Designing Safe Highly Automated Human-Machine Systems Using an STPA Approach: A Case Study

Presentation to the Tennessee Valey Chapter of ISSS

Dr Justin Poh, Dr Nancy G. Leveson
Massachusetts Institute of Technology (MIT)

Robert R. Copeland
US Army Combat Capabilities Development Command Aviation & Missile Center

22 July 2025

Massachusetts Institute of Technology

Engineering Systems Laboratory

ISAP
International Symposium on Aviation Psychology

About Me

2016 – 2020: Autonomous Driving Industry

- Hardware & Test Engineer, then systems engineer
- Developed physical and functional architectures for several versions of autonomous vehicle, accounting for safety and cybersecurity

2020 – 2025: MIT AeroAstro (Masters and PhD)

- Developed processes to enable safety-driven development of requirements and system architecture
- Applied approach to (1) pilot-automation architecture for a rotorcraft and (2) air traffic management architecture for urban air mobility
- Earned FAA private pilot license (PPL) in 2024

2025 - Present: Back in the autonomous driving industry as a software safety engineer

Agenda Overview

1. Challenges in designing highly automated human-machine systems
2. A new approach to architecting the role of human operators and automation/autonomy and the interactions between them
3. Brief demonstration of the approach applied to a case study
4. Conclusions & references for further information

Trend: Increasing Use of Software & Automation


[1]


[2]

[3]

Older Aircraft:

- Manually Flown
- Minimal Automation

Current & Future Aircraft:

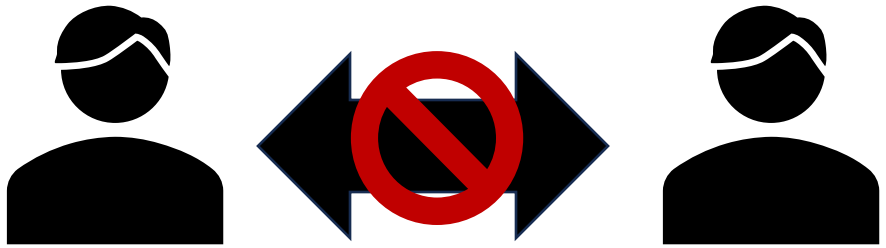
- Use of software-enabled autonomous functions
- Pilots work with (or supervise) automation

Increasing use of automation changes role of human pilot

Critical that aircraft design enables safe and effective human-automation interactions

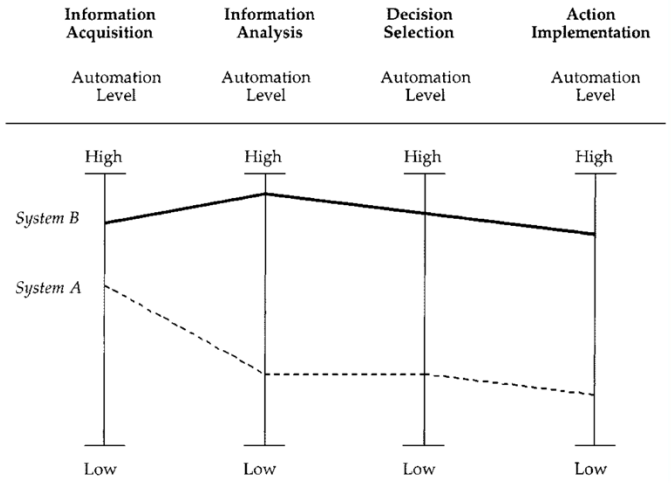
[1] Image from: https://piperowner.org/bold-warrior/
[2] Image from: https://commons.wikimedia.org/wiki/File:Boeing_787-8_N787BA_cockpit_%28cropped%29.jpg
[3] Image from: https://www.helicoptersmagazine.com/wisk-expands-evtol-plans-in-los-angeles/

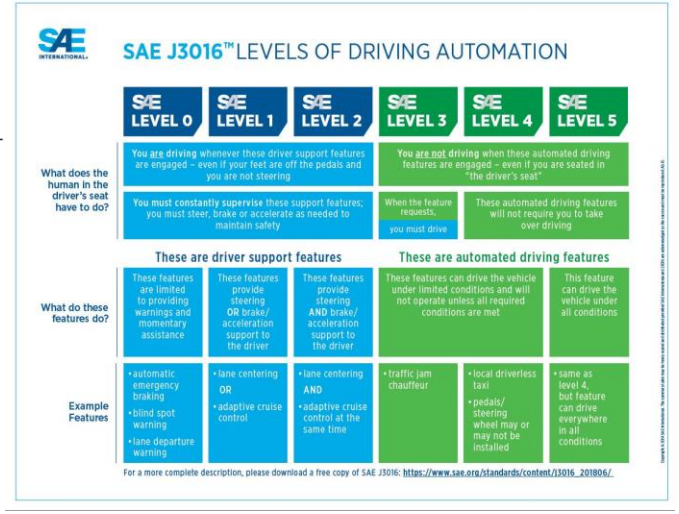
Challenges in Designing Human-Automation Interactions



Hardware & Software Engineers

Human Factors Engineers





Human factors engineering separated from hardware and software design

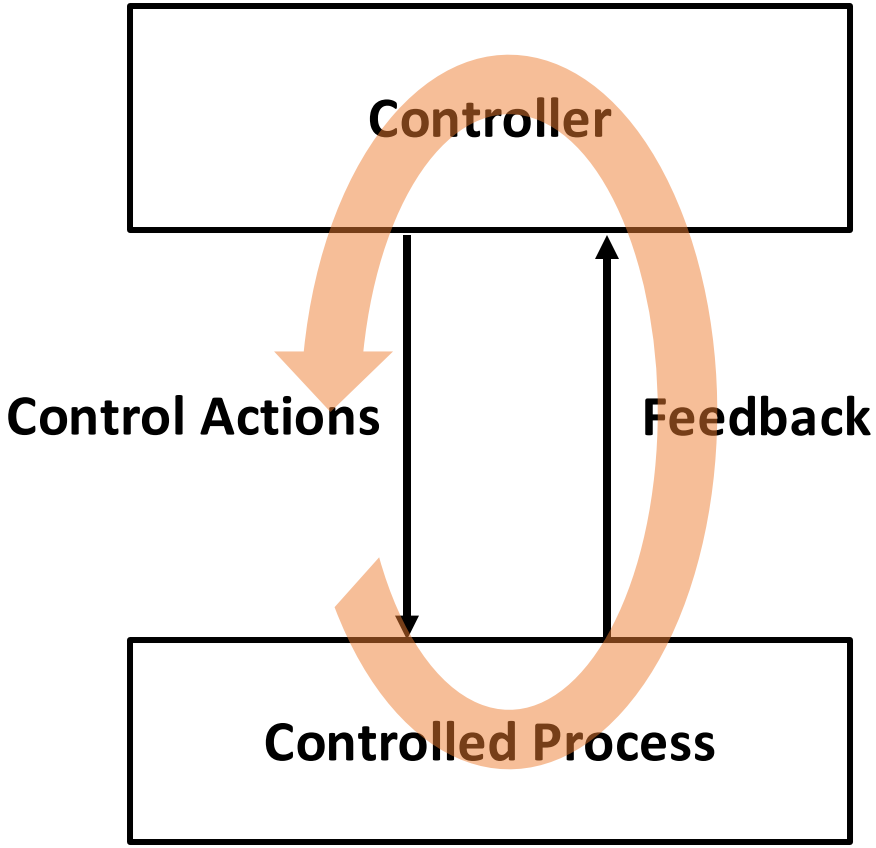
Levels of automation frameworks provide limited guidance on identifying required human-automation interactions

Research Objective: Demonstrate a new approach to system design that enables earlier and more integrated consideration of human factors and safety during design

[1] Image from: Parasuraman, Sheridan, and Wickens (2000)
[2] Image from: SAE International (2021)

Safety-Driven Design: A Control Problem

- Safety is an **Emergent Property** (arises from the interactions between system components)
- System design needs to include sufficient controls to **prevent unsafe behavior**
- System is modeled using a **Control Structure** containing:
  - Controlled Process & Controller
  - Control Actions & Feedback



Systems Theoretic Process Analysis* (STPA) analyzes the control loops in a system to identify how unsafe behavior could occur

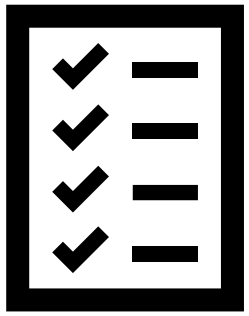
**See Leveson and Thomas (2018) for details*

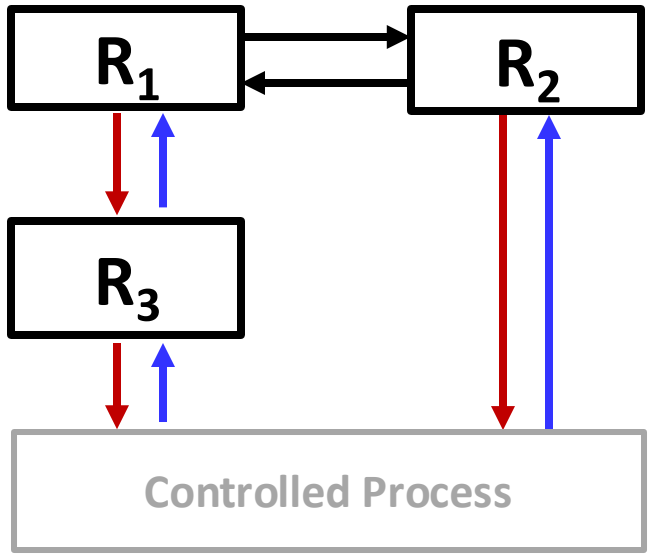
STPA Results Drive Safety-Informed Design Decisions

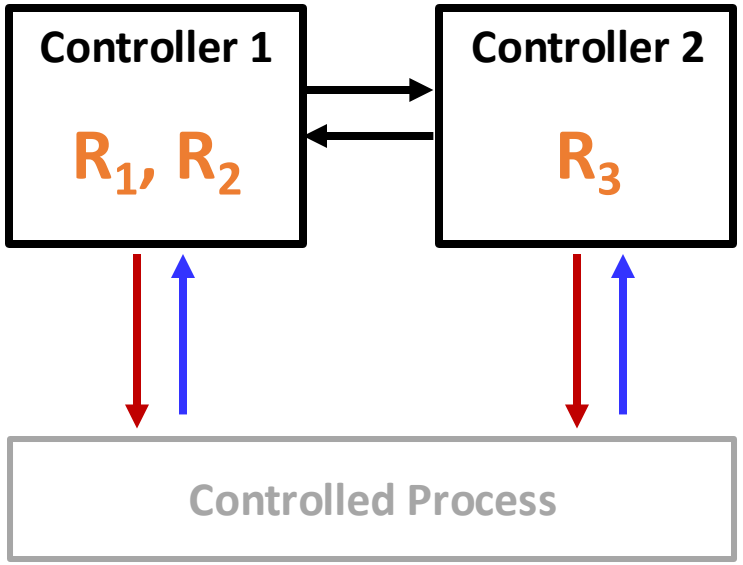
1. Identify System Requirements

2. Define Responsibilities (Functions) & Relationships (System-Level Behavior)

3. Allocate Responsibilities to System Elements (System Architecture)







STPA can help to make more informed <u>early</u> design decisions

Overview of Approach & Case Study

STPA* Analysis of System identifies how unsafe behavior could occur

↓

Define System Requirements to prevent unsafe behavior

↓

Define System-Level Behavior to meet requirements

↓

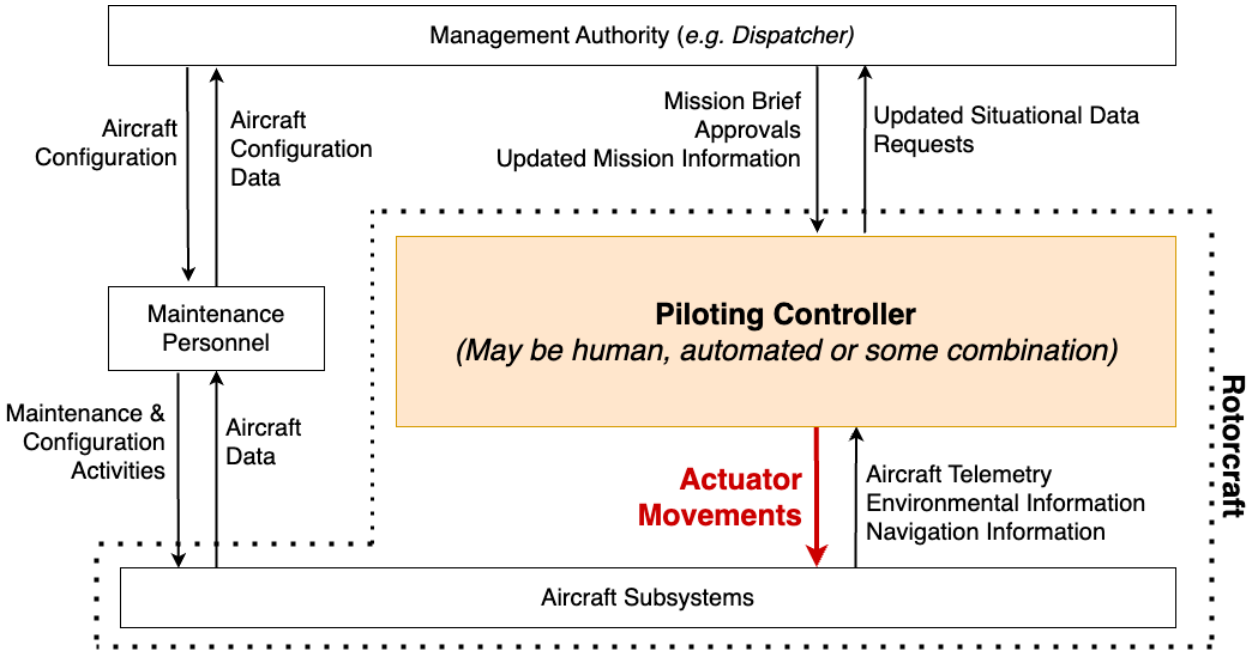
Create and Assess Architecture Options to implement system-level behavior

Case Study: Develop Pilot-Automation Architecture for a rotorcraft to be flown in Degraded Visual Environments (DVEs)

Result: Safety and human factors are considered upfront in the system architecture

[4] Image from:https://www.militaryaerospace.com/computers/article/16714693/sierra-nevada-demonstrates-helicopter-synthetic-vision-for-degraded-visibility-environments-dve

System Requirements Derived From STPA Analysis



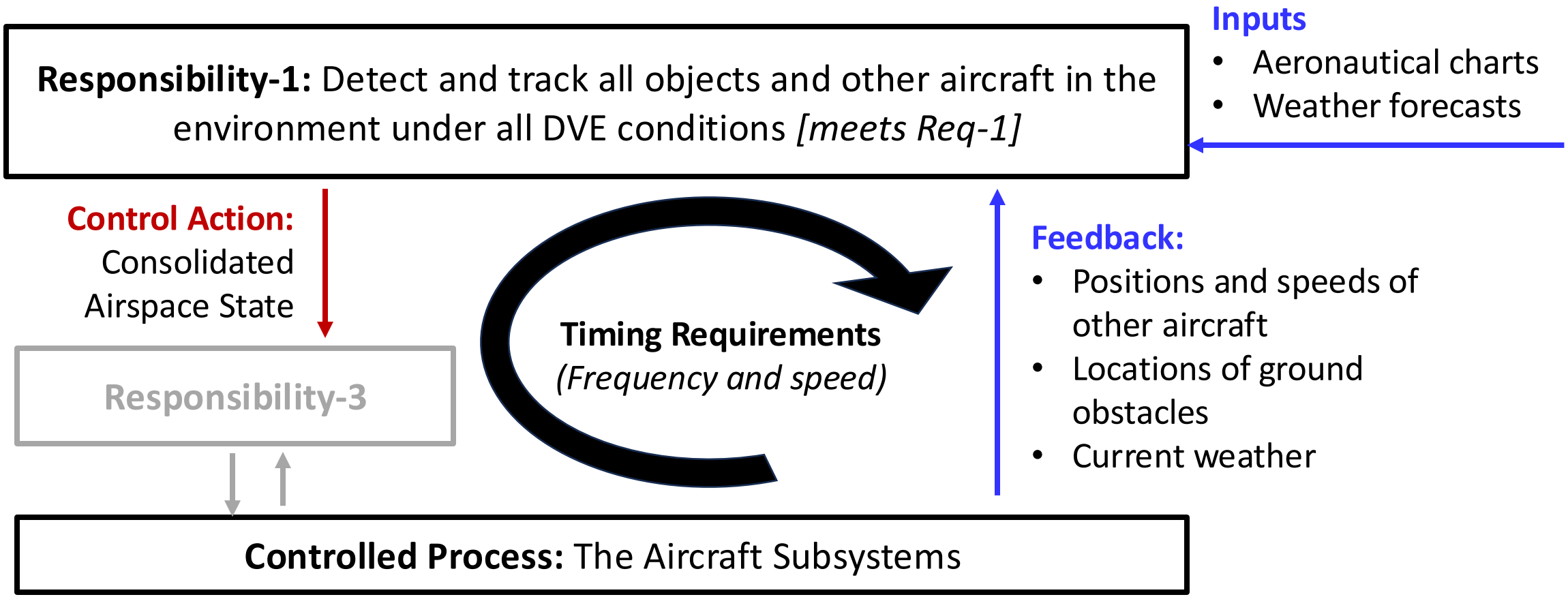
Example Unsafe Control Action: Piloting controller provides actuator movements that steer the aircraft toward another aircraft or object

Example Causal Scenario: Inaccurate sensor feedback wrongly indicates no aircraft or objects nearby. Piloting controller wrongly believes the airspace is clear and steers the aircraft toward the object or other aircraft.

Req-1: The aircraft system must be able to detect and track all objects and other aircraft in the environment under all DVE conditions.

STPA-derived system requirements account for safety and human factors considerations early in development

System-Level Behavior Defines Required Control Loops

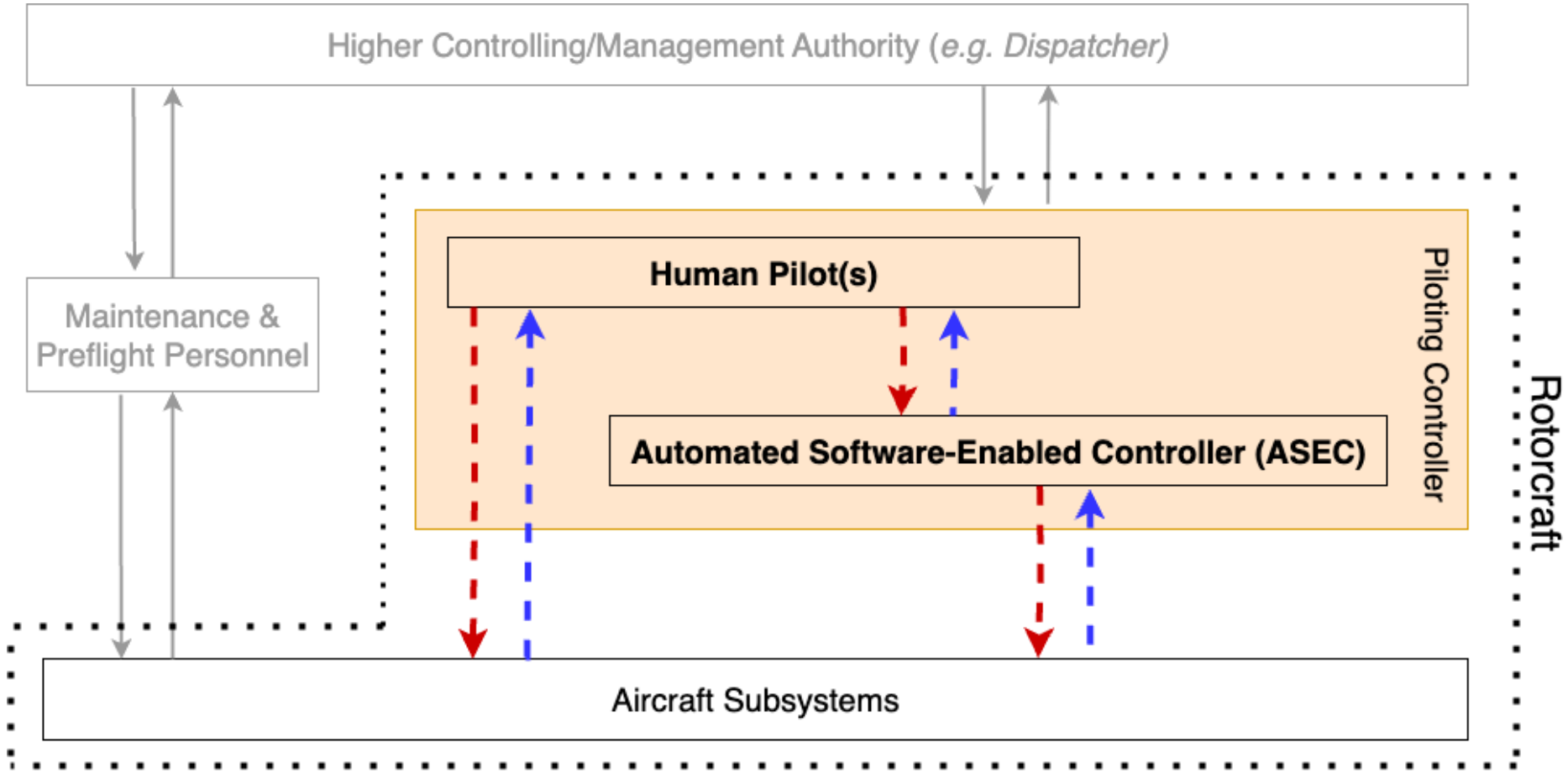


Control loops define required system elements

Can inform responsibility assignments to human operators

Creating Architecture Options By Assigning Responsibilities

Case Study Goal: Develop Pilot-Automation Architecture for a rotorcraft to be flown in Degraded Visual Environments (DVEs)



Architecture Creation By Assigning Responsibilities

Architecture options represent possible assignments of responsibilities

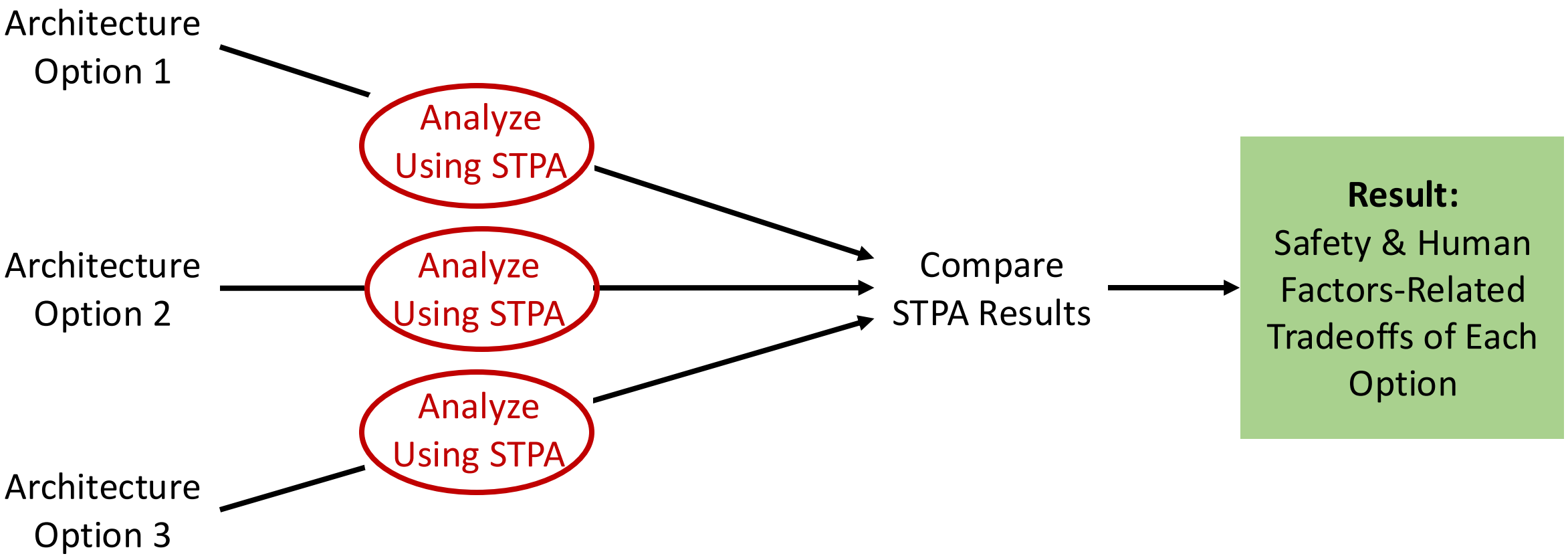
Responsibility ID	Option 1: Low Automation		Option 2: Medium Automation		Option 3: High Automation	
	Pilot	ASEC*	Pilot	ASEC*	Pilot	ASEC*
Resp-1: Detect and Track All Objects	•	•	•	•	•	•
Resp-2: Ensure collision-free flight path is available	•		•	•	•	•
Resp-3: Select Appropriate Flight Path	•		•	•	•	•
Resp-4: Provide control inputs quickly enough and with appropriate magnitude	•	•	•	•		•

Role of the human pilot in each architecture option is clearly defined

Easier to assess architecture's impact on human performance

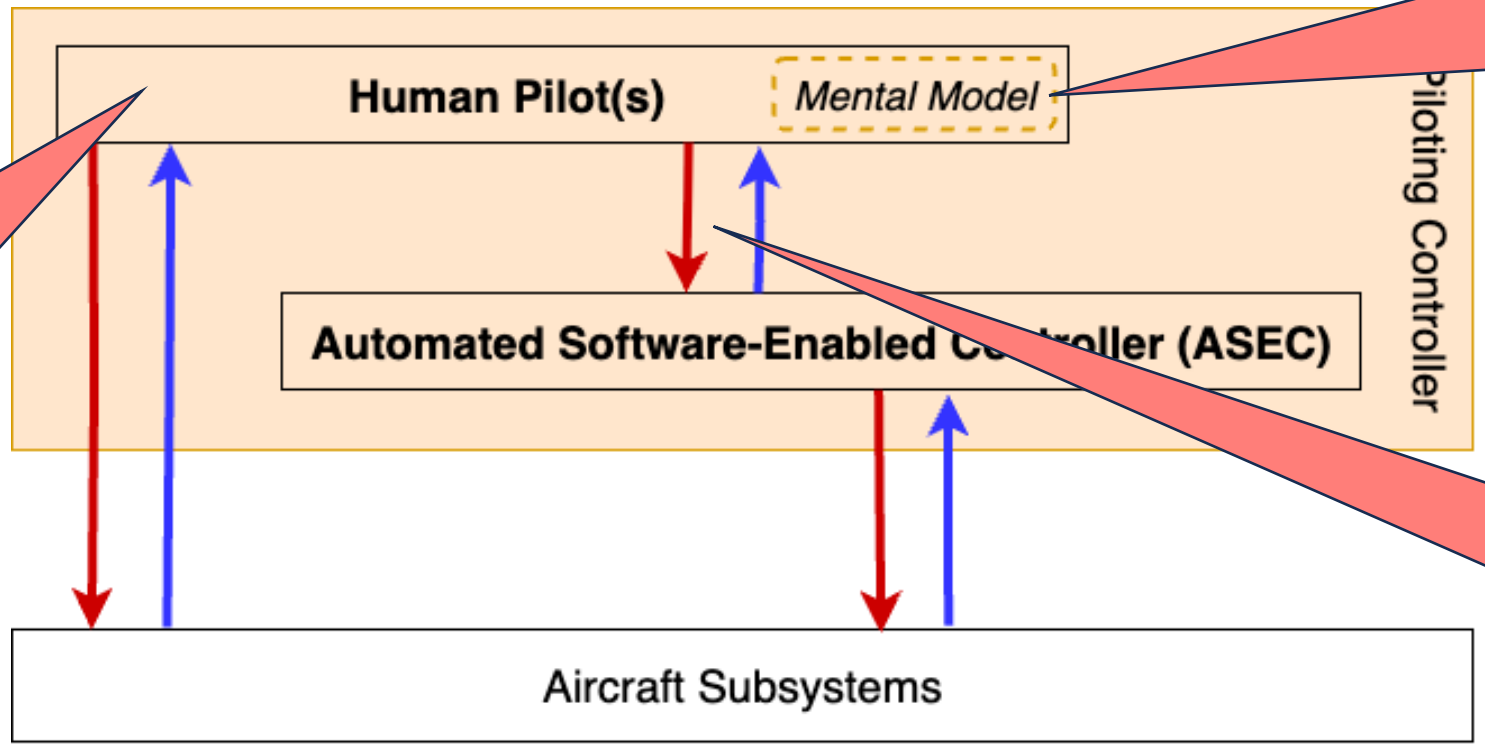
*ASEC: Automated Software-Enabled Controller

Comparing Architecture Options Using STPA



Comparison Result 1: Pilot Workload

Automation provides pilots with more assistance but expected reduction in pilot workload may not achieve



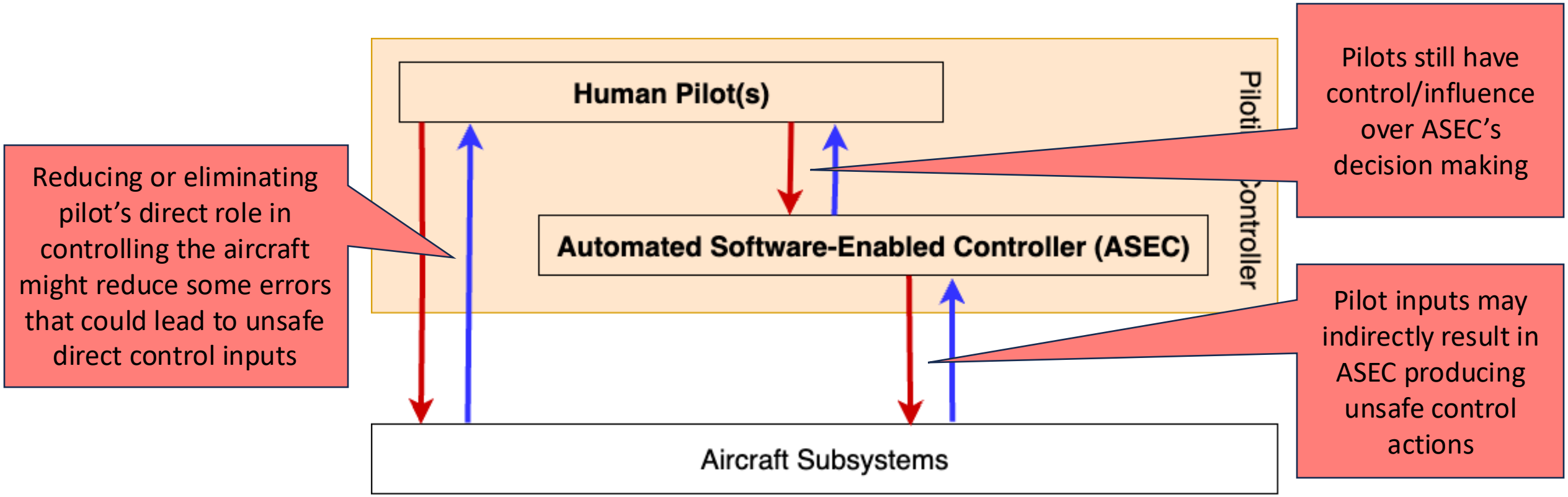
Difficult to reduce pilot's required situational awareness (requires large reduction in responsibilities)

Maintaining awareness of ASEC state and mental model of its behavior increases workload

Pilot's required interactions with automation can further increase workload

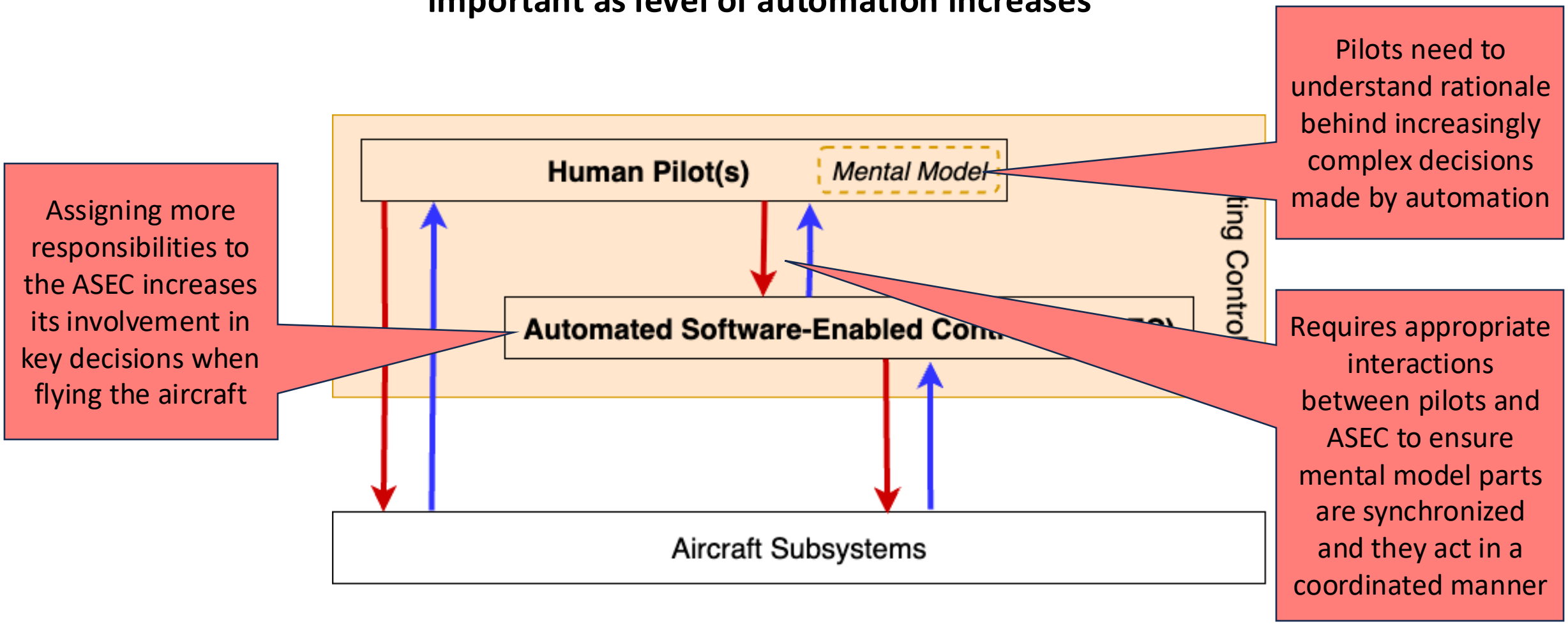
Comparison Result 2: Human Decision-Making Biases

System must always be designed to avoid unsafe human decision-making biases and heuristics, regardless of how much automation is employed



Comparison Result 3: Human-Automation Coordination

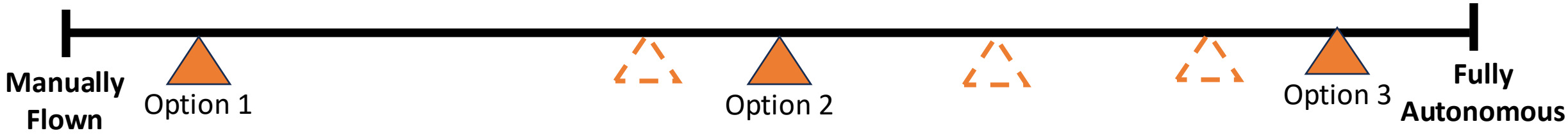
Good coordination between human pilot and automation is increasingly important as level of automation increases



Comparison Results Inform Preferred Architecture

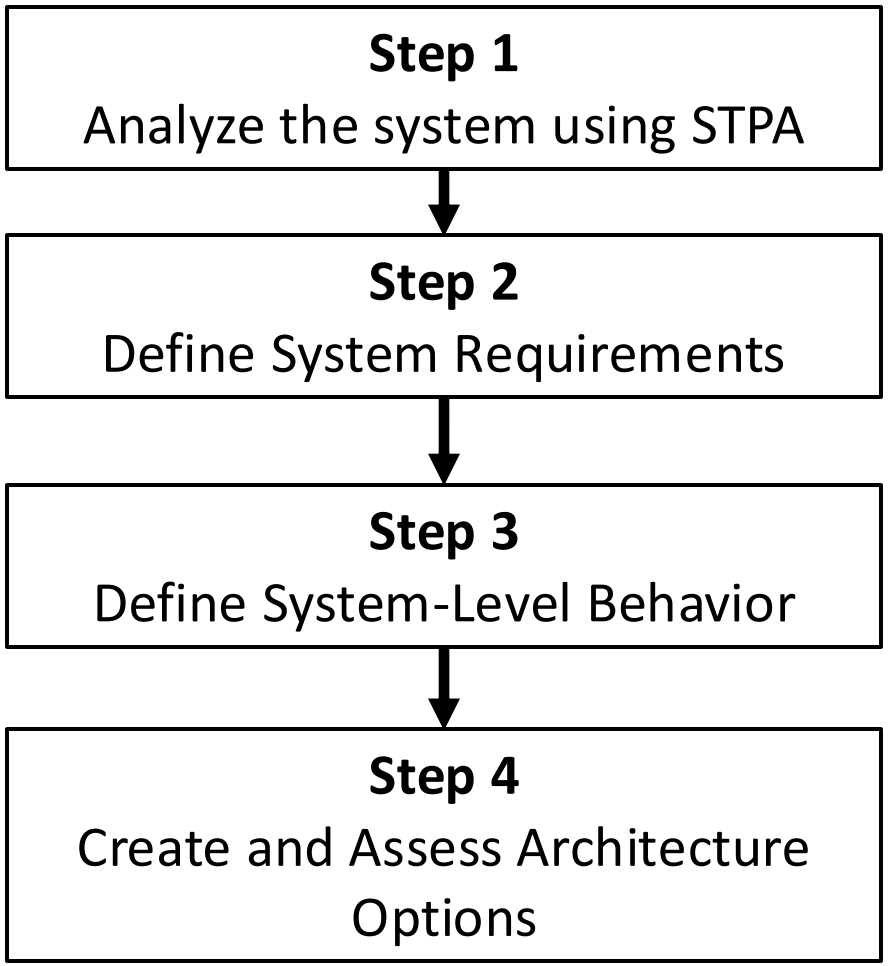
These safety and human factors tradeoffs can inform which responsibilities *should* be assigned to the human pilot or automation

Responsibility ID	Option 4	
	Pilot	ASEC*
Resp-1: Detect and Track All Objects		
Resp-2: Ensure collision-free flight path is available	?	
Resp-3: Select Appropriate Flight Path		
Resp-4: Provide control inputs quickly enough and with appropriate magnitude		



*ASEC: Automated Software-Enabled Controller

Summary



Research Objective: Demonstrate a new approach to system design that enables earlier and more integrated consideration of human factors and safety

- Using STPA ensures **system requirements account for human factors and safety considerations early in design**
- System-level behavior **identifies better information requirements to inform responsibility assignments**
- Comparing architecture options **highlights human factors-related benefits and tradeoffs**

Additional Related Publications

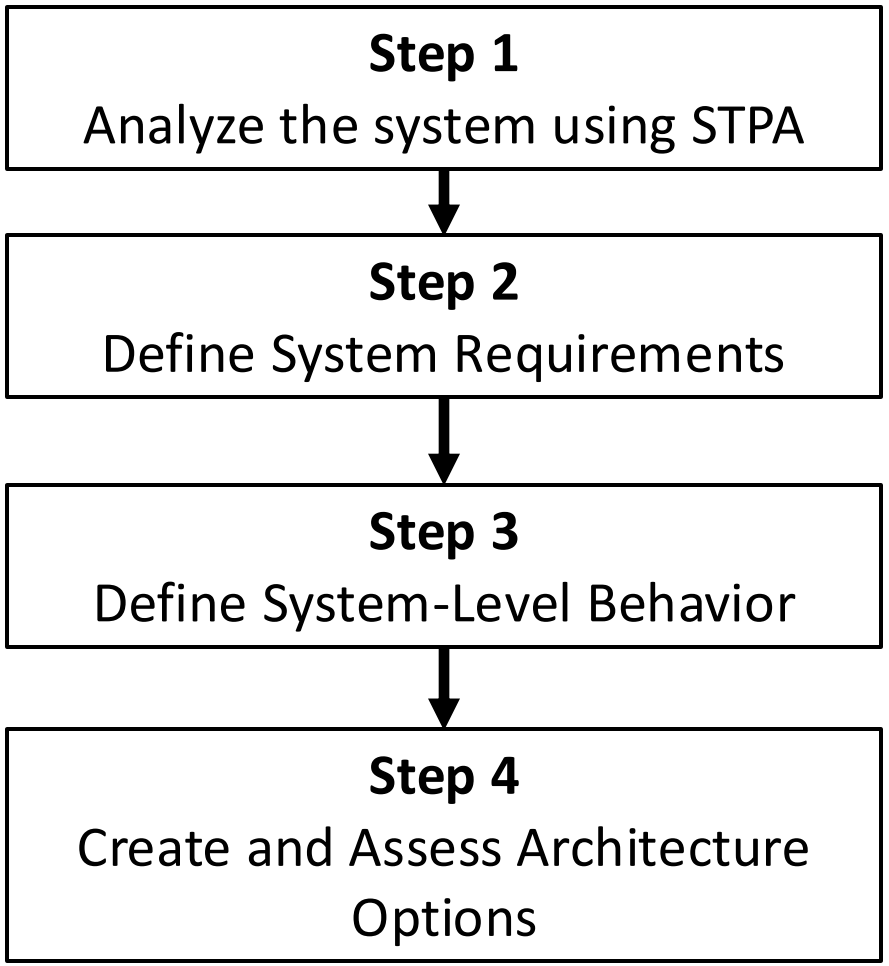
1. **More detailed information on this work** – J. Poh, “A Top-Down, Safety-Driven Approach to Architecture Development”, January 2022

2. **Application of this approach to air traffic control architecture** - J. Poh, N.G. Leveson, N.A. Neogi, "A Safety-Driven Approach to Exploring and Comparing Air Traffic Management Concepts for Enabling Urban Air Mobility", Proceedings of the International Conference on Research in Air Transportation (ICRAT), July 2024

3. **Refinement and extension of this approach** - J. Poh “A Systems-Theoretic Framework For Safety-Driven Development of System Architectures, December 2024

All publications available at https://www.justinpoh.com/publications--presentations.html

Summary



Research Objective: Demonstrate a new approach to system design that enables earlier and more integrated consideration of human factors and safety

- Using STPA ensures **system requirements account for human factors and safety considerations early in design**
- System-level behavior **identifies better information requirements to inform responsibility assignments**
- Comparing architecture options **highlights human factors-related benefits and tradeoffs**

Questions?

justin@justinpoh.com

References

- Copeland, R. (2019). *An Analysis and Classification Process towards the Qualification of Autonomous Systems in Army Aviation*. Vertical Flight Society's 75th Annual Forum & Technology Display.
- Copeland, R., Carter, H., & Mulder, J. (2024). *Information Management and Information Fusion: Human Engineering Approaches and Constructs (Revisited)* [US Army, DEVCOM AvMC, Redstone Arsenal, AL (Whitepaper)].
- Flanigen, P., Copeland, R., Sarter, N., & Atkins, E. (2022). Current challenges and mitigations for airborne detection of vertical obstacles. *Proceedings for International Society for Optics and Photonics (SPIE) Defense and Commercial Sensing*.
- Leveson, N. (2011). *Engineering a safer world: Systems thinking applied to safety*. MIT Press.
- Leveson, N., & Thomas, J. P. (2018, March). *STPA Handbook*. psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *30*(3), 286–297.
- Poh, J. (2022). *A Top-Down, Safety-Driven Approach to Architecture Development for Complex Systems* [Masters]. MIT.
- Proctor, R. W., & Van Zandt, T. (2018). *Human factors in simple and complex systems* (3rd ed.). CRC Press.
- SAE International. (2021). *J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*.
- Wickens, C. D. (Ed.). (2004). *An introduction to human factors engineering* (2nd ed.). Pearson/Prentice Hall.