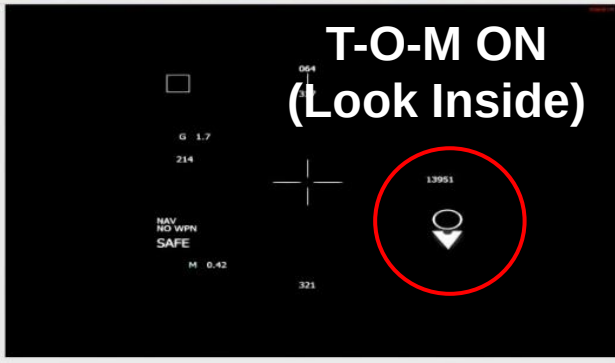



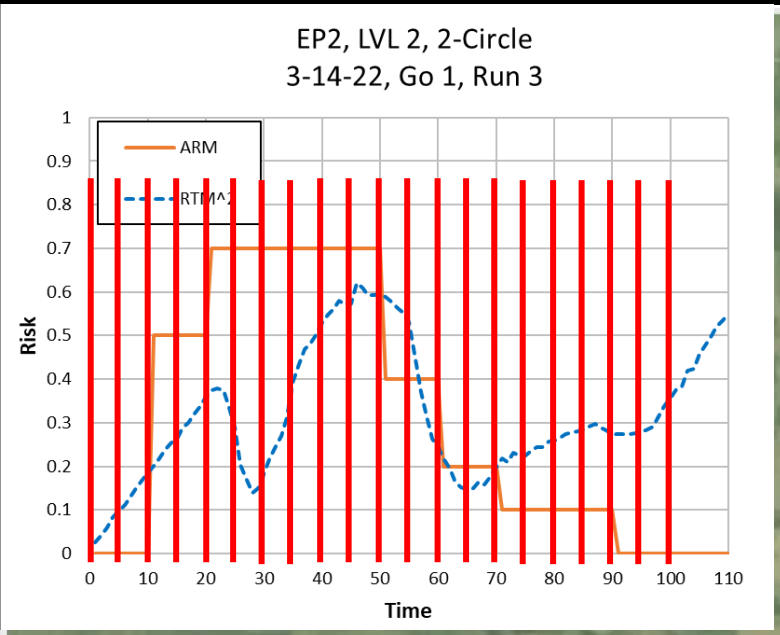
# LVL 2 TGT, T-O-M ON

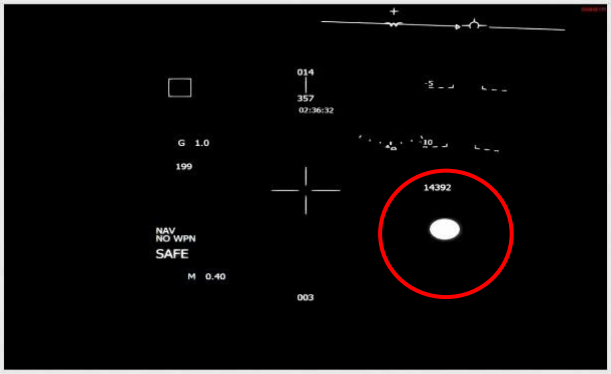


Battle Management Task



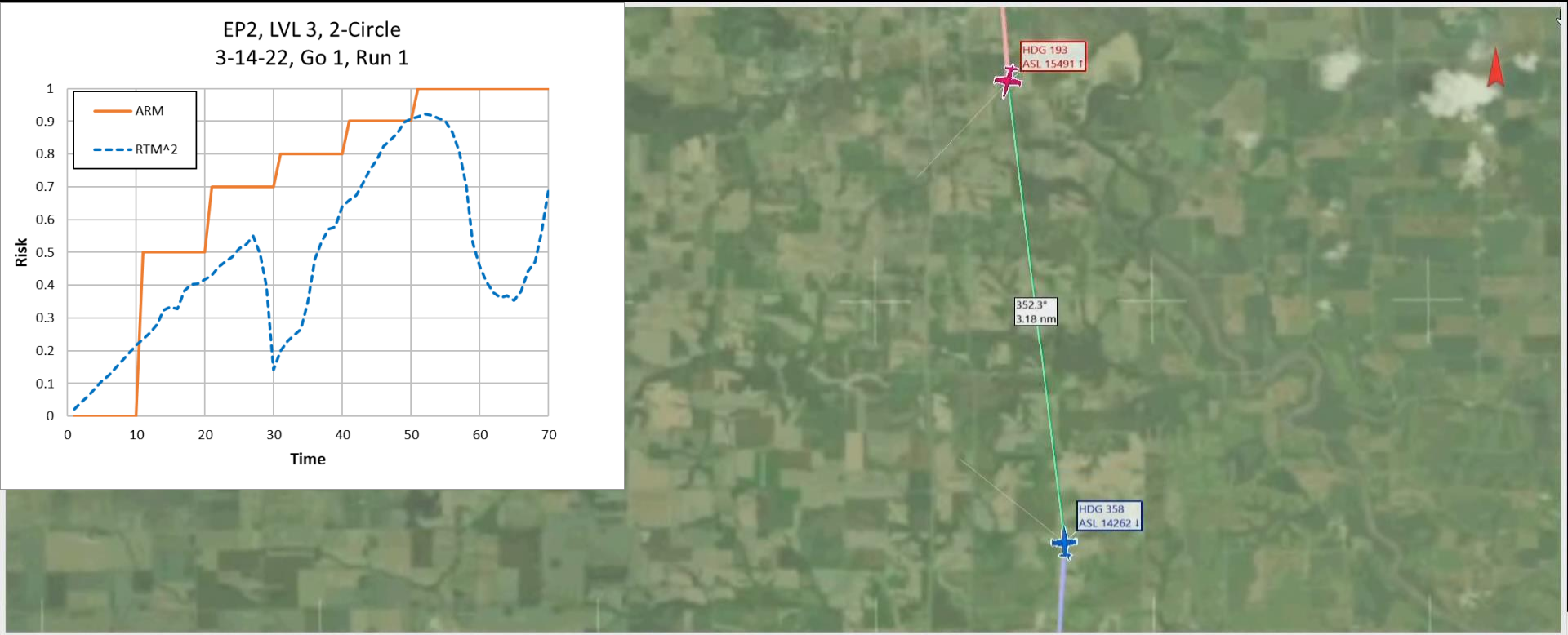
# LVL 2 TGT, T-O-M ON

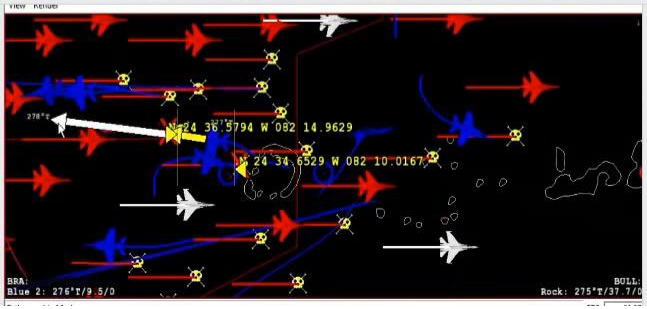


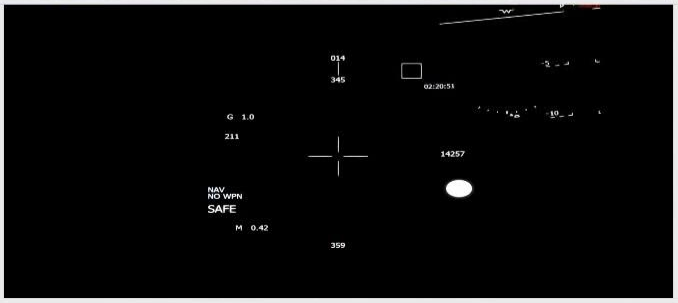


# LVL 3 TGT, T-O-M ON





# Results & Lessons Learned

## Demonstrate a Framework for Measuring and Displaying Appropriate Trust with Respect to Own Ship Autonomy

1. Create Mission Priorities
   - Tailored to Mission
2. Create a Risk Model
   - Must be complex enough to build trust
3. Distract the Operator to Measure Trust
   - Must be scorable based on effort
4. Create a “Trust-o-Meter”
   - Risk meter better?

# HAVE R2D2

**Recommendation: Field L-29 as AI dogfighter of the future**