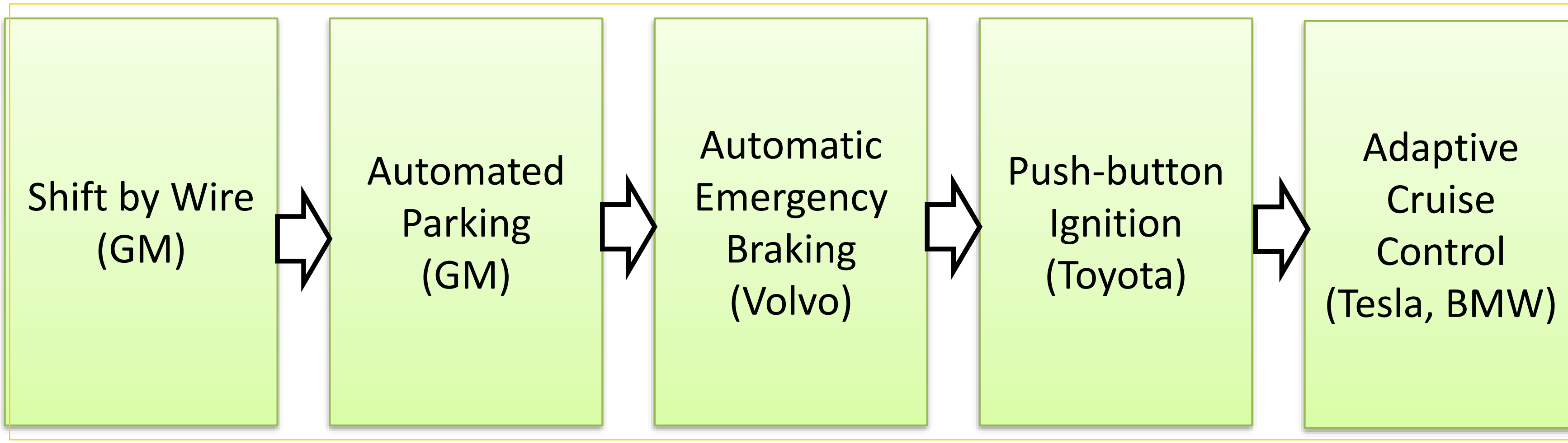
Massachusetts
Institute of
Technology

System-Theoretic Process Analysis (STPA): Engineering for Humans

Dr. John Thomas

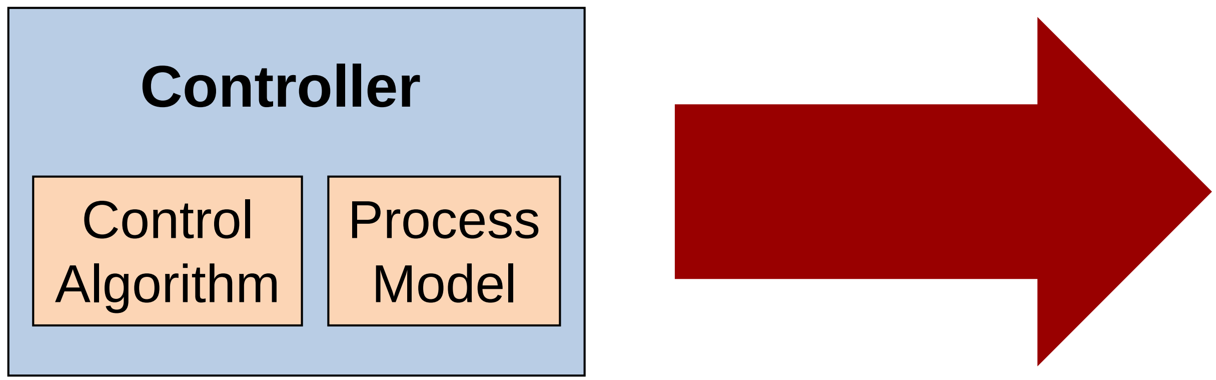
Any questions? Please email: jthomas4@mit.edu

Past Applications, Progression



Acknowledgements
Mark A. Vernacchia
Charles A. Green
Padma Sundaram
Joseph D’Ambrosio
Matt Boesch
Megan France
Jeramiah Robertson

Controller model



Existing systems-theoretic controller model

- Generic
- Not specific to humans

Leveson, 2012

HUMAN FACTORS MODELS

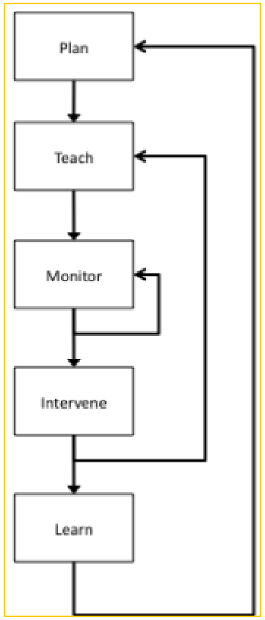


Figure 3. Sheridan's (1992) Supervisory Control Model.

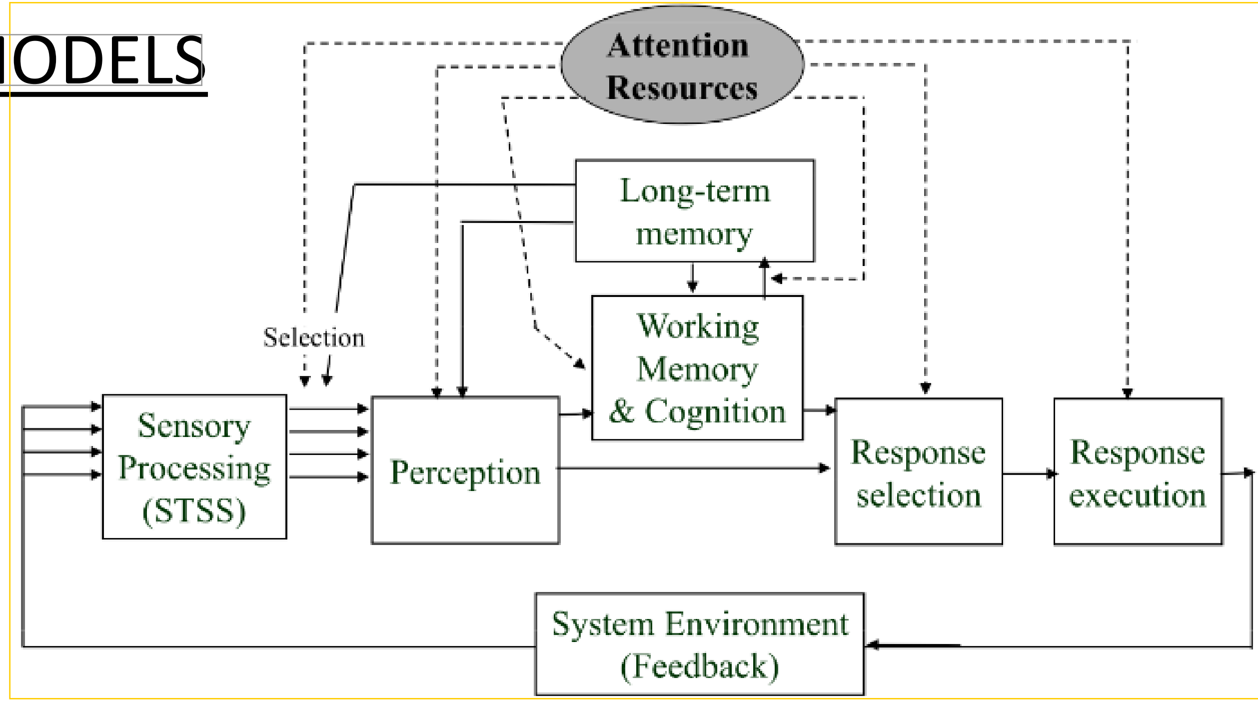


Figure 2. Modified from Wickens and Hollands (2000, p.11).

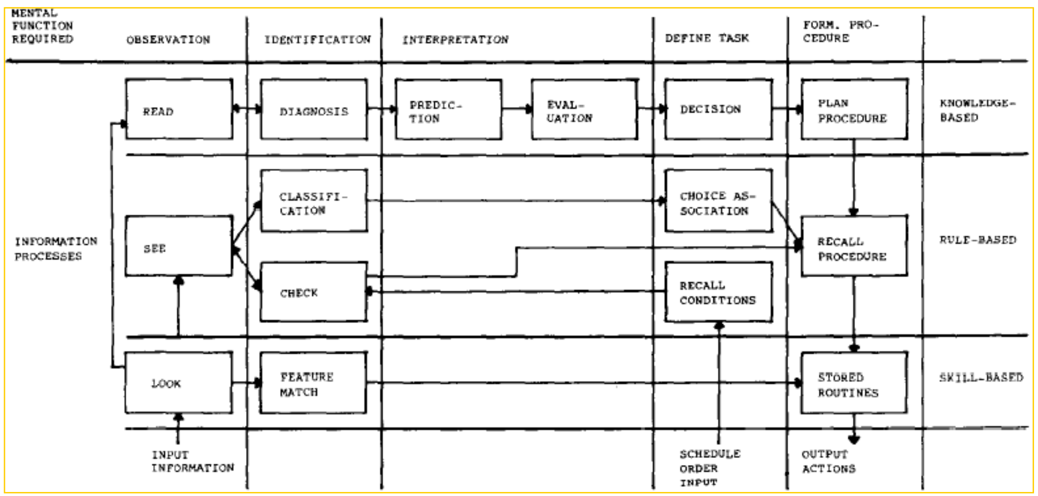


Fig. 3. The diagram illustrates how the same required mental function can be served by different information processes - each with particular error mechanisms.

164 *C. D. Wickens*

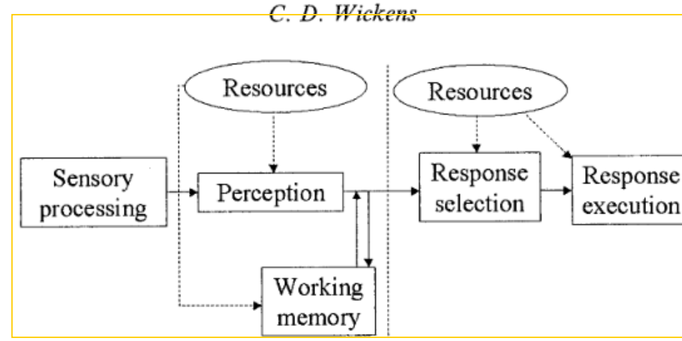
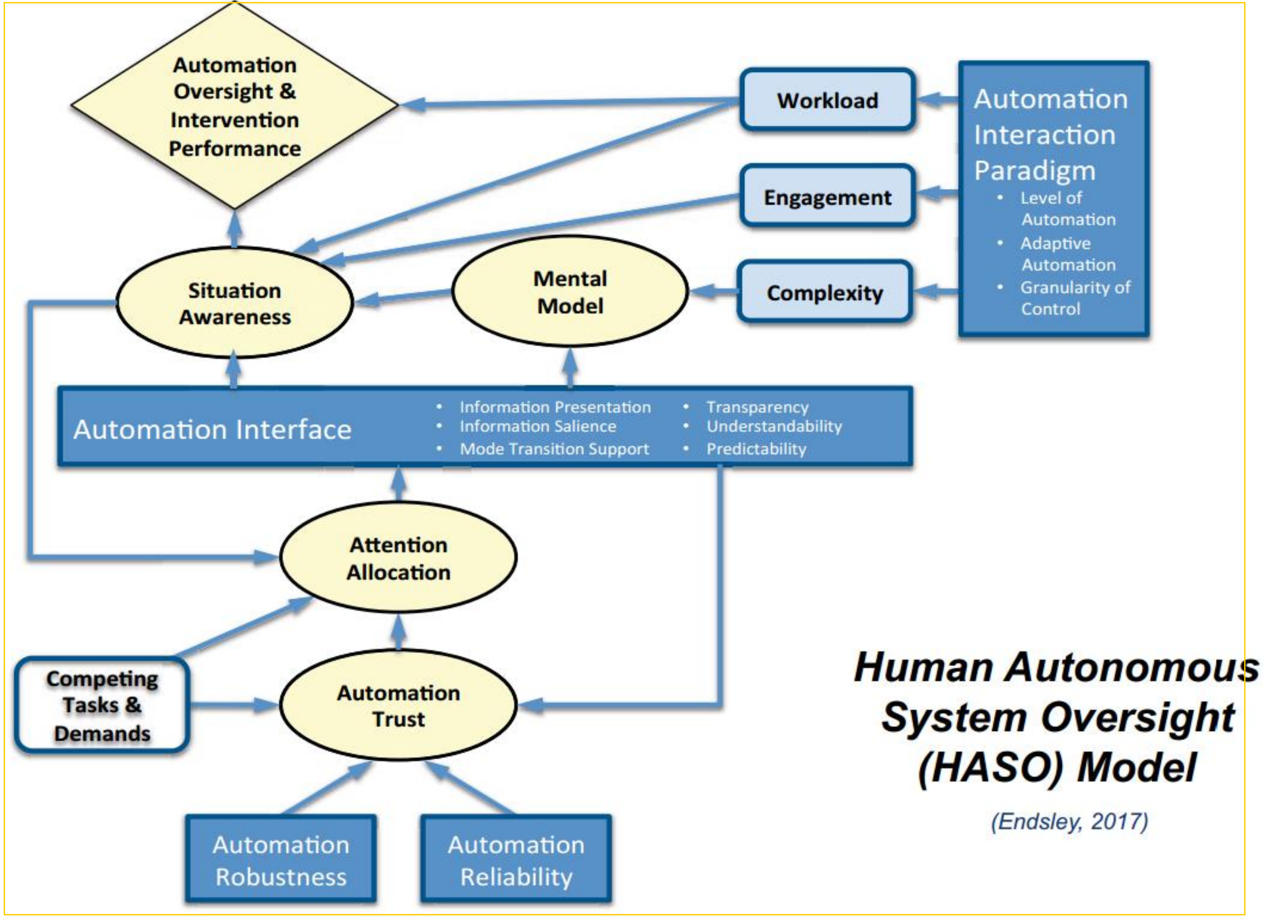
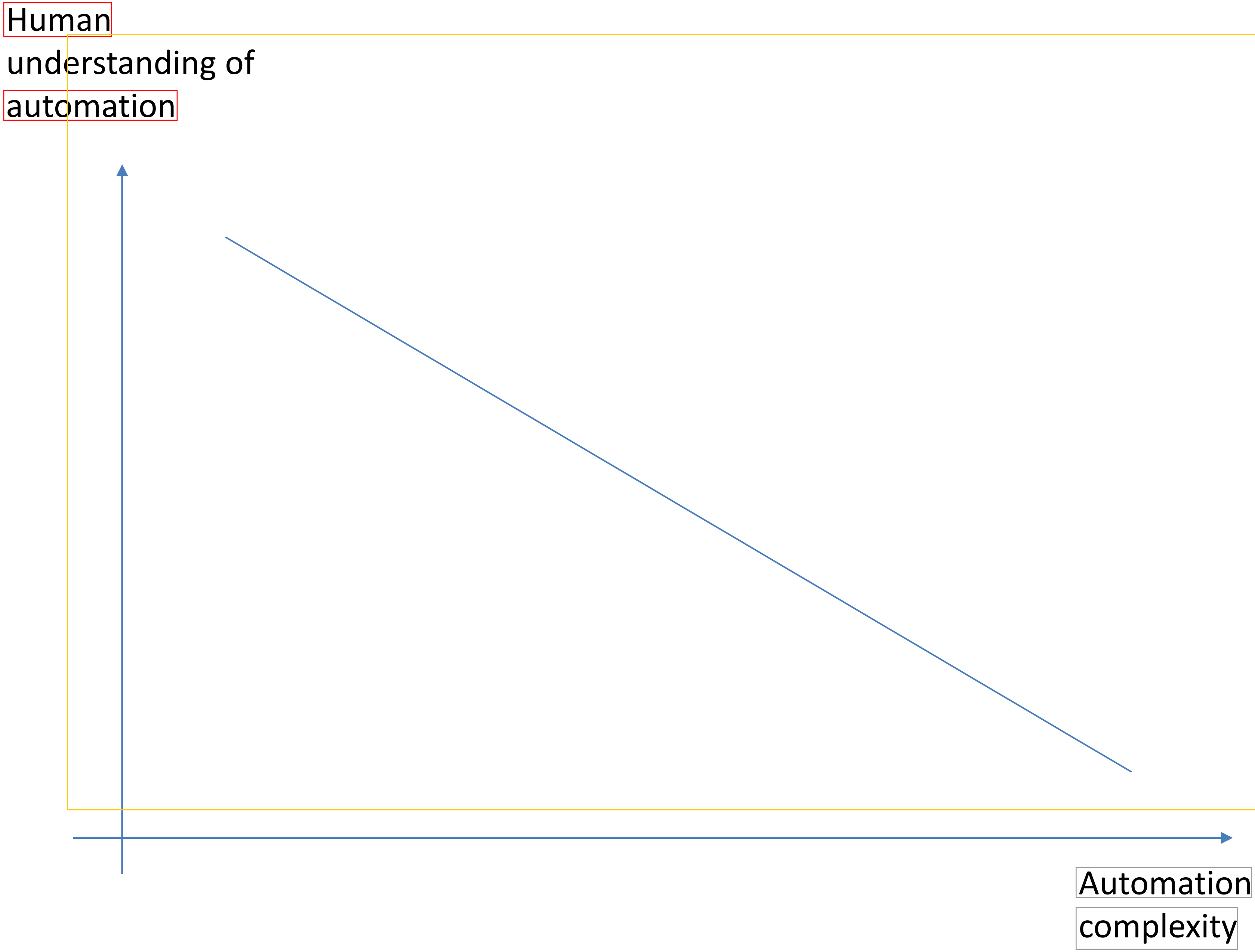


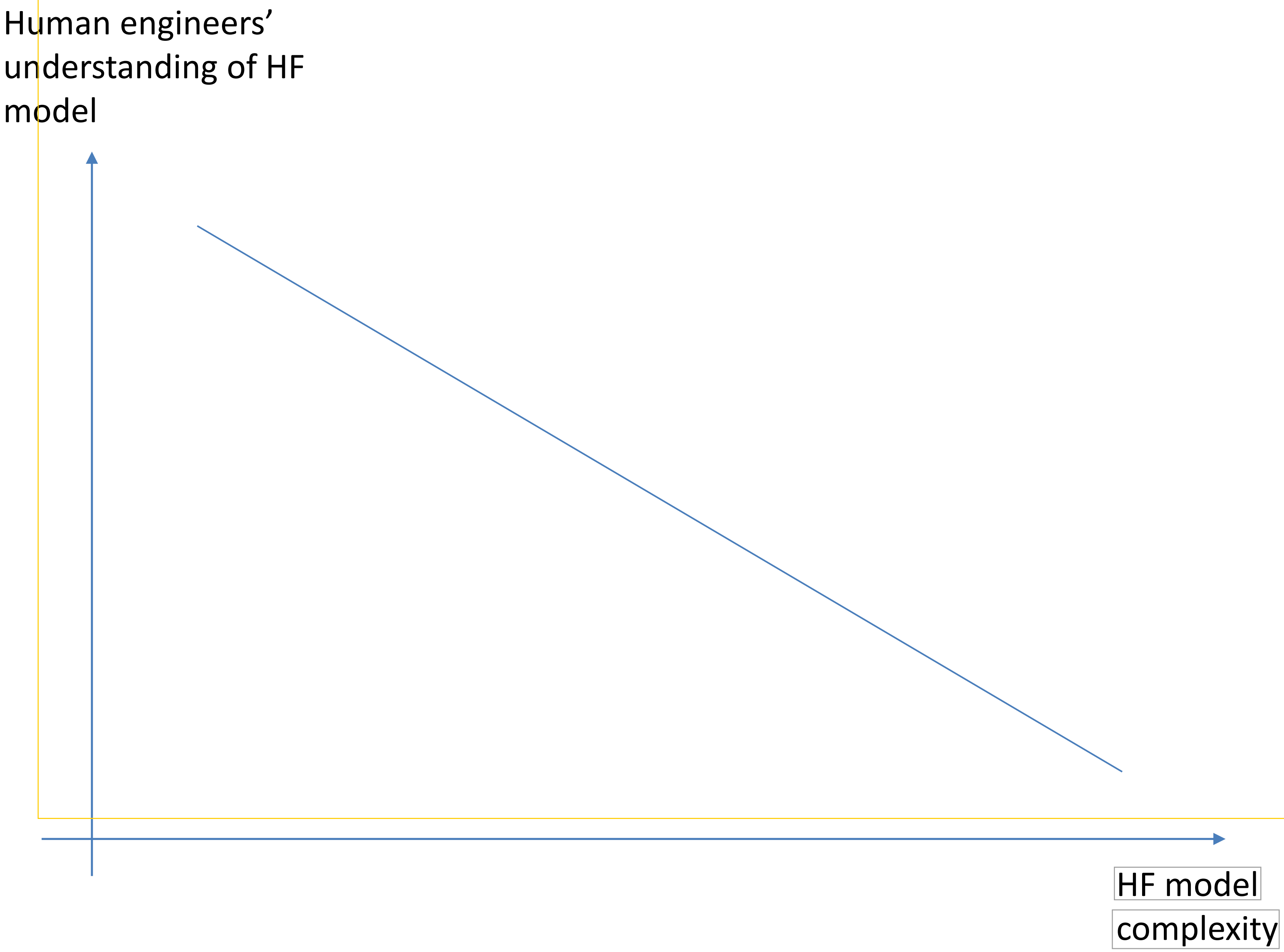
Figure 2. Representation of two resources, supplying the different stages of information processing. Sensory processing, the operation of the peripheral visual and auditory systems, is assumed to be relatively resource-free (after Wickens and Hollands 2000).



"This is really complicated, just doesn't make sense to me"

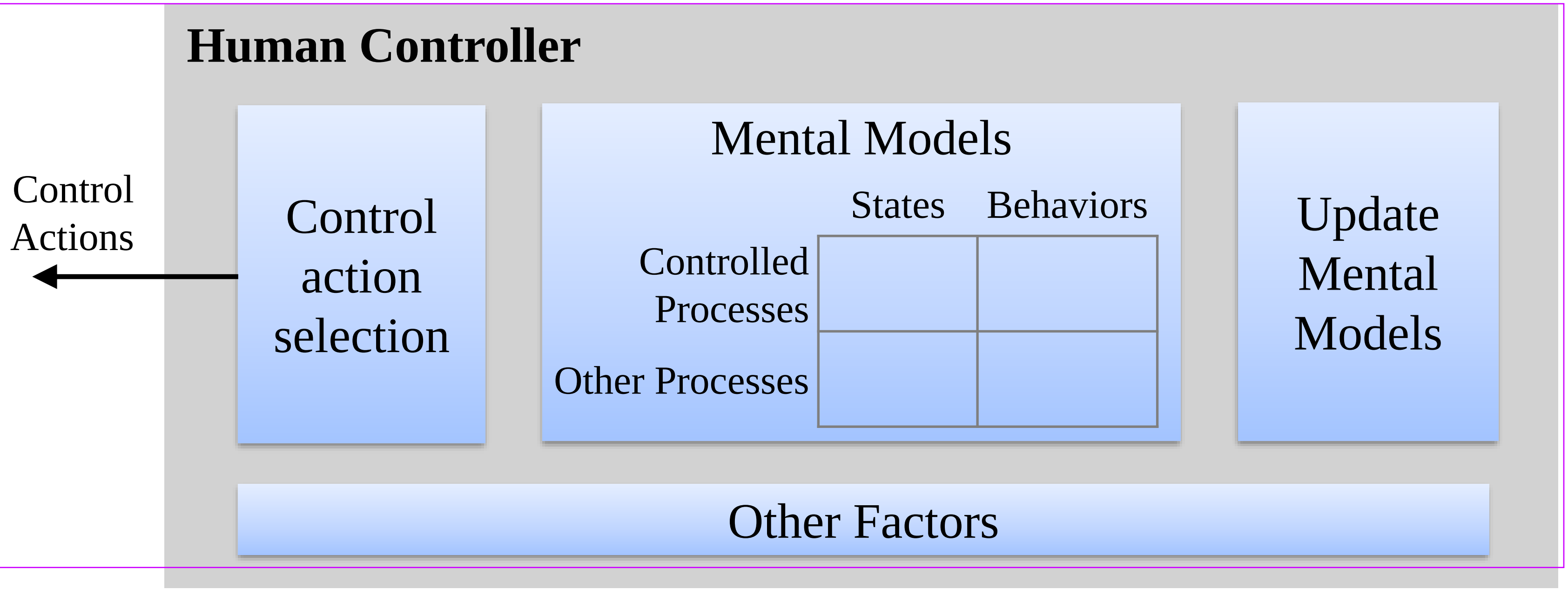
– Fredrik Matheson, "Promoting trust in AI applications"


Human
understanding of
automation
Automation
complexity

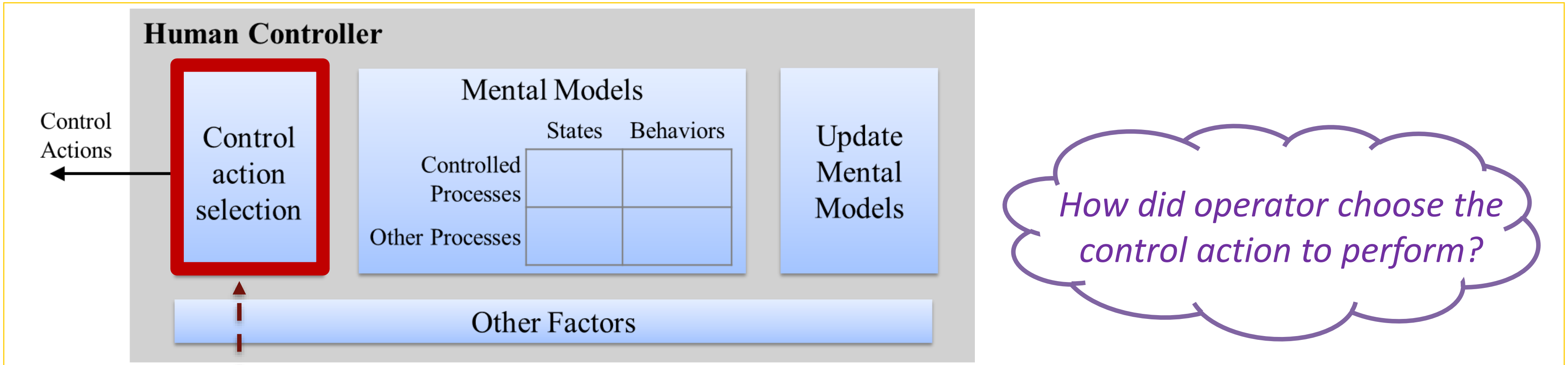

Human engineers' understanding of HF model
HF model complexity

Tradeoff

STPA Human Model



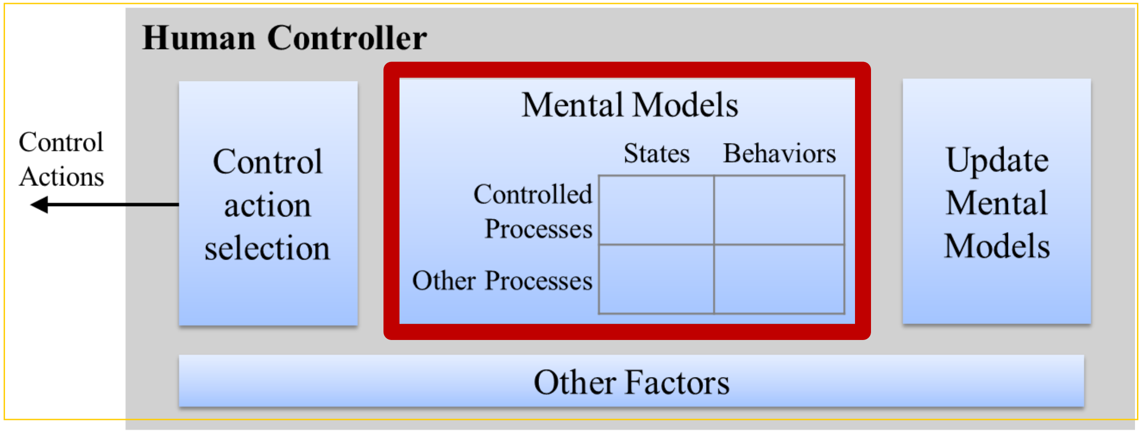
Control Action Selection

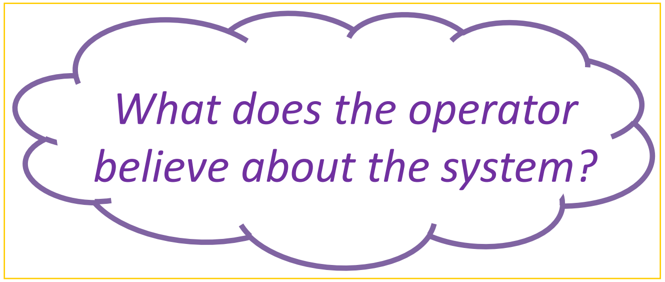


Control Action Selection

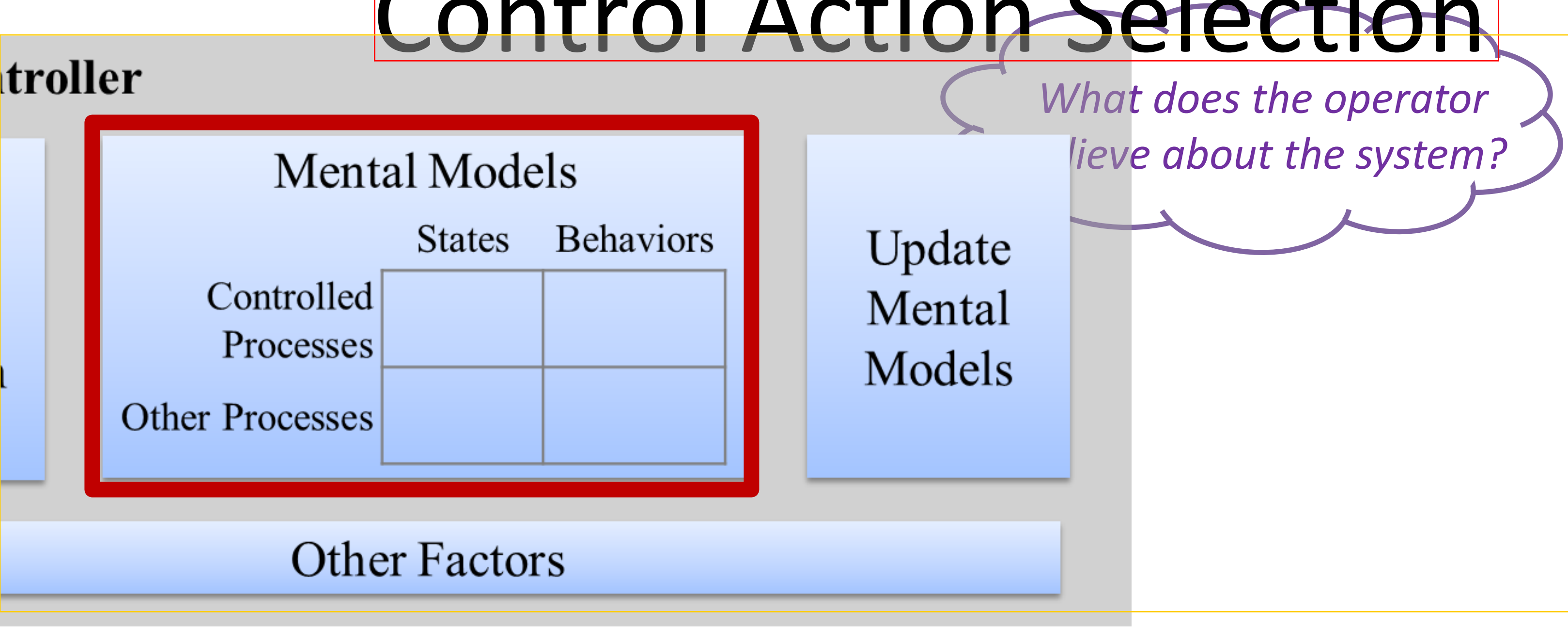
- What were the operator's goals?
- What alternatives was the operator choosing between?
- How automatic or novel was the behavior?
- How might the operator's mental models affect their decision?
- What external factors (*eg. time pressure)* might affect their decision?

Control Action Selection





Control Action Selection



Mental models

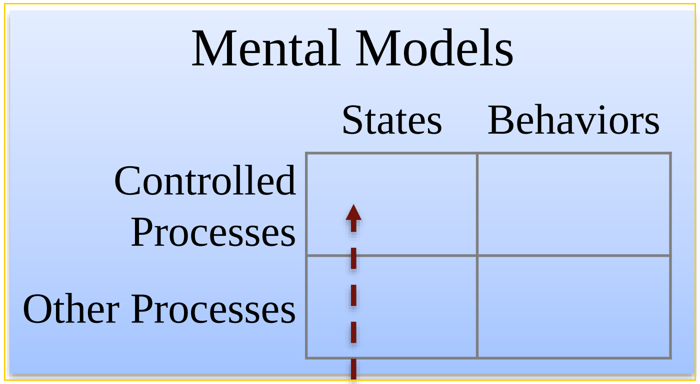
Mental Models

	States	Behaviors
Controlled Processes		
Other Processes		

What does the operator believe about the system?

Mental models

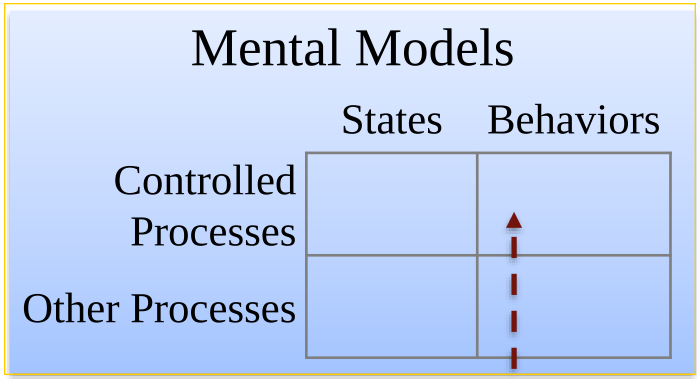
What does the operator believe about the system?

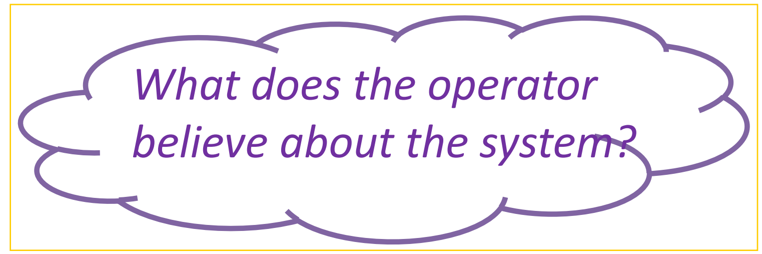


Mental Model of Controlled Process States

- Controlled processes: directly or indirectly controlled (e.g. automation, aircraft, engines, etc.)
- Beliefs about modes and mode changes
- Believes about the current process stage, for processes with multiple stages
- Beliefs about system variables (eg. true/false)

Mental models

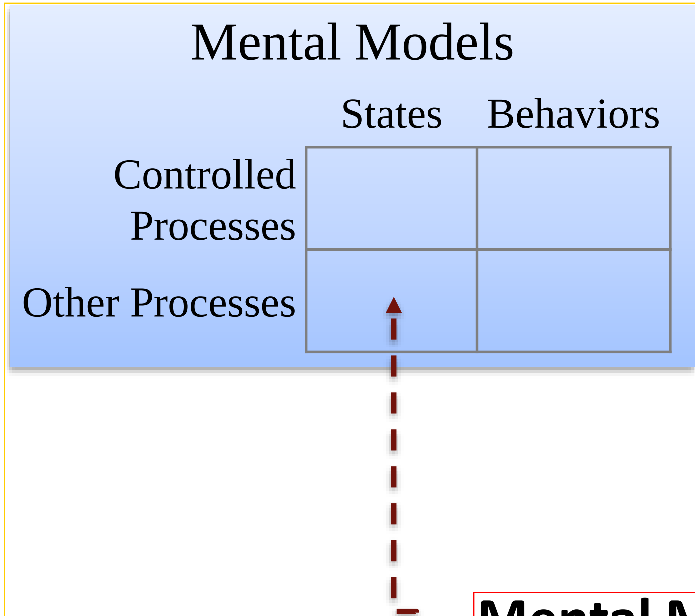


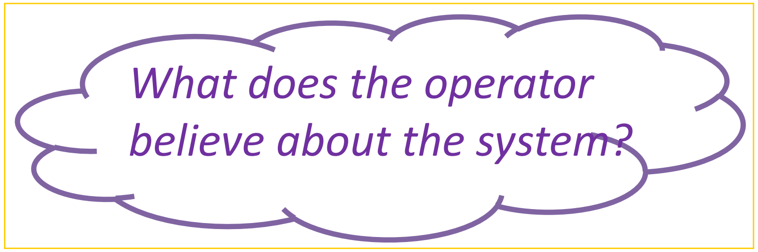


Mental Model of Controlled Process Behavior

- Beliefs about what processes can do
- Beliefs about how processes will behave in a particular mode or stage of operation
- Beliefs about if-then relationships between operator input and process output

Mental models

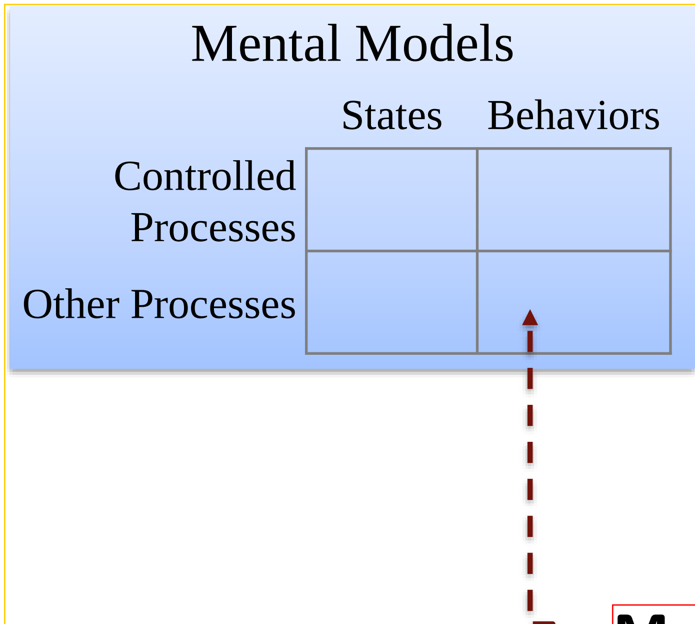




Mental Model of Other Process States

- Changes in environmental conditions
- Familiar or unfamiliar environments
- State of outside controllers (e.g. other pilots, ATC)
- Social and organizational conditions

Mental models

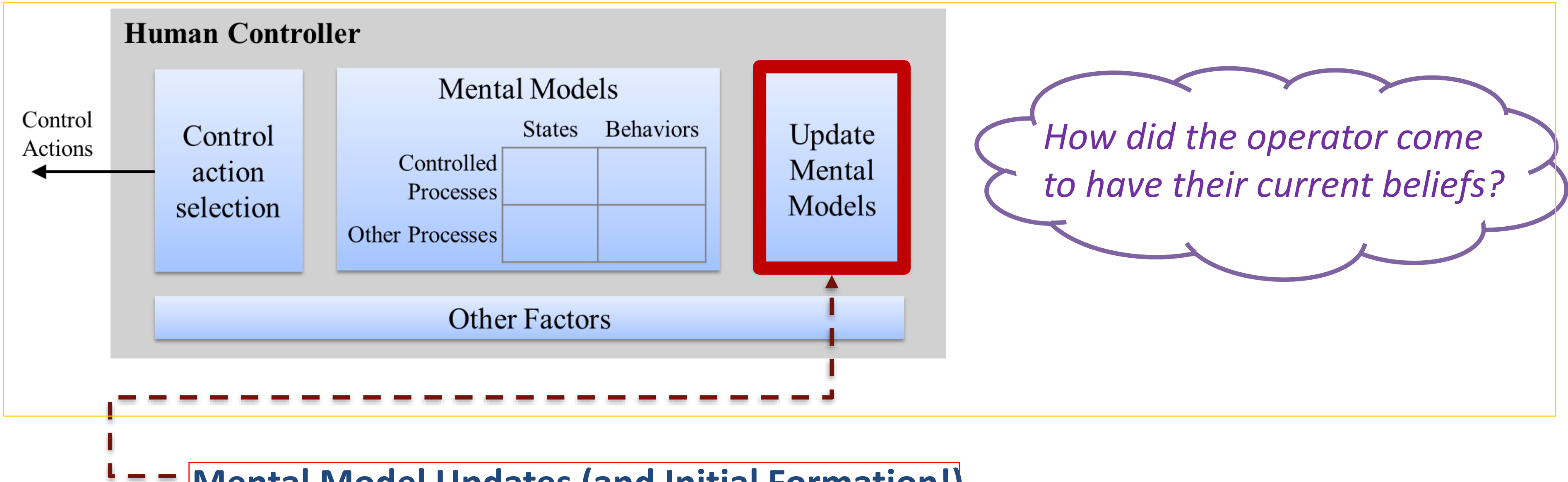


What does the operator believe about the system?

Mental Model of Other Process States

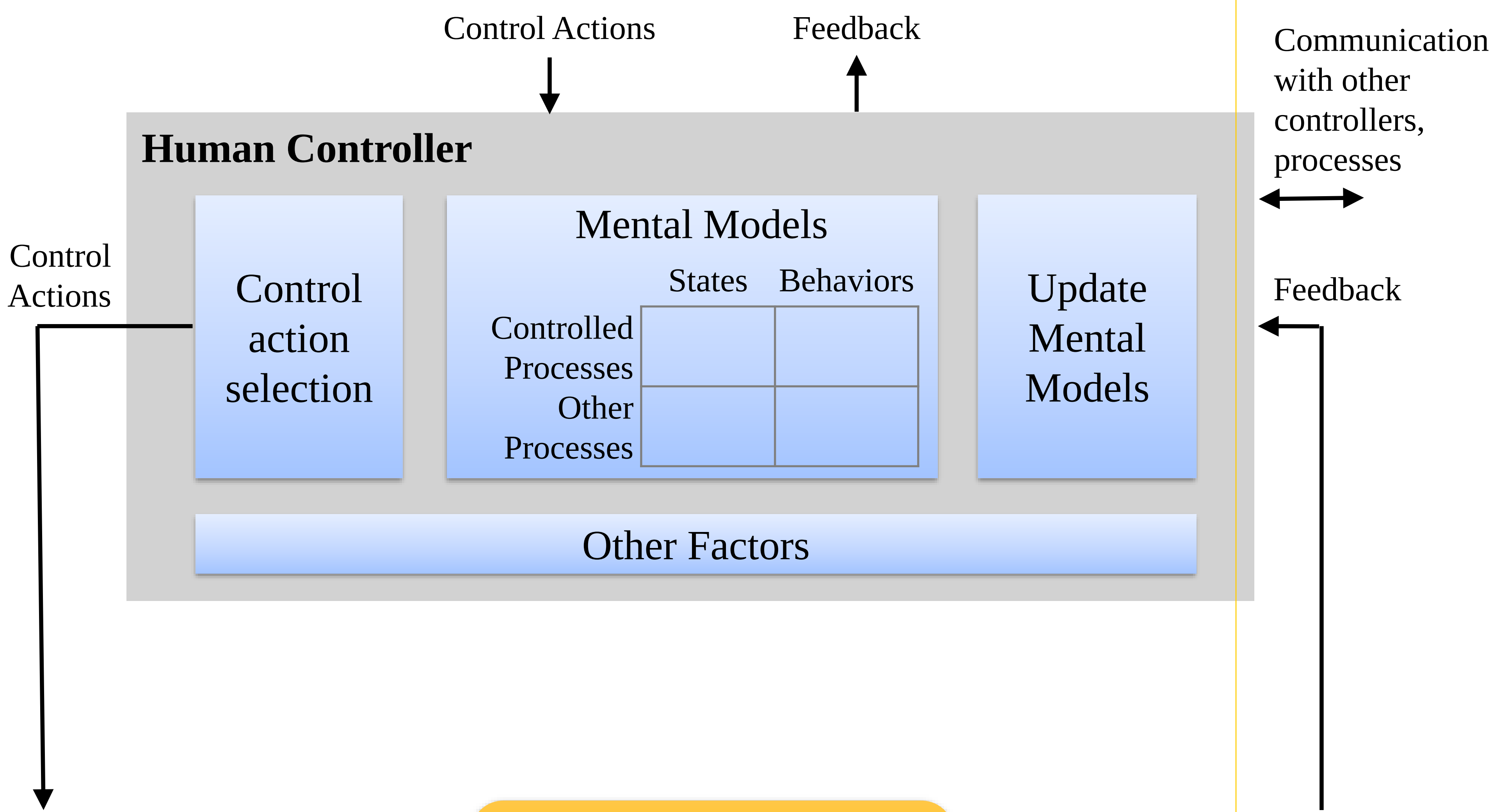
- Behavior and expectations of environment
- Capabilities of outside controllers (e.g. other pilots, ATC)
- Social and organizational expectations

Mental Model updates

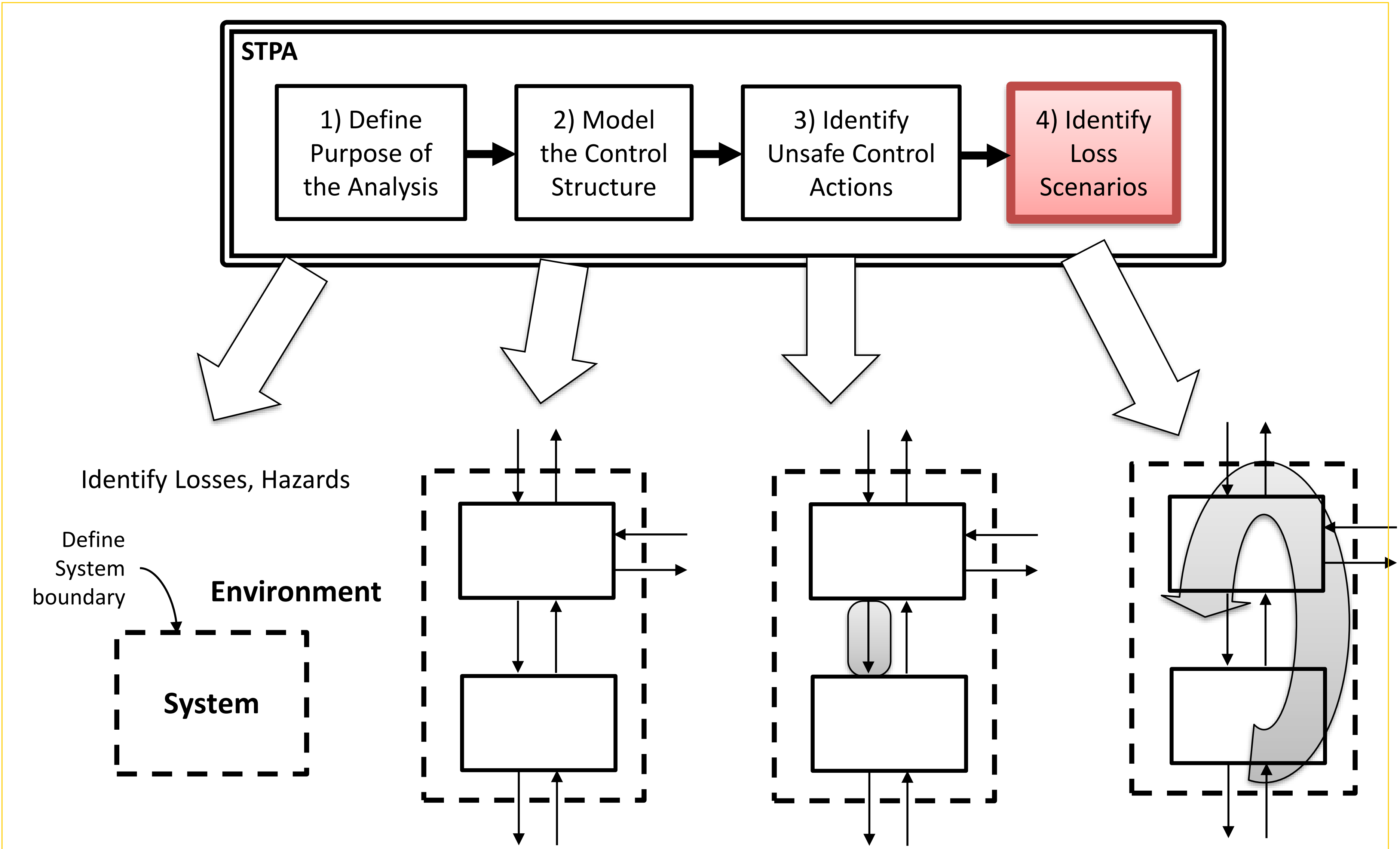


Mental Model Updates (and Initial Formation!)

- Consider initial formation of mental model vs. later updates
- Consider non-feedback inputs such as training programs and documentation
- Consider whether input/feedback was observed (*salience, expectations)*
- Consider whether input/feedback was correctly perceived & interpreted



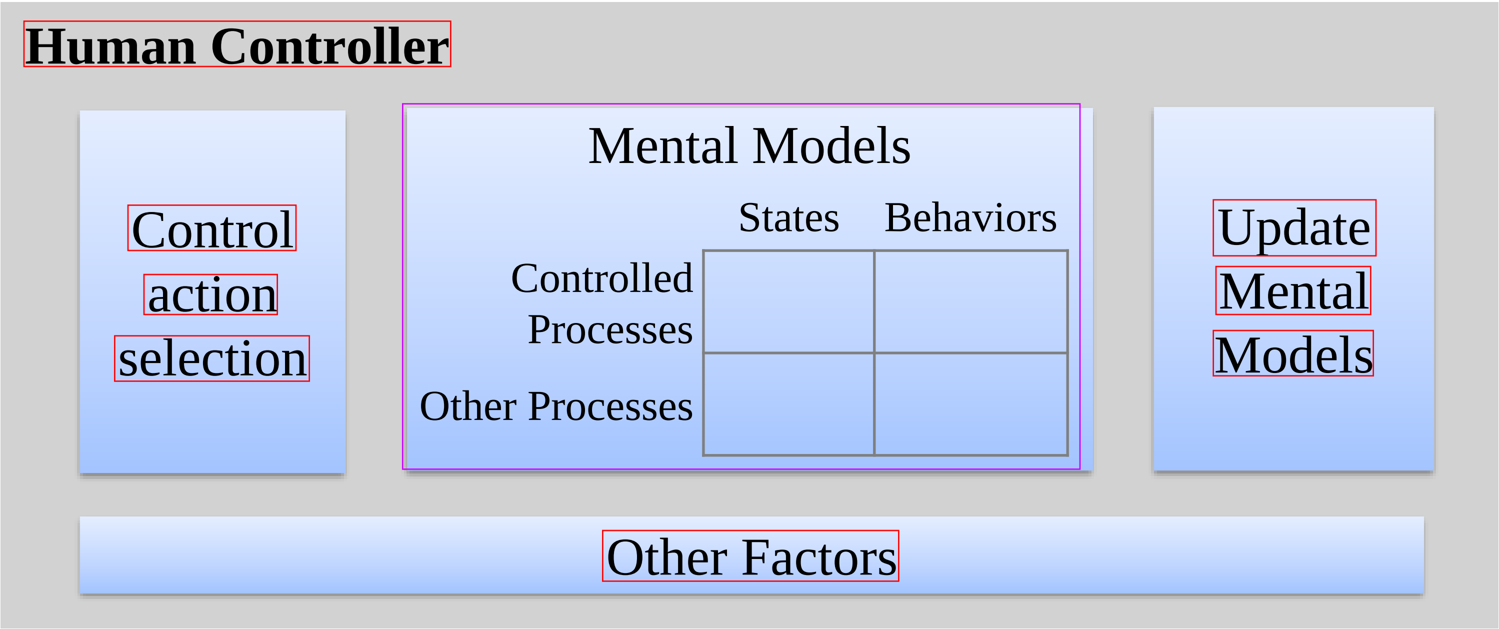
This model is based on accidents!


STPA
1) Define Purpose of the Analysis
2) Model the Control Structure
3) Identify Unsafe Control Actions
4) Identify Loss Scenarios
Identify Losses, Hazards
Define System boundary
Environment
System

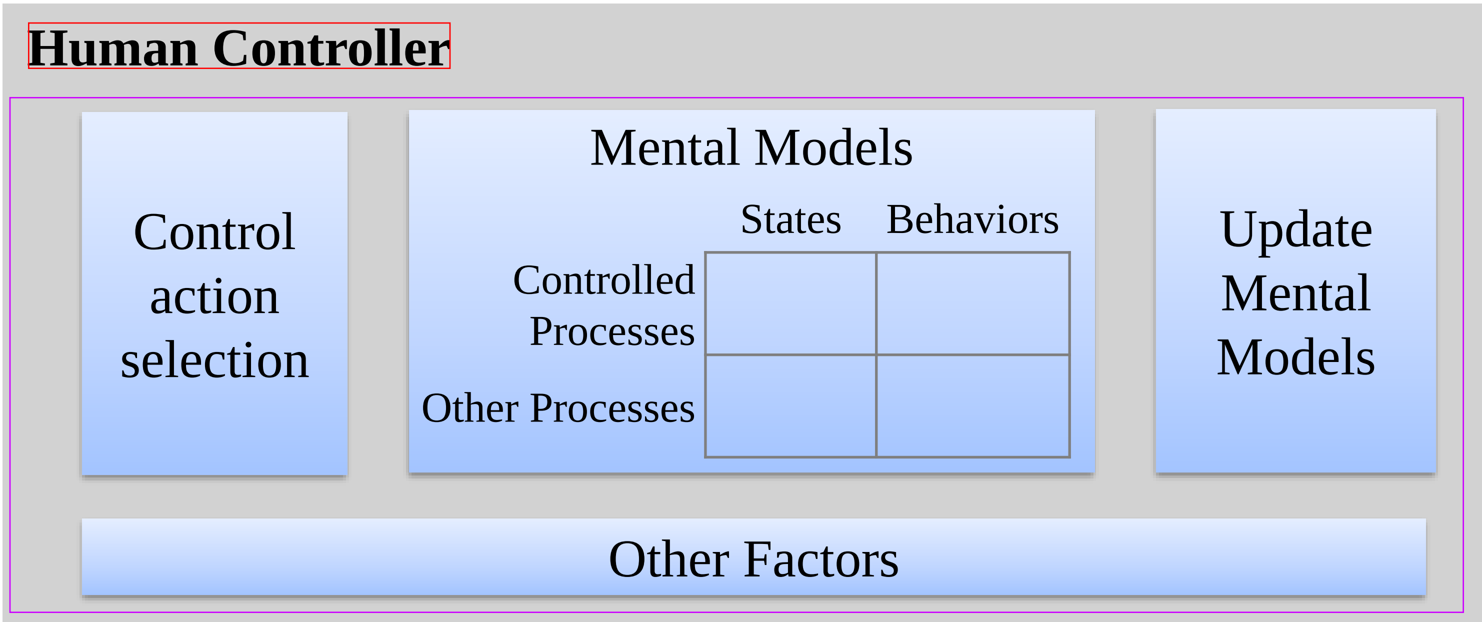
(Thomas, 2018)

ENGINEERING/ANALYSIS METHOD

- Losses, Hazards
- Control structure
- UCAs
- Build scenarios
  - Identify Mental Model variables
  - Identify Mental Model Flaws
  - Identify flaws in Mental Model Updates
  - Identify unsafe decisions (Control Action Selections)

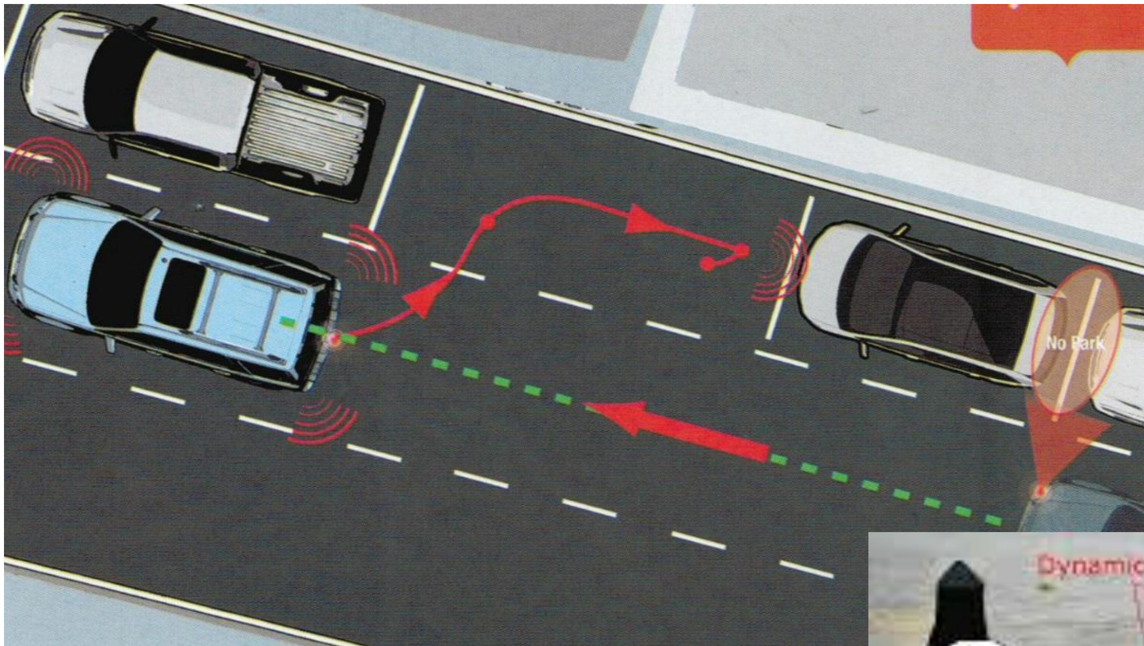


BENEFITS



- The new Engineering for Humans approach is **simple to apply**, and each part of the new model provides important insight into human behavior
- It provides **additional guidance** human scenarios, and can be used **early** in the design process
- Most importantly, it fits well into existing processes and provides a **"common language"** for engineers across disciplines to discuss issues

Automated parking assist



KEY ASSUMPTIONS ABOUT OUR SYSTEM

- The automation is capable of steering, braking, shifting, and accelerating.
- The driver is expected to monitor the system to respond to unexpected events and obstacles.
- The driver may temporarily override the APA computer's actions by braking or accelerating for short periods of time.
- If the driver
  - grabs the wheel
  - accelerates above a given maximum speed
  - brakes for more than 2 seconds
  - or presses the APA button

the automation will be fully disabled.

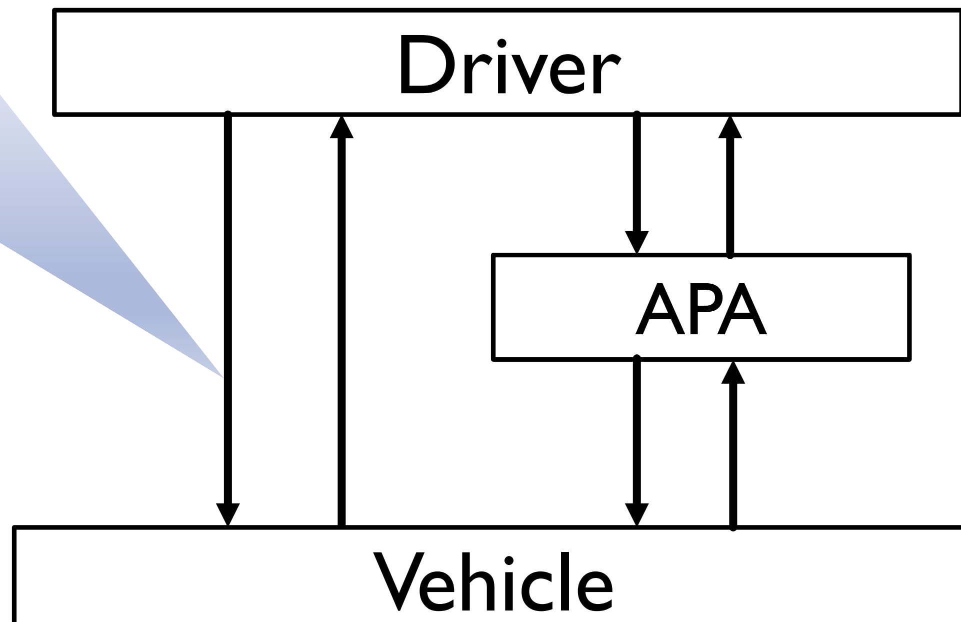
ACCIDENTS AND HAZARDS

System Level Accidents	
A-1	Death, injury, or property damage resulting from a collision with a person, vehicle, object, or terrain.
A-2	Injury or property damage occurring within the vehicle, without a collision.
A-3	Loss of customer satisfaction with automated parking, without injury or property damage.

System Level Hazards	
H-1	The vehicle does not maintain a safe minimum distance between itself and obstacles such as pedestrians, vehicles, objects, and terrain. [A-1]
H-2	Occupants or cargo are subjected to sudden high forces that may result in injury or property damage. [A-2]
H-3	The vehicle parks inappropriately, either in an unsuitable space (e.g. blocking a fire hydrant) or in violation of parking guidelines (e.g. excessively far from the curb). [A-3]

UNSAFE CONTROL ACTIONS

	Not Provided	Provided	Too early, too late, out of order	Stopped too soon, applied too long
Brake	**UCA-1: Driver does not brake when auto-parking and computer doesn't react to an obstacle**			

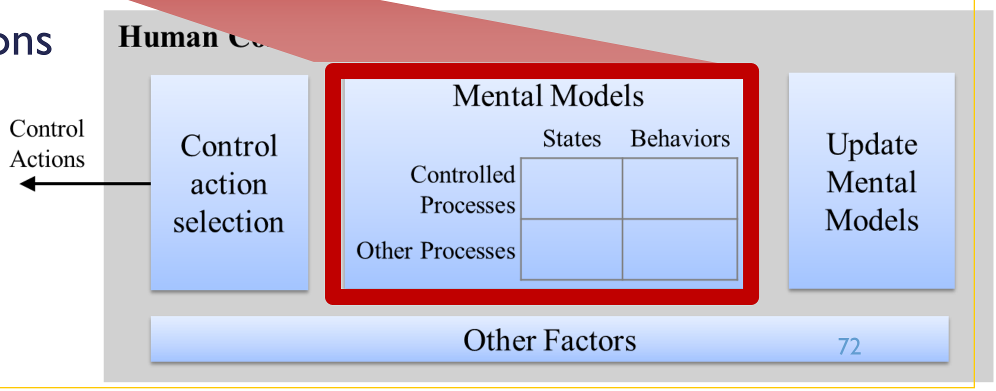


NEW PROCESS

- Identify UCAs
  - UCA-1: Driver does not brake when auto-parking and computer doesn't react to an obstacle
- Identify Mental Model variables
  - MM-1: APA is enabled/disabled
  - MM-2: APA computer reacting appropriately/inappropriately
  - MM-3: Obstacle on collision path
- Identify Mental Model Flaws
- Identify flaws in Mental Model Updates
- Identify unsafe Control Action Selections

NEW PROCESS

- Identify UCAs
  - UCA-1: Driver does not brake when auto-parking and computer doesn't react to an obstacle
- Identify Mental Model variables
  - MM-1: APA is enabled/disabled
  - MM-2: APA computer reacting appropriately/inappropriately
  - MM-3: Obstacle on collision path
- Identify Mental Model Flaws
- Identify flaws in Mental Model Updates
- Identify unsafe Control Action Selections



NEW PROCESS

- Identify UCAs
- Identify Mental Model variables
  - MM-1: APA is enabled/disabled
  - MM-2: APA computer reacting appropriately/inappropriately
  - MM-3: Obstacle on collision path
- Identify Mental Model Flaws
- Identify unsafe decisions (Control Action Selections)
- Identify inadequate Mental Model Updates

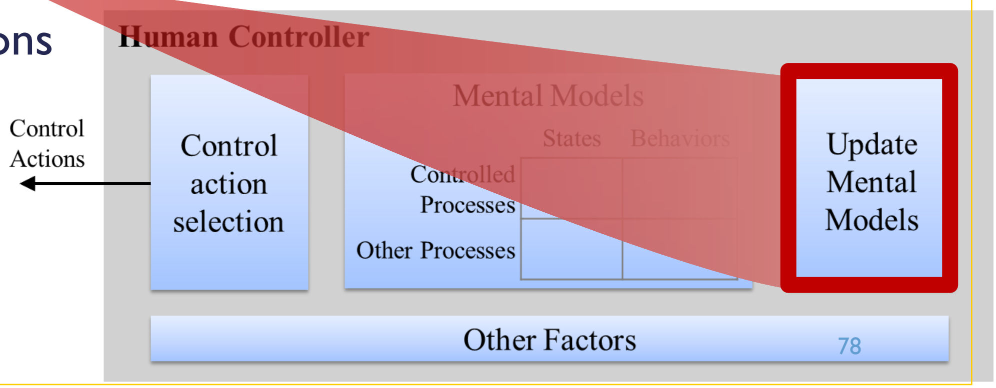
Mental Models

	States	Behaviors
Controlled Processes	1.	2.
Other Processes	3.	4.

Type of MM flaw	Examples
1) Incorrect beliefs about controlled process state (including modes)	Driver thinks APA is enabled when APA is really disabled
2) Incorrect beliefs about controlled process behaviors	Driver thinks APA is reacting properly and will brake automatically
3) Incorrect beliefs about other process state (e.g. environment)	Driver thinks there is no obstacle when there is one
4) Incorrect beliefs about other process behavior (e.g. environment)	Driver knows there is an obstacle, but thinks it won’t move on a collision path

NEW PROCESS

- Identify UCAs
  - UCA-1: Driver does not brake when auto-parking and computer doesn't react to an obstacle
- Identify Mental Model variables
  - MM-1: APA is enabled/disabled
  - MM-2: APA computer reacting appropriately/inappropriately
  - MM-3: Obstacle on collision path
- Identify Mental Model Flaws
- Identify flaws in Mental Model Updates
- Identify unsafe Control Action Selections



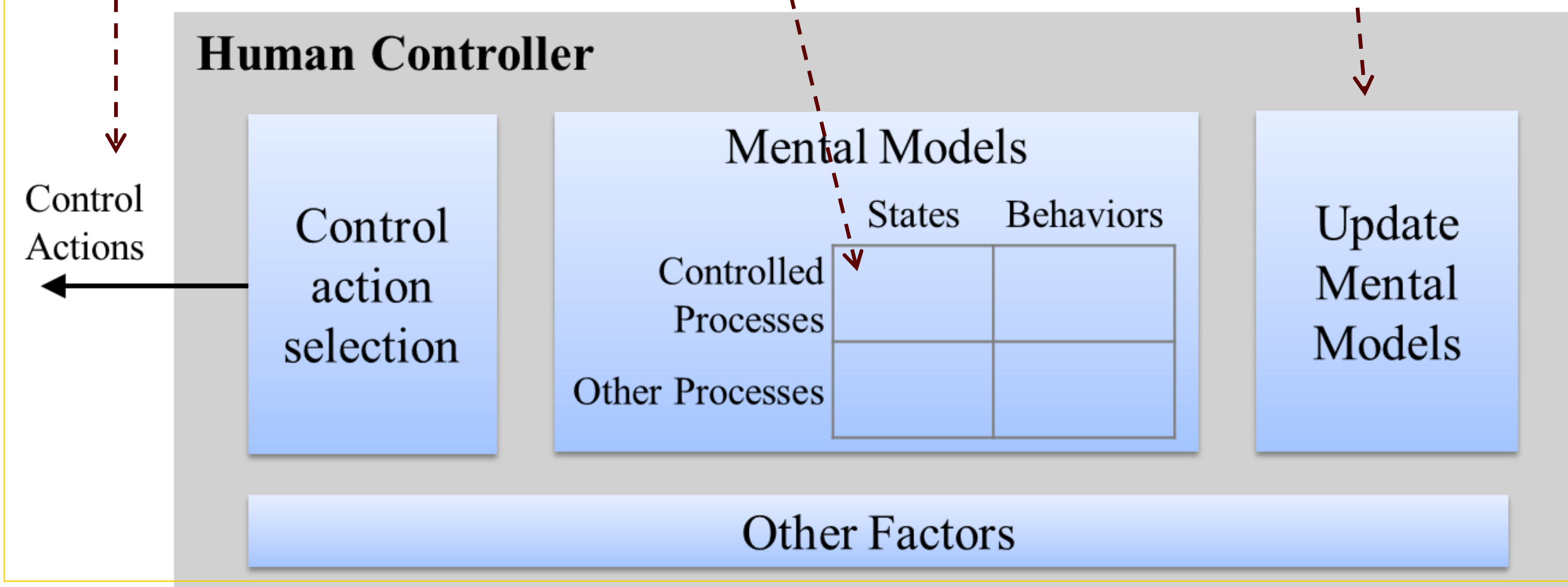
STPA: ENGINEERING FOR HUMANS

Driver does not brake when auto-parking and computer doesn’t react to an obstacle (UCA-1)

Driver believes APA is enabled when APA is really disabled (MM-1)

Consider:

1. Automatic mode changes
2. Previous cmds ignored
3. Phases of operation
4. Etc.

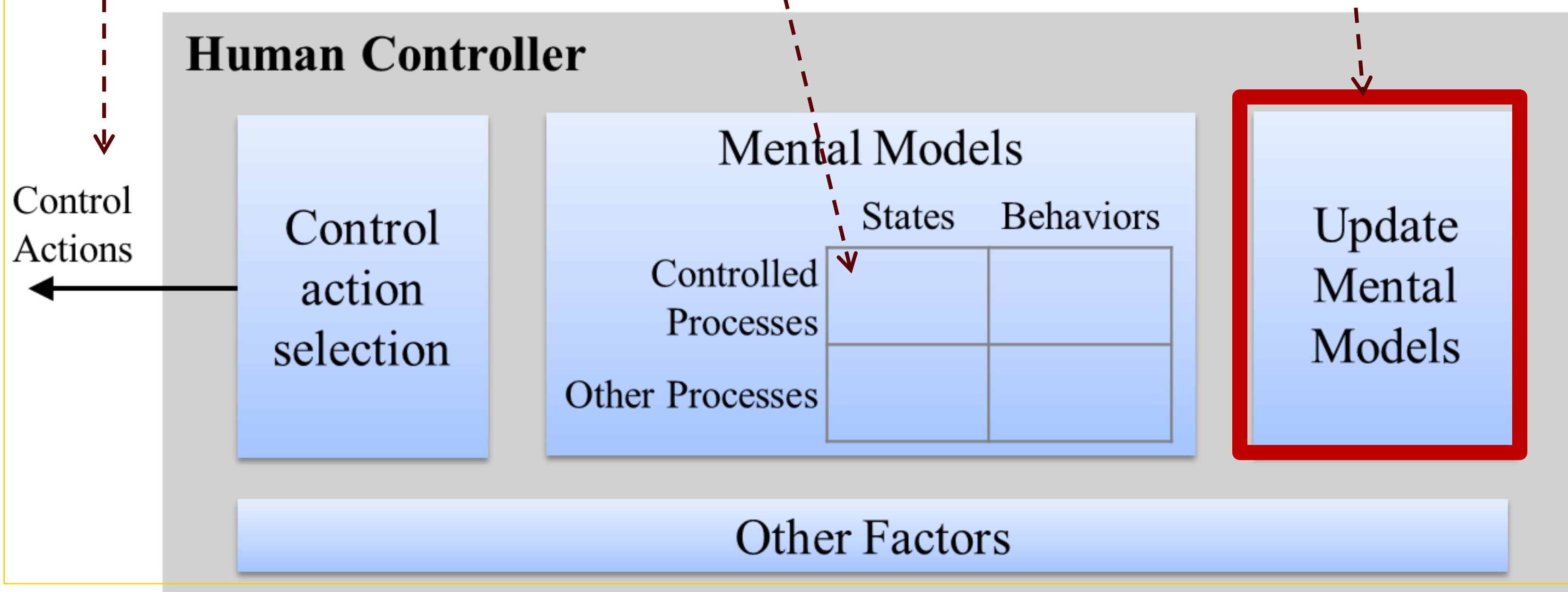


STPA: ENGINEERING FOR HUMANS

Driver does not brake when auto-parking and computer doesn't react to an obstacle (UCA-1)

Driver believes APA is enabled when APA is really disabled (MM-1)

APA automatically disabled itself but driver didn't notice the change

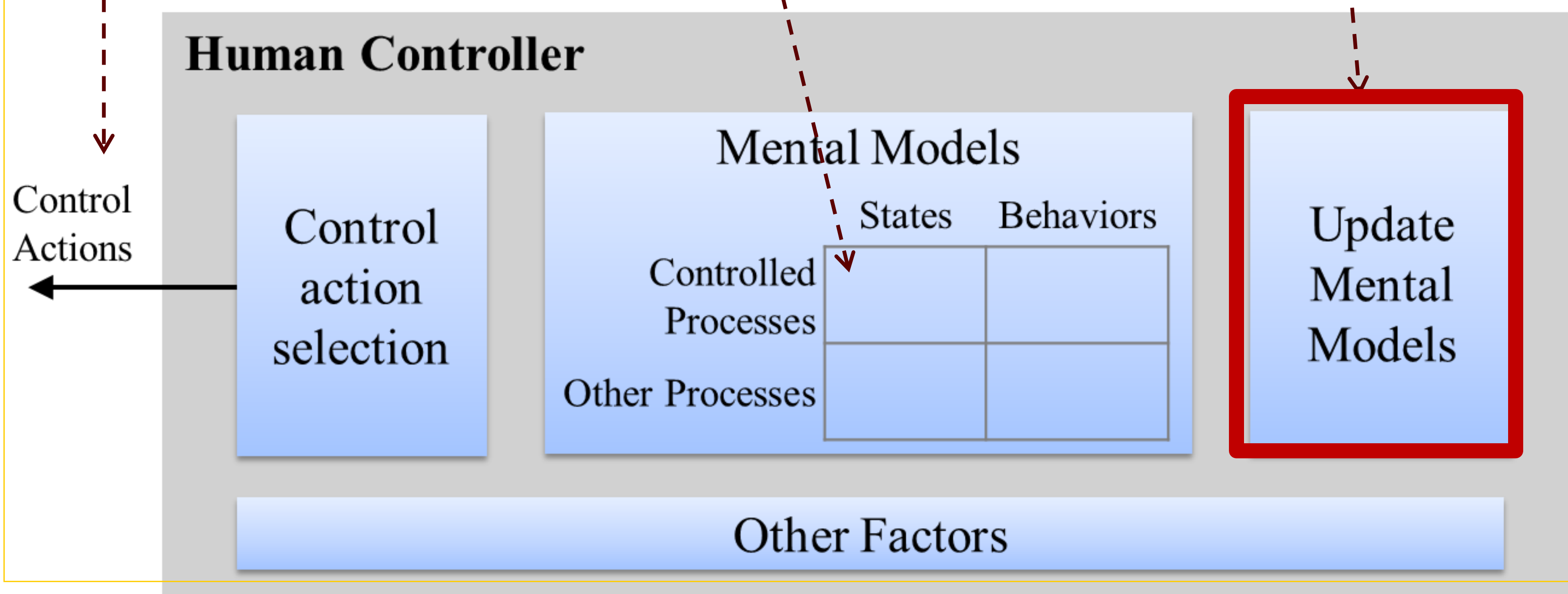


STPA: ENGINEERING FOR HUMANS

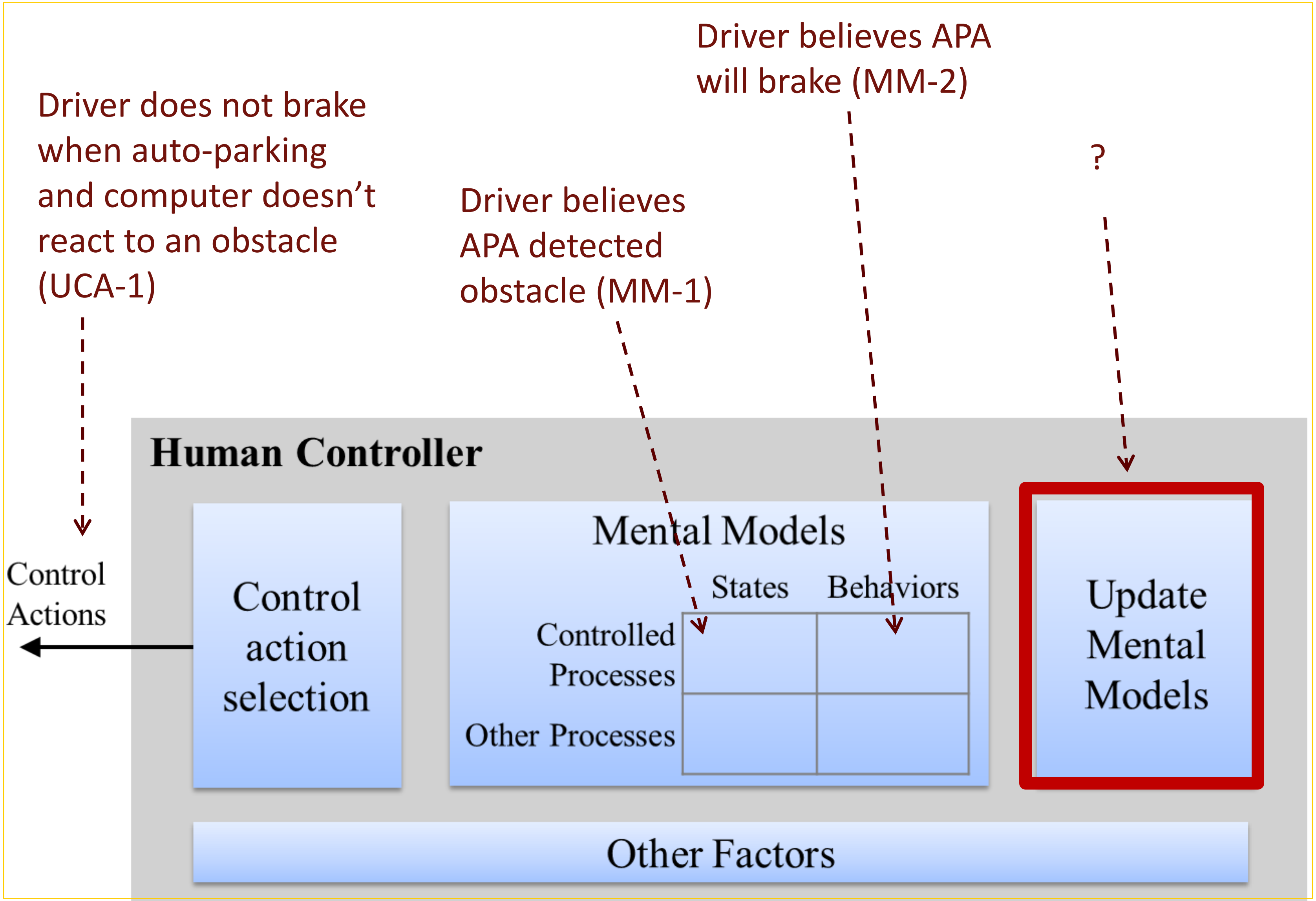
Driver does not brake when auto-parking and computer doesn’t react to an obstacle (UCA-1)

Driver believes APA is enabled when APA is really disabled (MM-1)

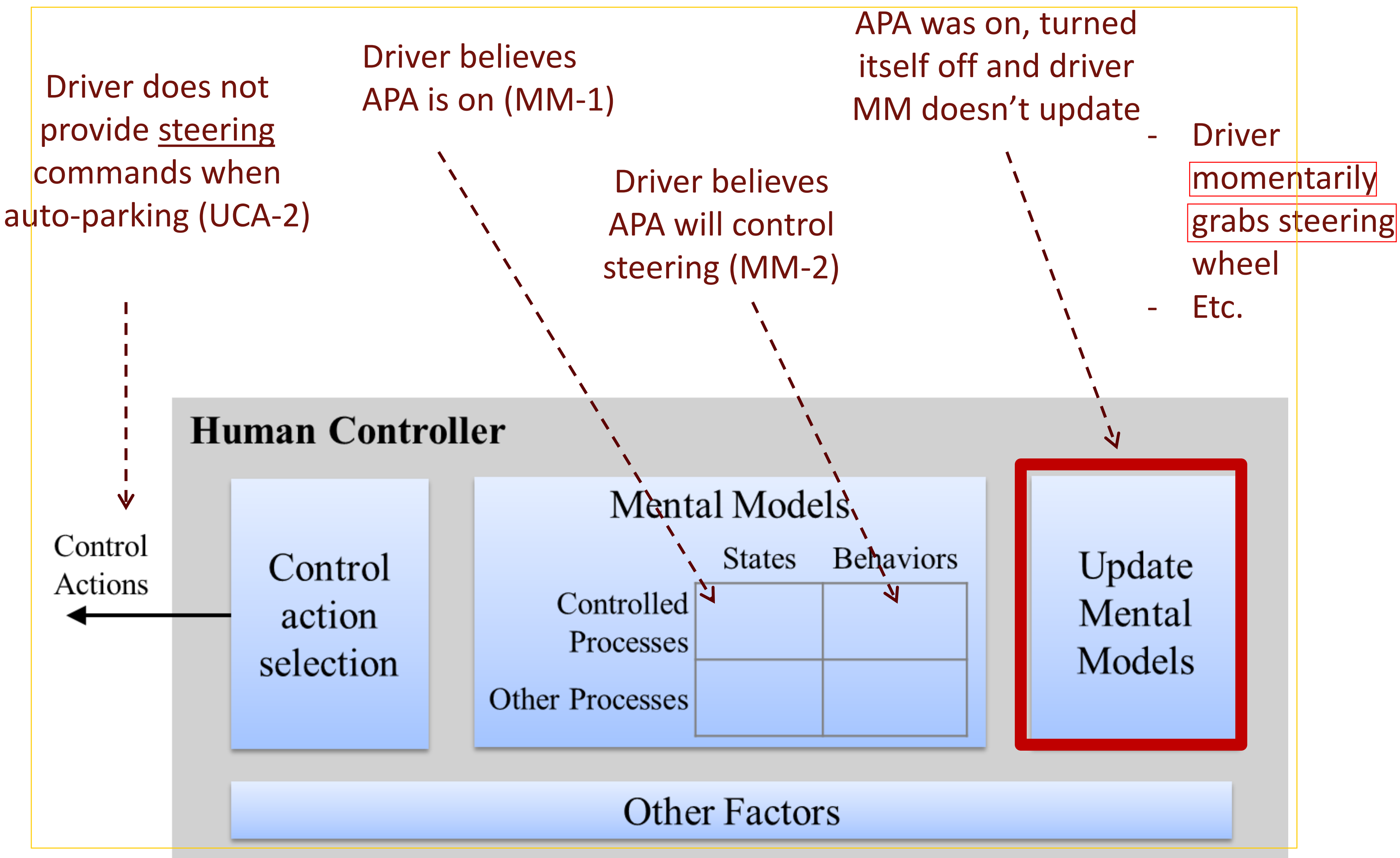
APA automatically disabled itself, driver noticed the change but didn’t understand it



STPA: ENGINEERING FOR HUMANS

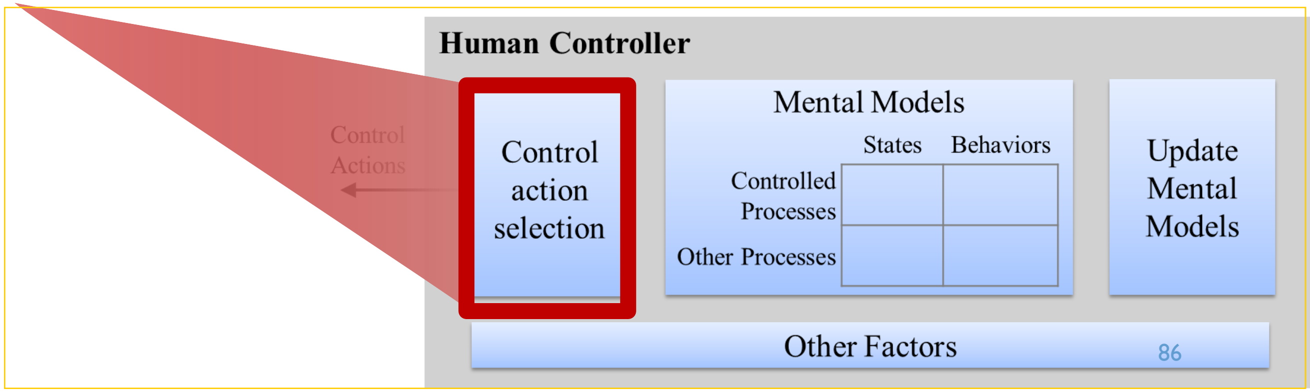


STPA: ENGINEERING FOR HUMANS



NEW PROCESS

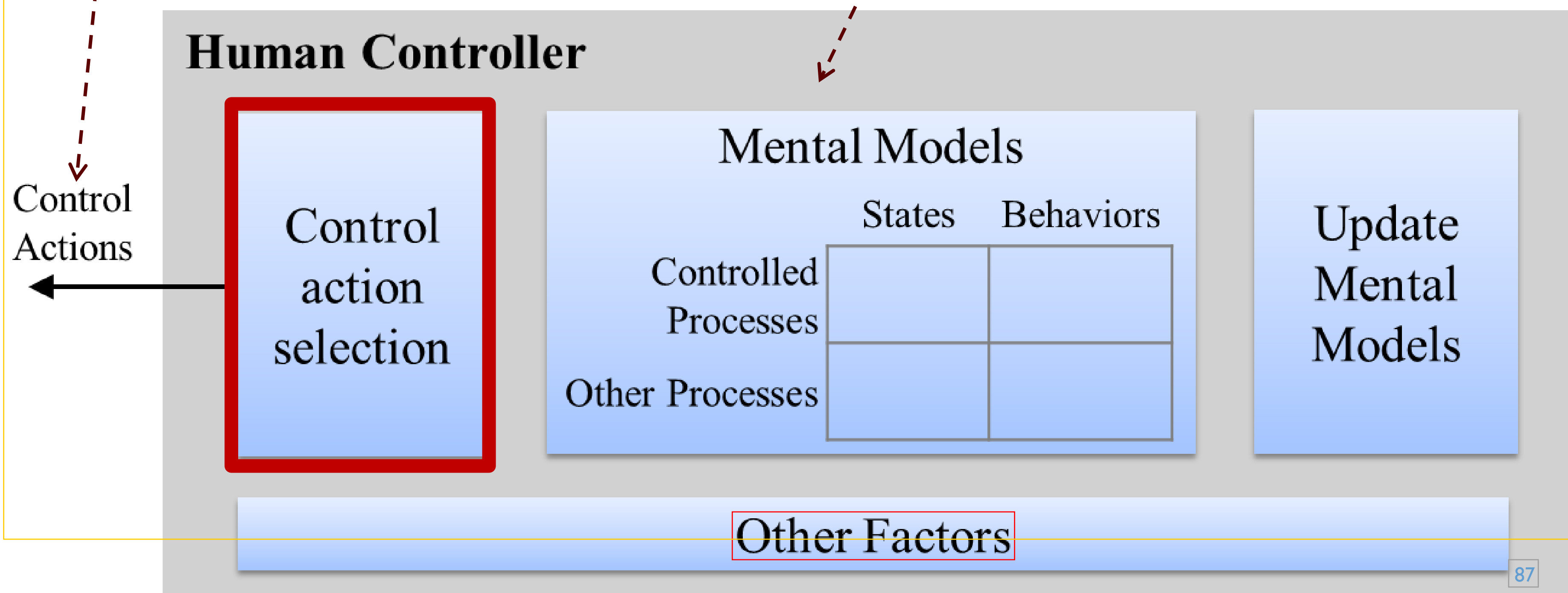
- Identify UCAs
  - UCA-1: Driver does not brake for an obstacle when computer does not react appropriately to the obstacle
- Identify Mental Model variables
  - MM-1: APA reacting appropriately/inappropriately
  - MM-2: Obstacle on collision path
- Identify Mental Model Flaws
- Identify flaws in Mental Model Updates
- Identify unsafe Control Action Selections



STPA: ENGINEERING FOR HUMANS

Driver knows APA is on
Driver knows APA hasn't reacted yet
Driver knows there is an obstacle in the way
Driver knows obstacle is on collision path

Driver does not brake when auto-parking and computer doesn't react to an obstacle (UCA-1)

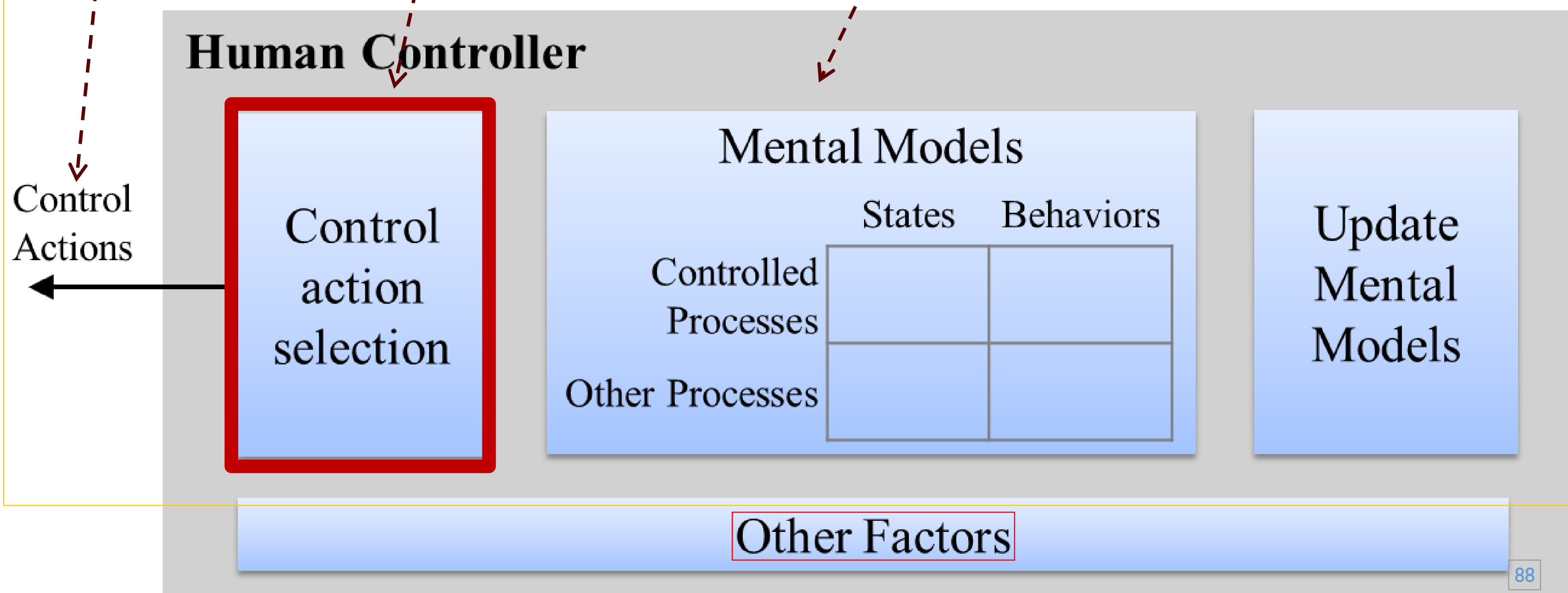


STPA: ENGINEERING FOR HUMANS

Driver knows APA is on
Driver knows APA hasn't reacted yet
Driver knows there is an obstacle in the way
Driver knows obstacle is on collision path

Driver does not brake when auto-parking and computer doesn't react to an obstacle (UCA-1)

Driver may not know they can control brake with APA on

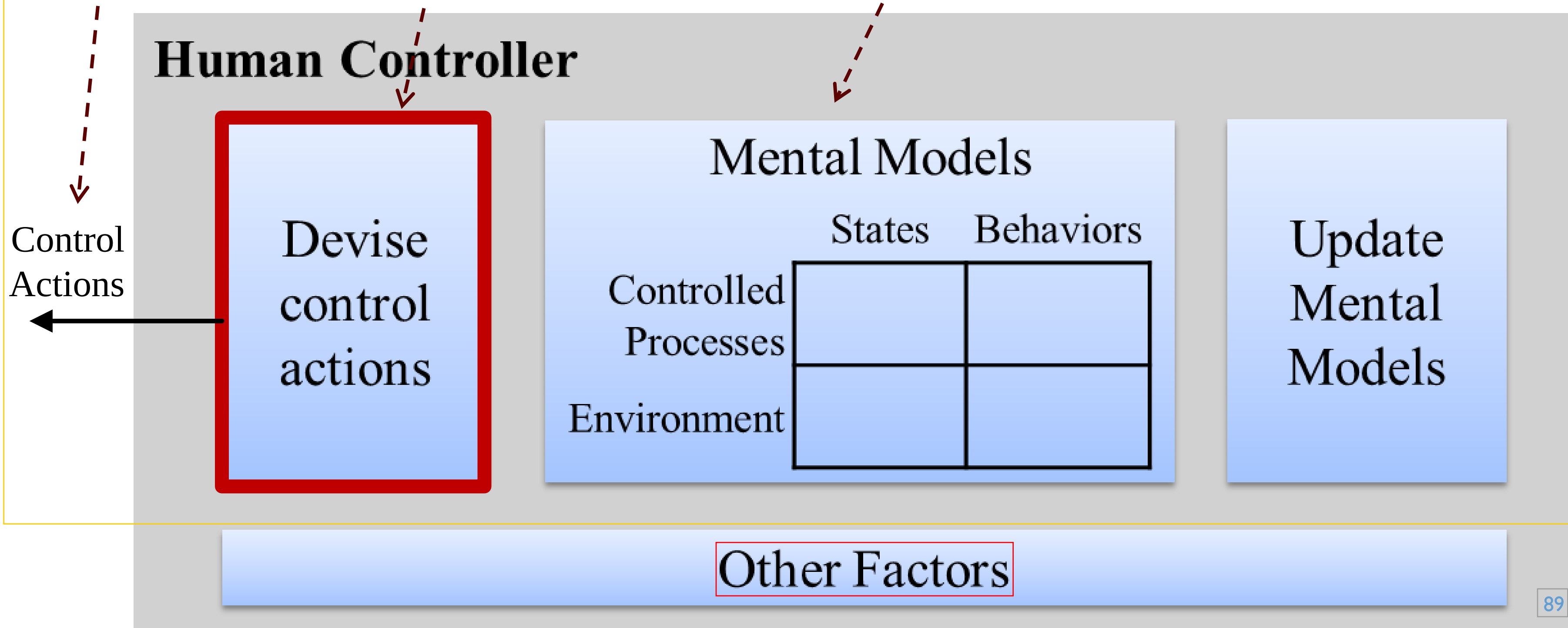


STPA: ENGINEERING FOR HUMANS

Driver knows APA is on
Driver knows APA hasn't reacted yet
Driver knows there is an obstacle in the way
Driver knows obstacle is on collision path

Driver does not brake when auto-parking and computer doesn't react to an obstacle (UCA-1)

Driver may decide to disable APA instead

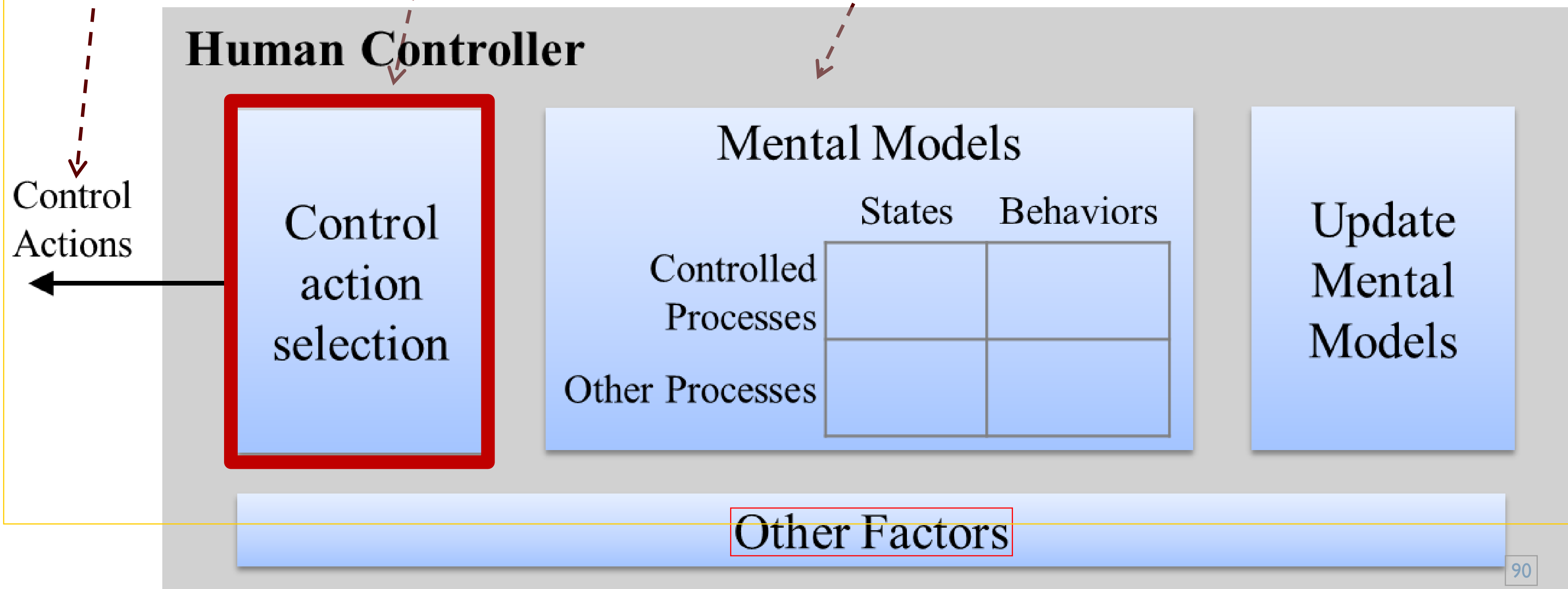


STPA: ENGINEERING FOR HUMANS

Driver does not brake when auto-parking and computer doesn't react to an obstacle (UCA-1)

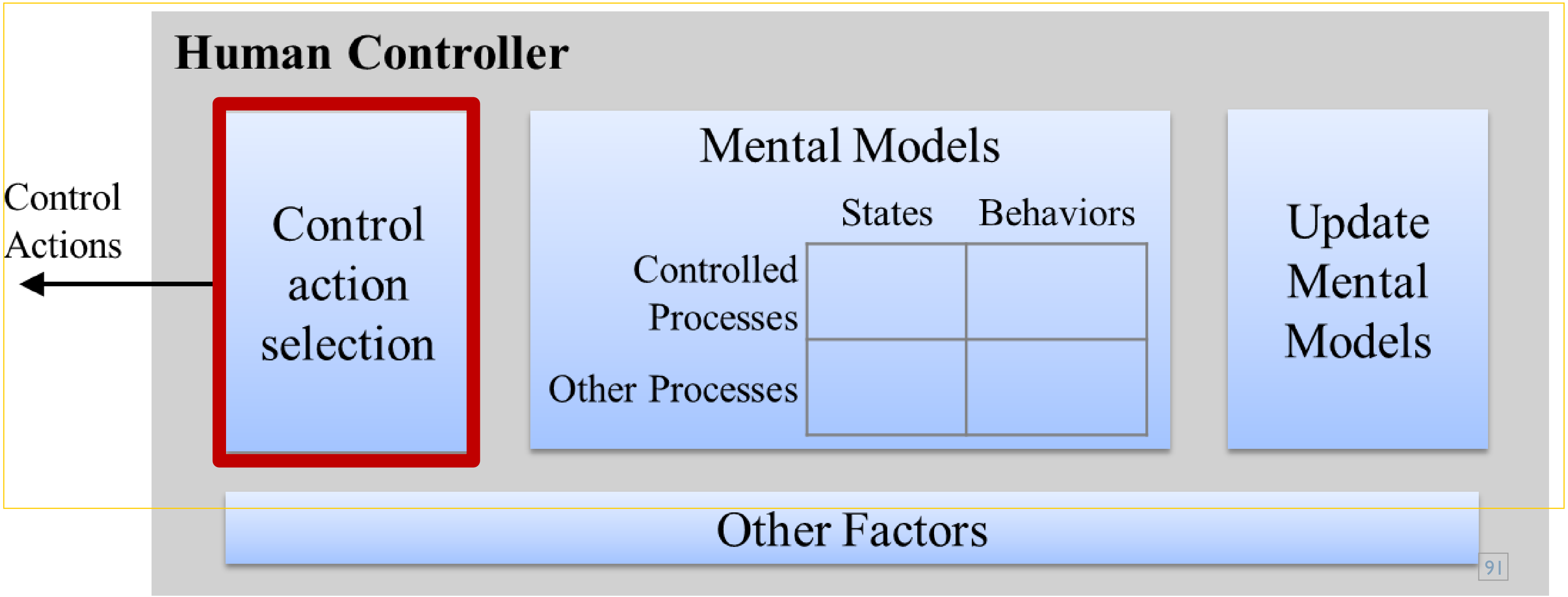
Driver may still be waiting for APA to act

Driver knows APA is on
Driver knows APA hasn't reacted yet
Driver knows there is an obstacle in the way
Driver knows obstacle is on collision path



STPA: ENGINEERING FOR HUMANS

- Identify unsafe Control Action Selections
  - Consider whether the driver is aware they can control X
  - Consider alternative driver controls/actions
  - Consider other driver goals



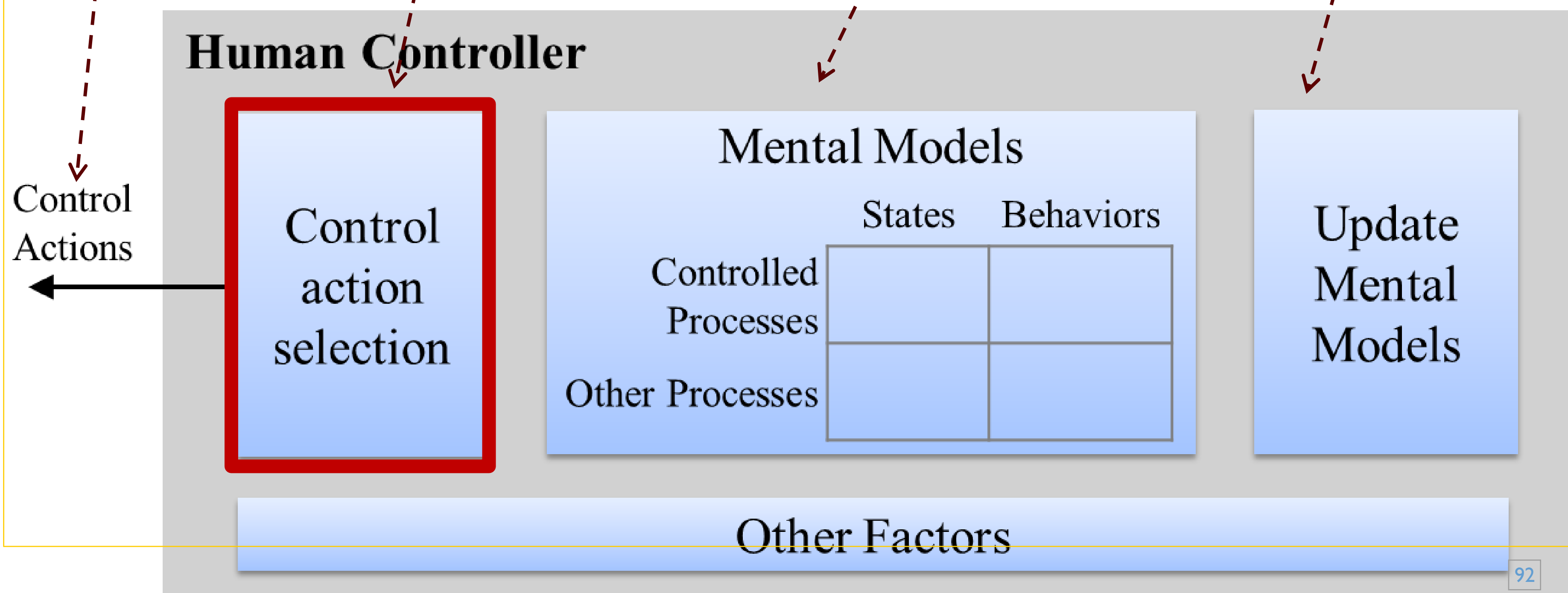
STPA: ENGINEERING FOR HUMANS

Driver does not brake when auto-parking and computer doesn't react to an obstacle (UCA-1)

Driver may still be waiting for APA to act

Driver knows APA is on
Driver knows APA hasn't reacted yet
Driver knows there is an obstacle in the way
Driver knows obstacle is on collision path

No mechanism provided to update expectation (will it brake?)



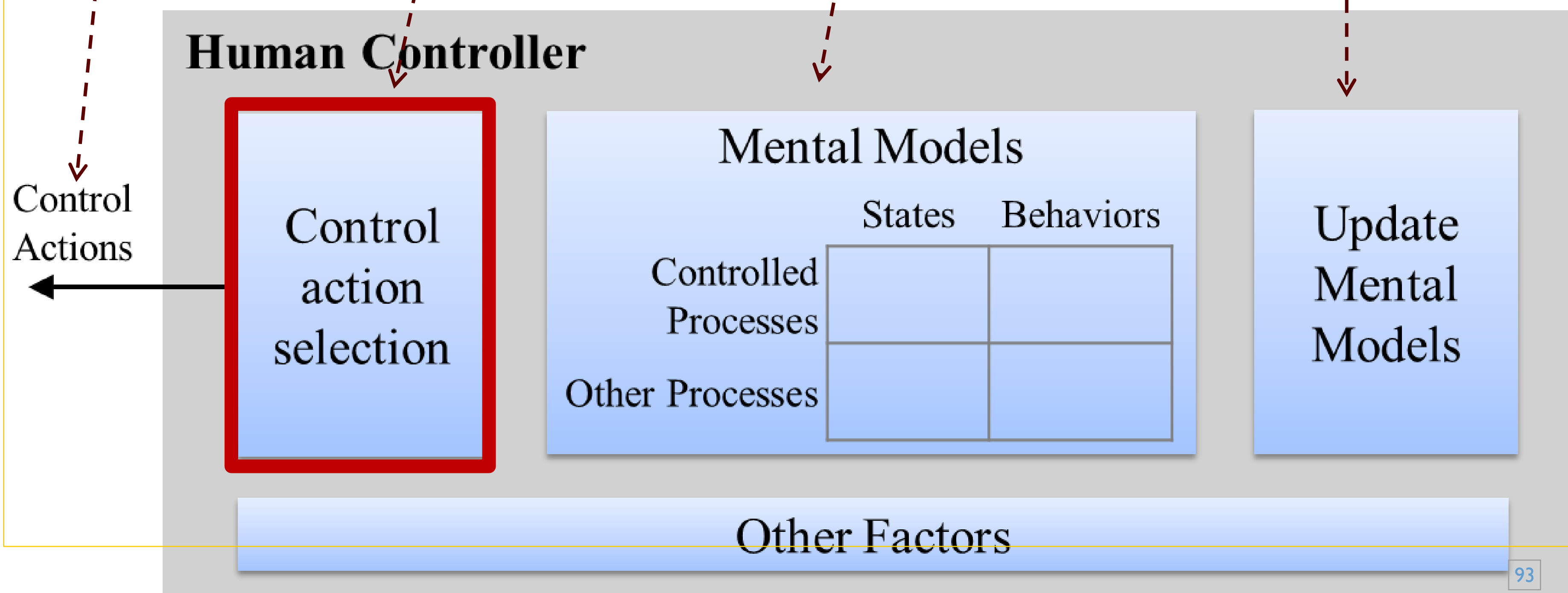
STPA: ENGINEERING FOR HUMANS

Driver does not brake when auto-parking and computer doesn't react to an obstacle (UCA-1)

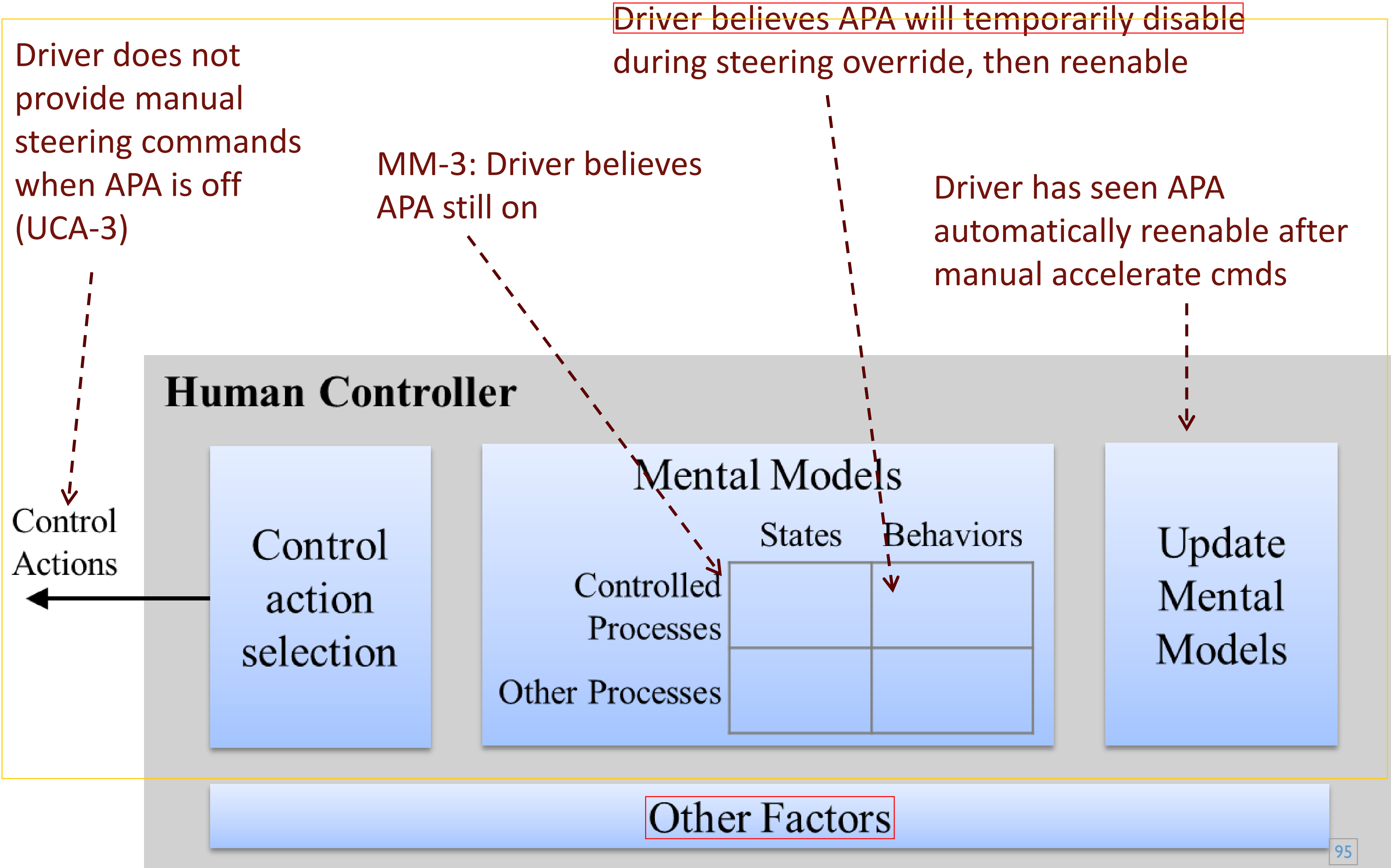
Driver rule: I don't need to brake when APA is on

Driver knows APA is on
Driver knows APA hasn't reacted yet
Driver knows there is an obstacle in the way
Driver knows obstacle is on collision path

Driver has seen APA brake for parked cars, assumes it can brake for all obstacle types

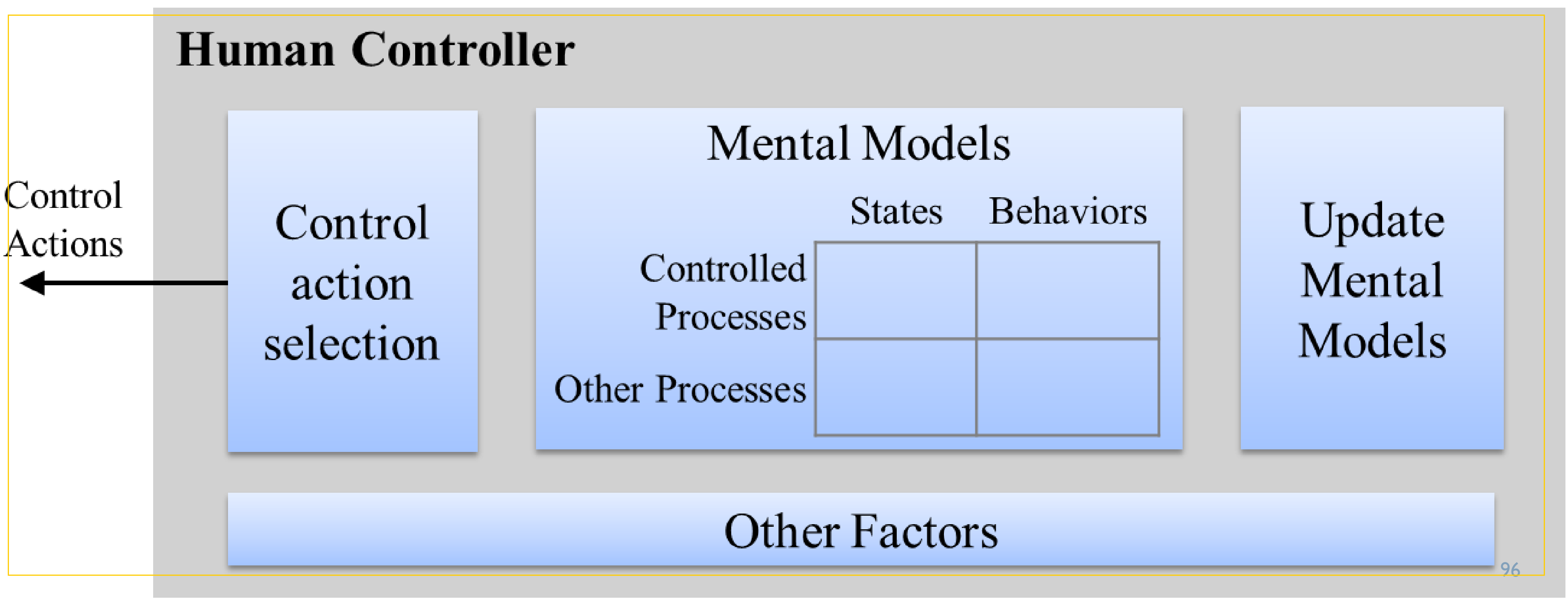


STPA: ENGINEERING FOR HUMANS



STPA: ENGINEERING FOR HUMANS

- Identify UCAs
- Identify Mental Model variables
- Identify Mental Model Flaws
- Identify flaws in Mental Model Updates
- Identify unsafe decisions (Control Action Selections)



Can it work for other systems?

VOLVO CITY SAFETY SYSTEM

From Volvo website:

- City Safety is a support system designed to help the driver avoid low speed collisions when driving in slow-moving, stop-and-go traffic.
- City Safety triggers brief, forceful braking if a low-speed collision is imminent.

VOLVO CITY SAFETY PREVENTING AN ACCIDENT

VOLVO CITY SAFETY PREVENTING AN ACCIDENT

ACCIDENT WITH CITY SAFETY

VOLVO RESPONSE

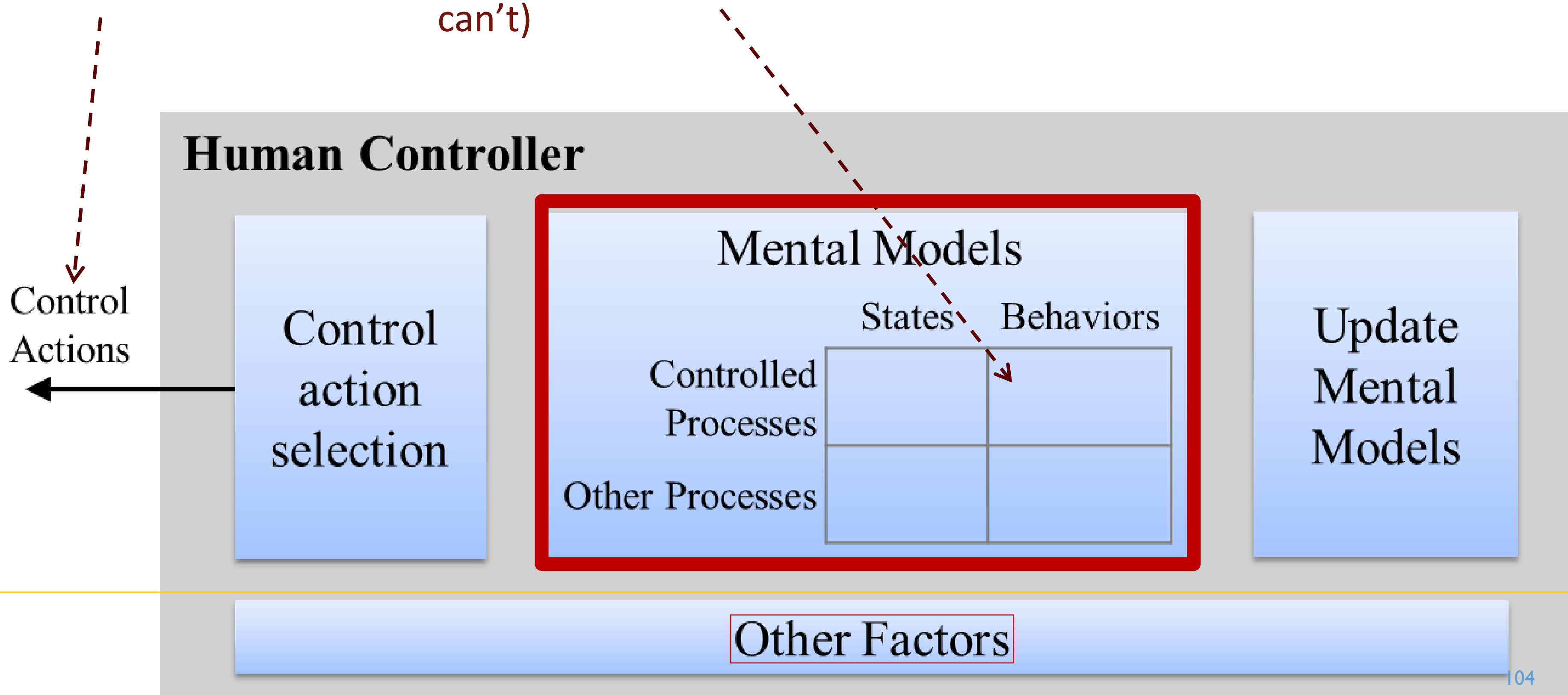
- "The Volvo XC60 comes with City Safety as a standard feature
- "however this does not include the Pedestrian detection functionality … this is sold as a separate package."
- Optional pedestrian detection functionality costs $3,000

Kashmir Hill, "Volvo says horrible 'self-parking car accident' happened because driver didn't have 'pedestrian detection'", Splinter news, May 2015
https://splinternews.com/volvo-says-horrible-self-parking-car-accident-happened-1793847943

STPA: ENGINEERING FOR HUMANS

Driver does not brake for pedestrian (UCA-1)

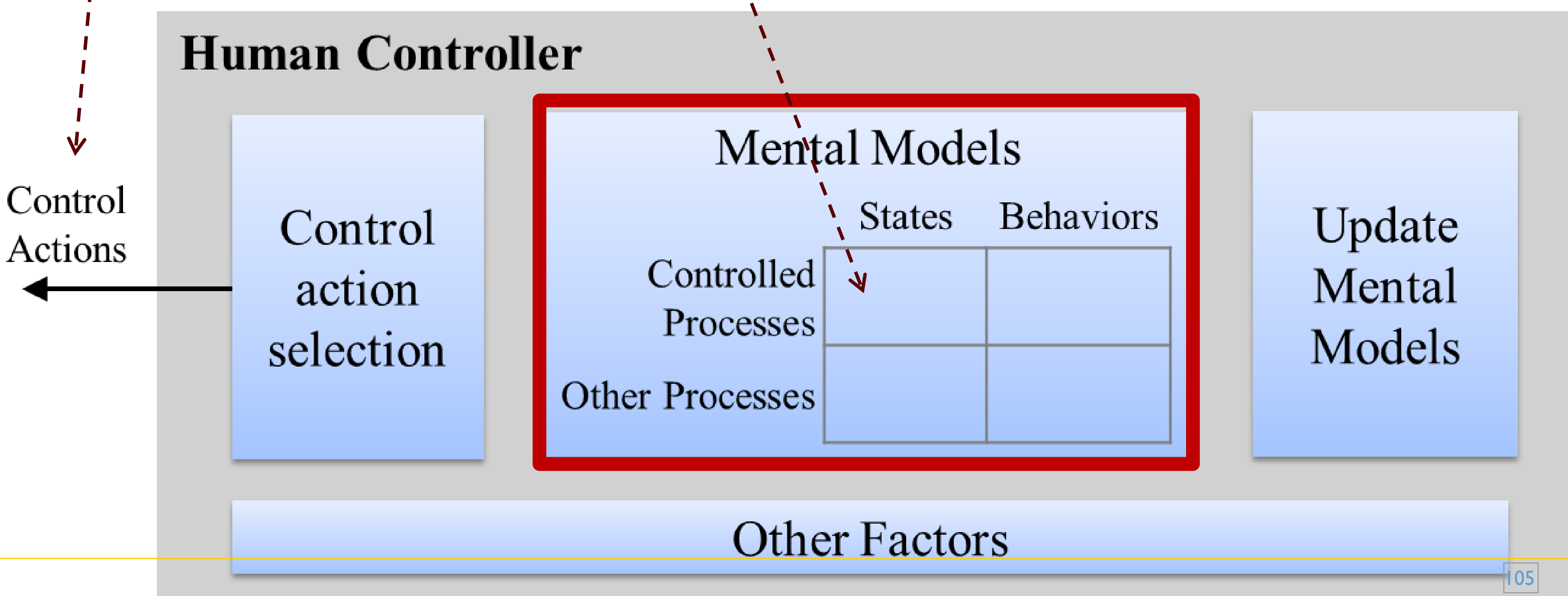
Driver believes City Safety System can automatically brake for pedestrians (it can't)



STPA: ENGINEERING FOR HUMANS

Driver does not brake for pedestrian (UCA-1)

Driver thinks City Safety System is on (it is really off)

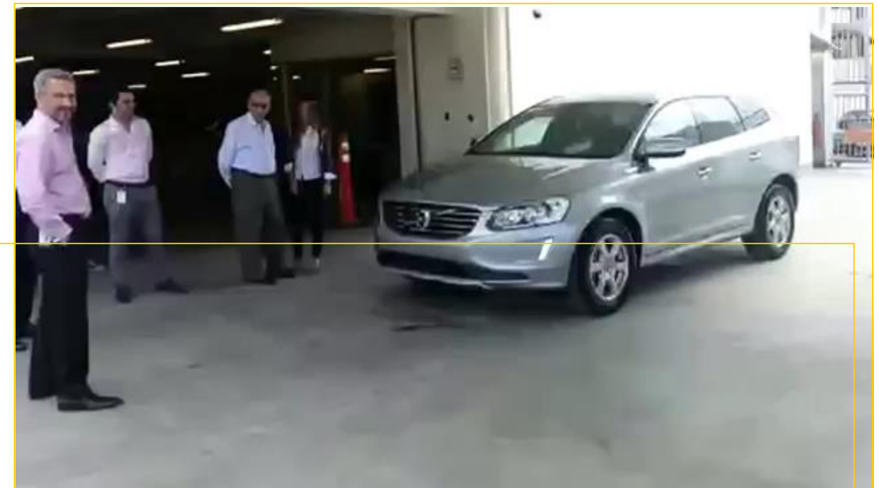


VOLVO RESPONSE

- "The Volvo XC60 comes with City Safety as a standard feature …
- "however this does not include the Pedestrian detection functionality … this is sold as a separate package."
- Optional pedestrian detection functionality costs $3,000
- Even with pedestrian detection, it mostly likely would not have worked because the driver accelerated

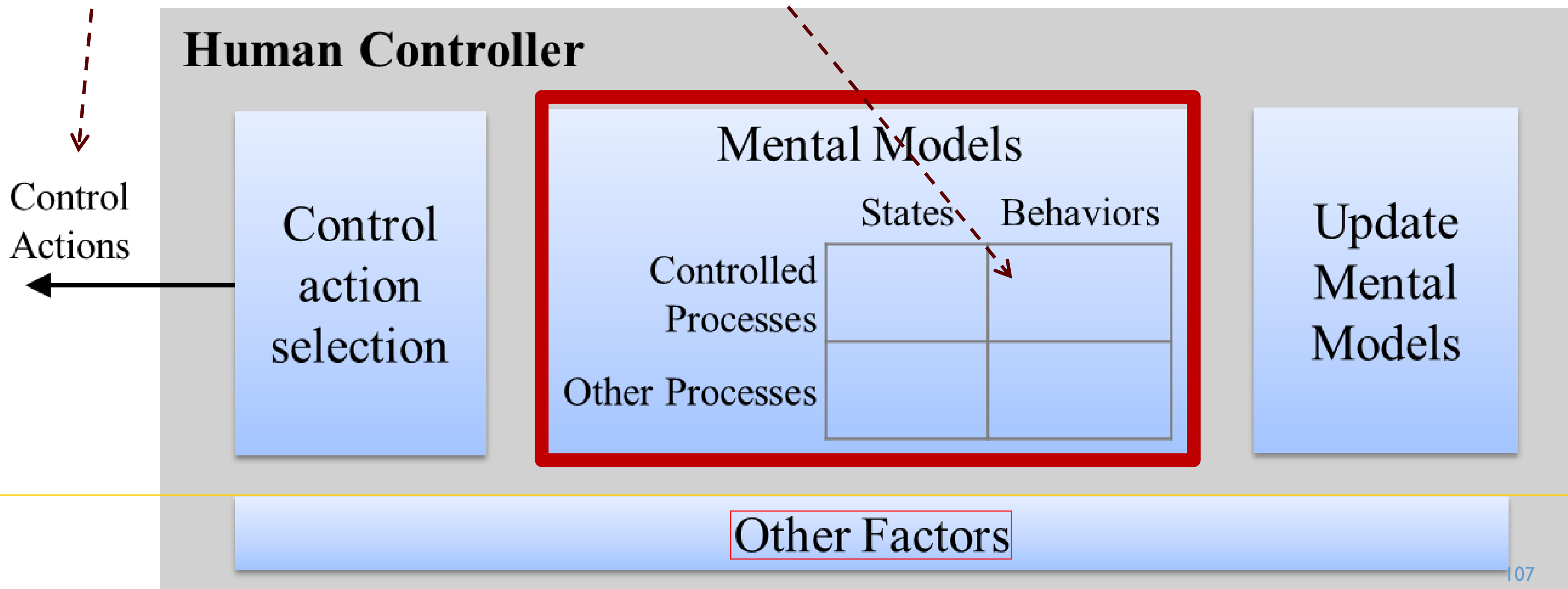
Kashmir Hill, "Volvo says horrible 'self-parking car accident' happened because driver didn't have 'pedestrian detection'", Splinter news, May 2015
https://splinternews.com/volvo-says-horrible-self-parking-car-accident-happened-1793847943

STPA: ENGINEERING FOR HUMANS



Driver does not brake for pedestrian (UCA-1)

Driver thinks City Safety System can intervene during acceleration (it can't)



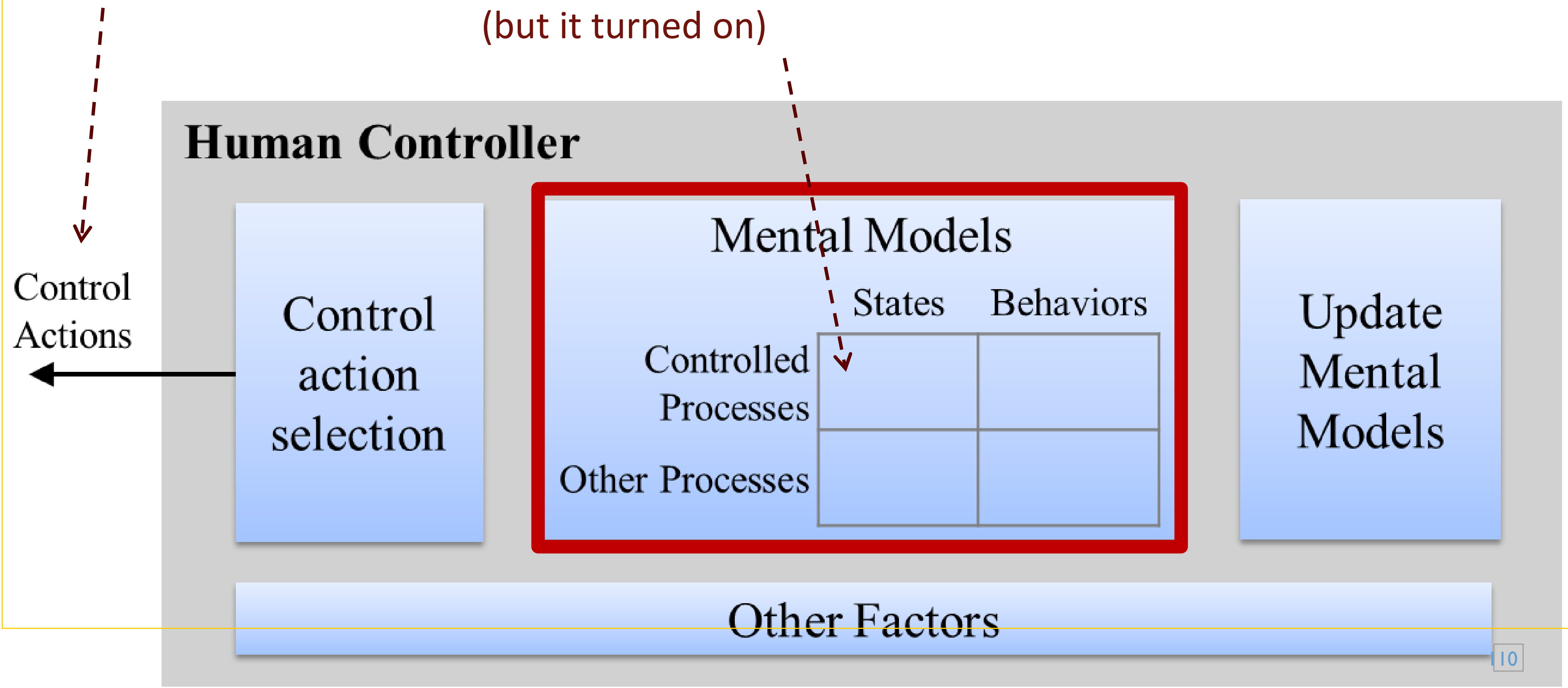
TESLA SUMMON

This feature will park Model S while the driver is outside the vehicle. Please note that the vehicle may not detect certain obstacles, including those that are very narrow (e.g., bikes), lower than the fascia, or hanging from the ceiling. As such, Summon requires that you continually monitor your vehicle's movement and surroundings while it is in progress and that you remain prepared to stop the vehicle at any time using your key fob or mobile app or by pressing any door handle.

STPA: ENGINEERING FOR HUMANS

Driver does not provide manual override when obstacle in path (UCA-1)

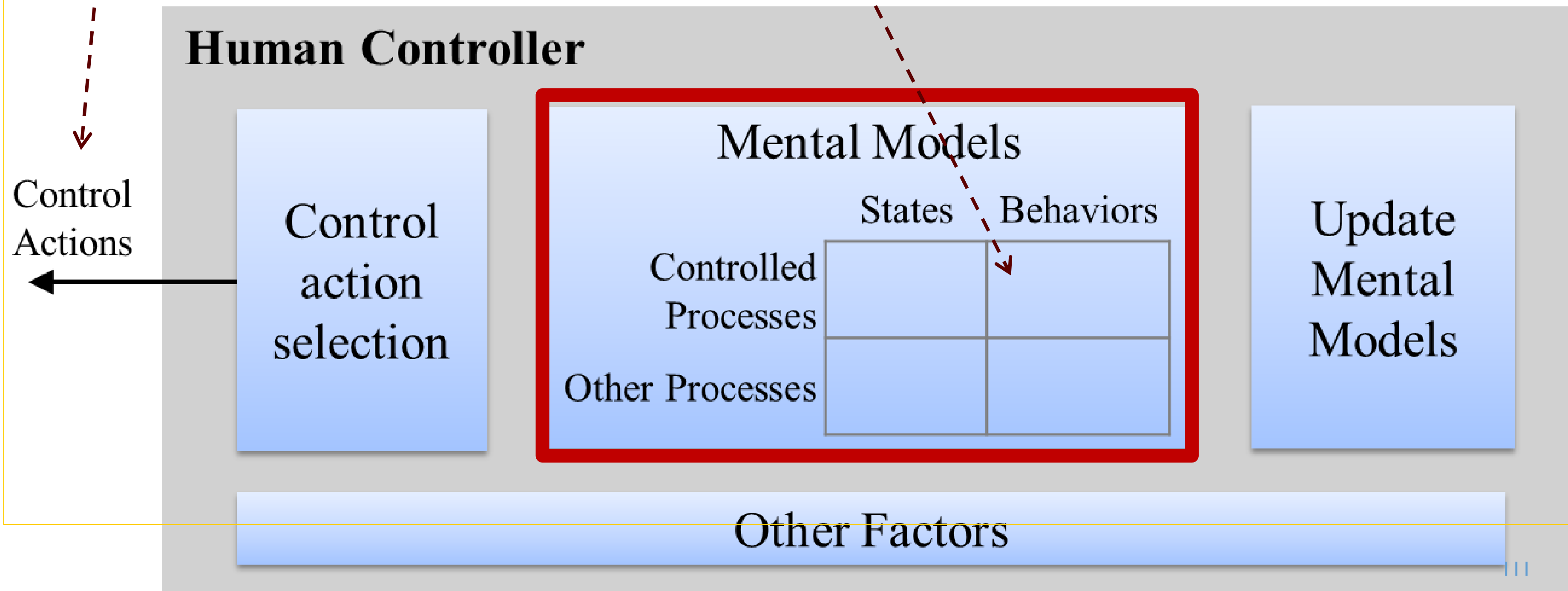
Driver thinks Summon is off (but it turned on)



STPA: ENGINEERING FOR HUMANS

Driver does not provide manual override when obstacle in path (UCA-1)

Driver thinks Summon will detect raised obstacles, apply brakes (it won't)



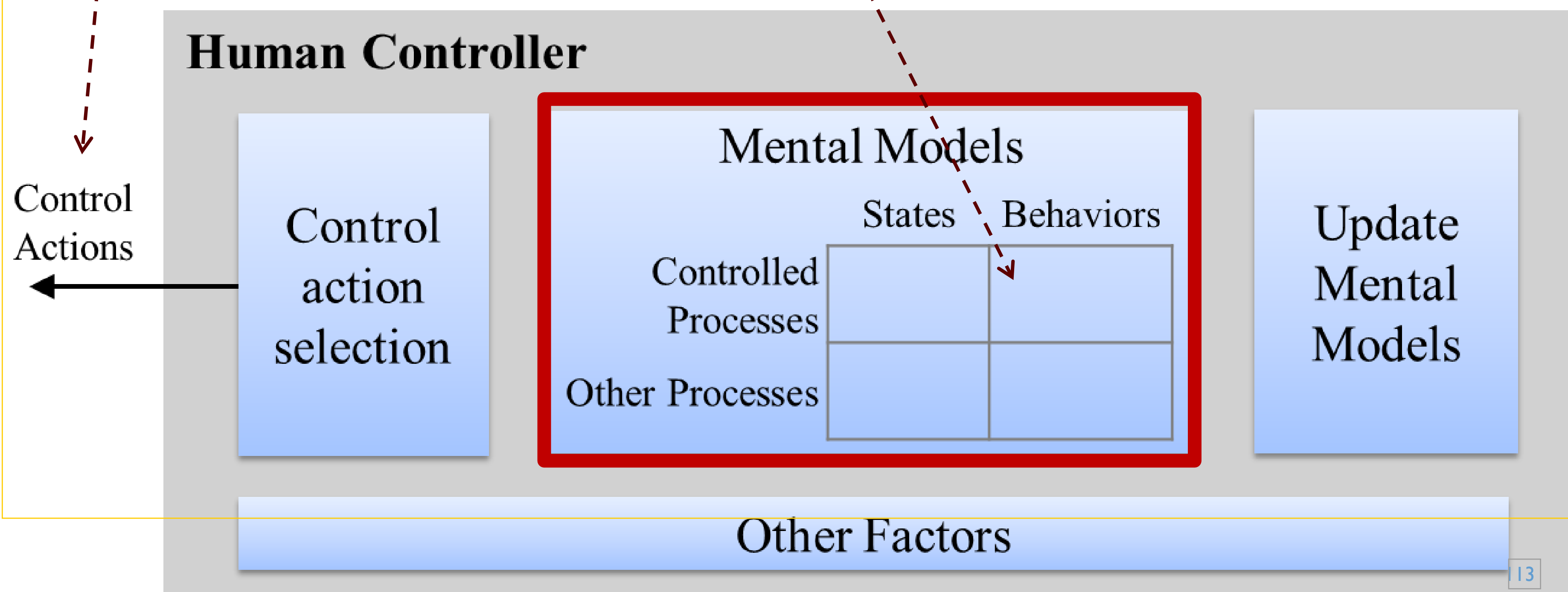
MONOSTABLE SHIFTER DESIGN

Audi A8: Similar design, but SW will automatically activate electronic park brake if driver exits

STPA: ENGINEERING FOR HUMANS

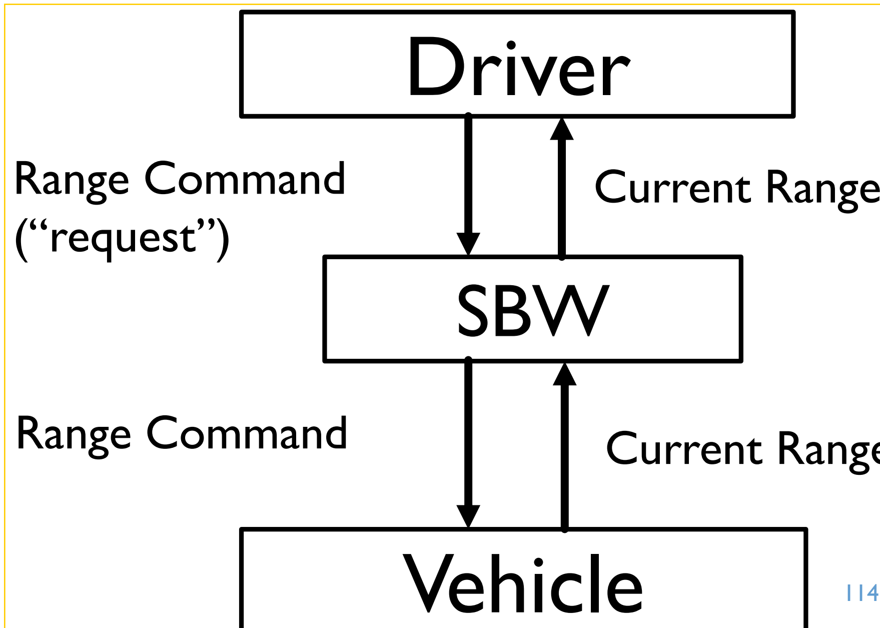
Driver does not provide Park cmd before exiting vehicle (UCA-1)

Driver believes vehicle will automatically shift to park (it won't)



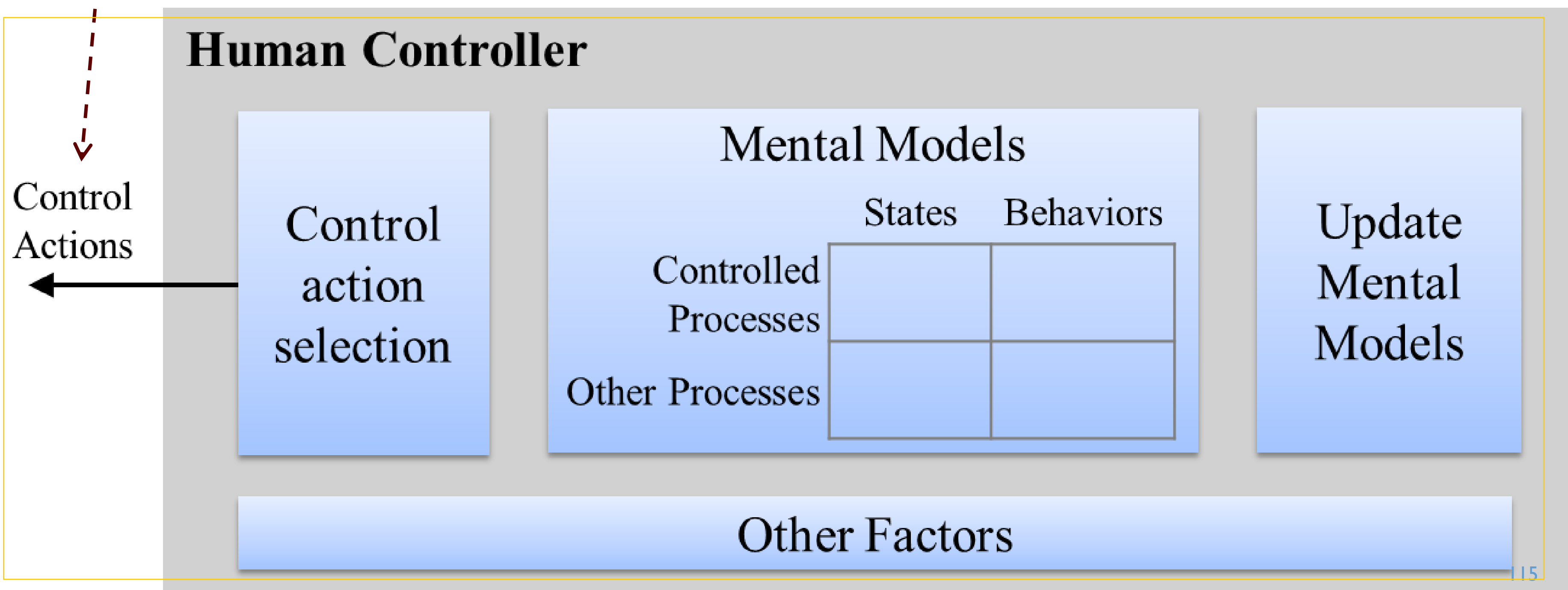
Range =

- Park
- Reverse
- Neutral
- Drive
- Etc.



STPA: ENGINEERING FOR HUMANS

Driver exits vehicle when vehicle is not in park (UCA-1)



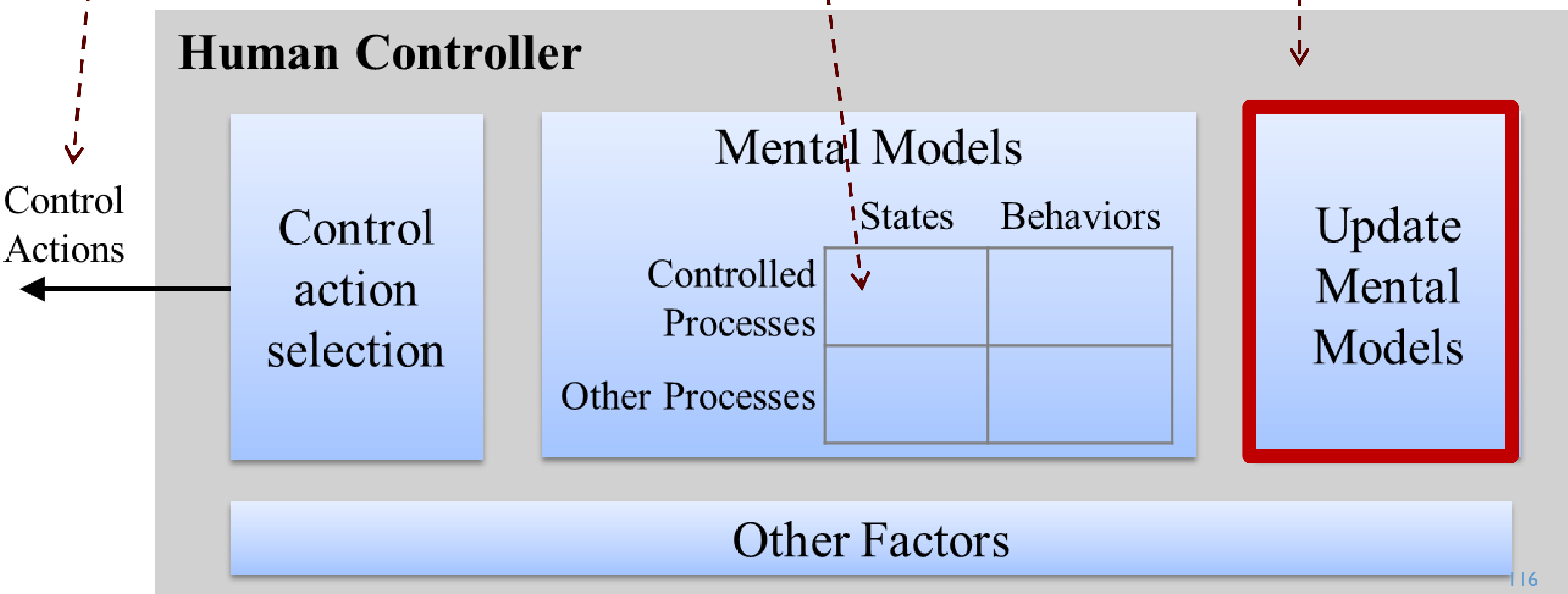
STPA: ENGINEERING FOR HUMANS

Driver exits vehicle when vehicle is not in park (UCA-1)

Driver incorrectly believes vehicle is in Park (MM-1)

Consider:

1. Automatic mode changes
2. Previous cmds ignored
3. Phases of operation
4. Etc.

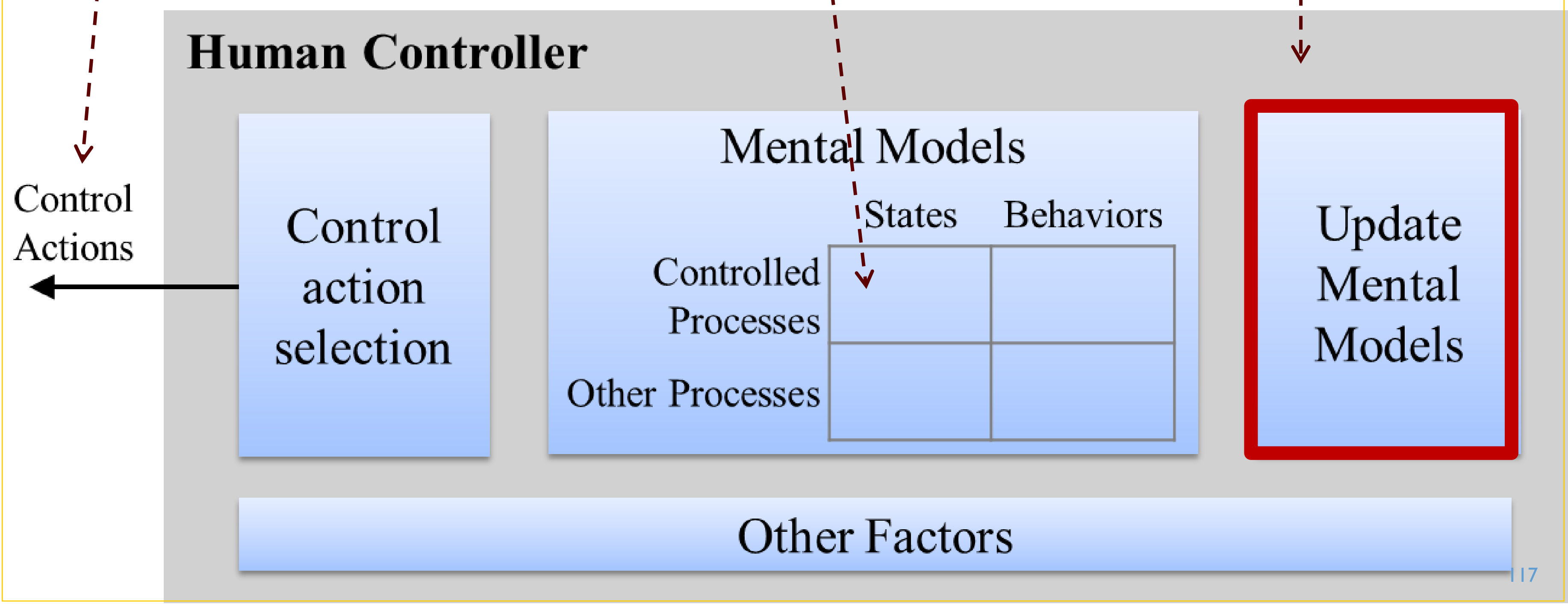


STPA: ENGINEERING FOR HUMANS

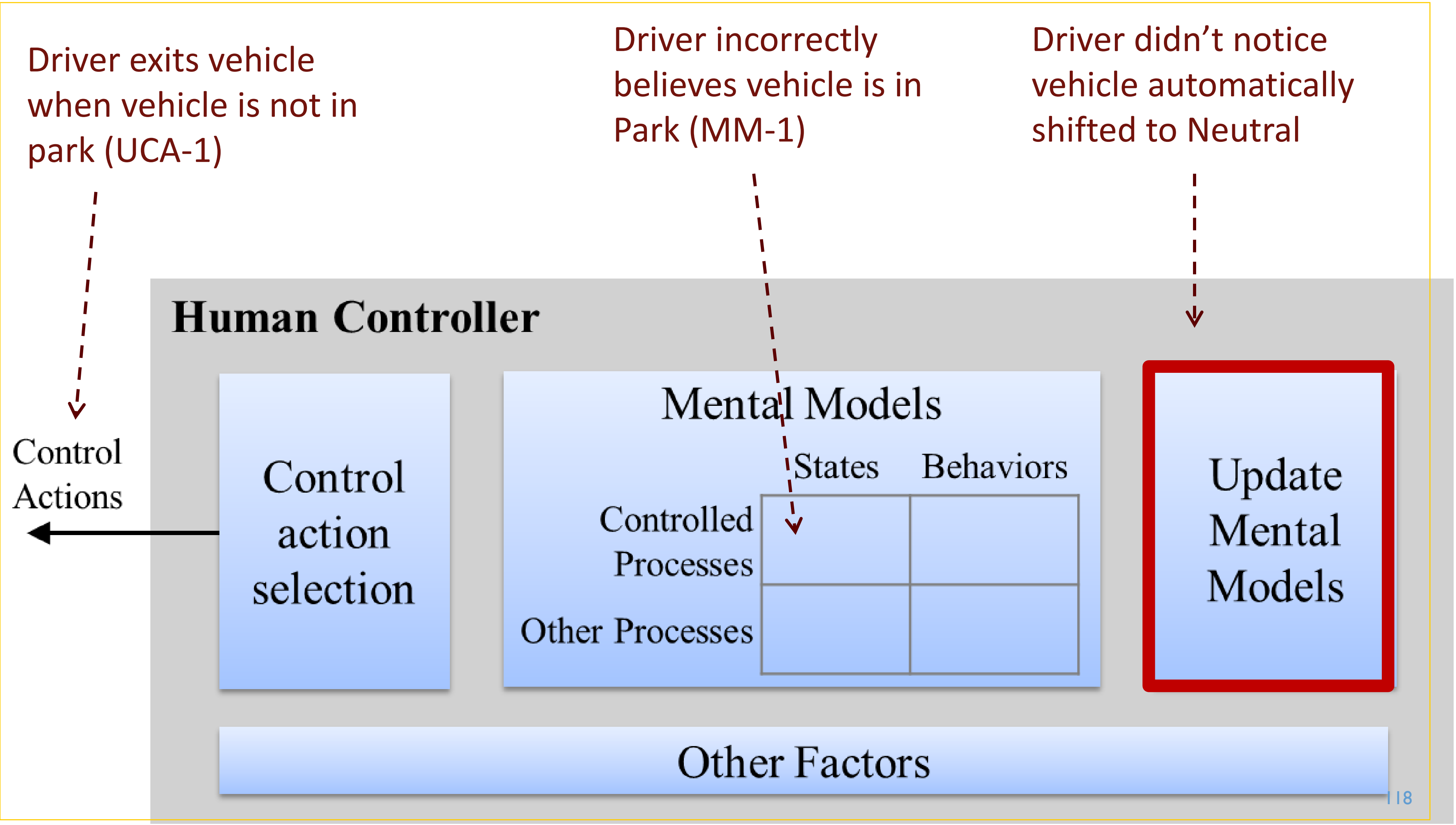
Driver exits vehicle when vehicle is not in park (UCA-1)

Driver incorrectly believes vehicle is in Park (MM-1)

Driver didn't notice vehicle ignored previous Park command



STPA: ENGINEERING FOR HUMANS

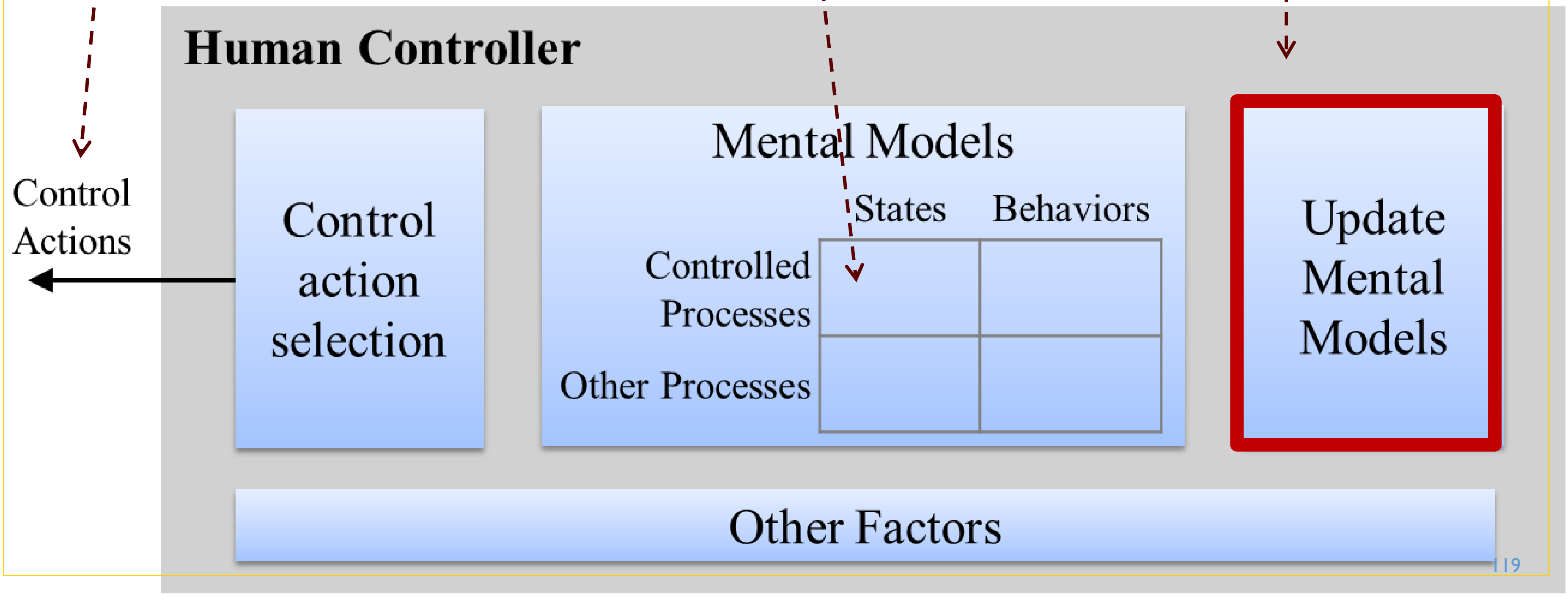


STPA: ENGINEERING FOR HUMANS

Driver accelerates when vehicle is not in appropriate range (e.g. reverse instead of drive)

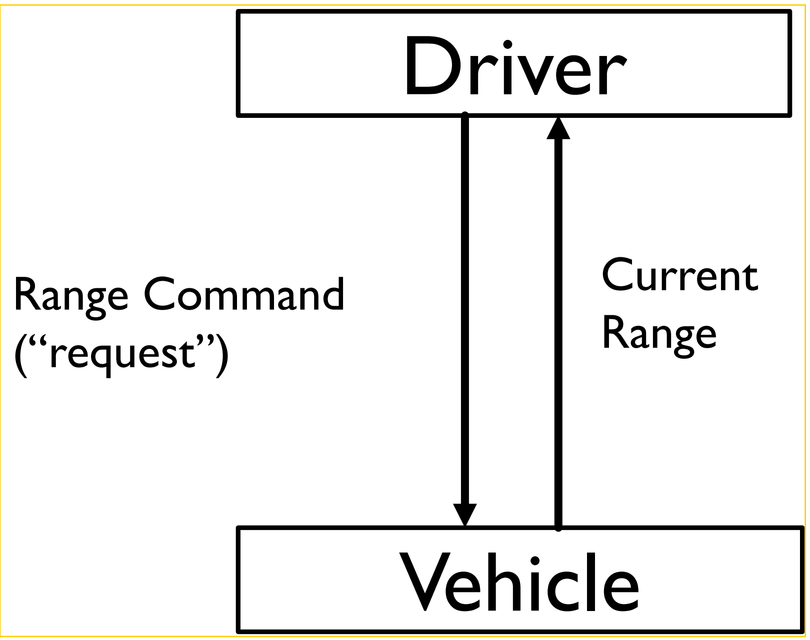
Driver incorrectly believes vehicle is in Drive (MM-1)

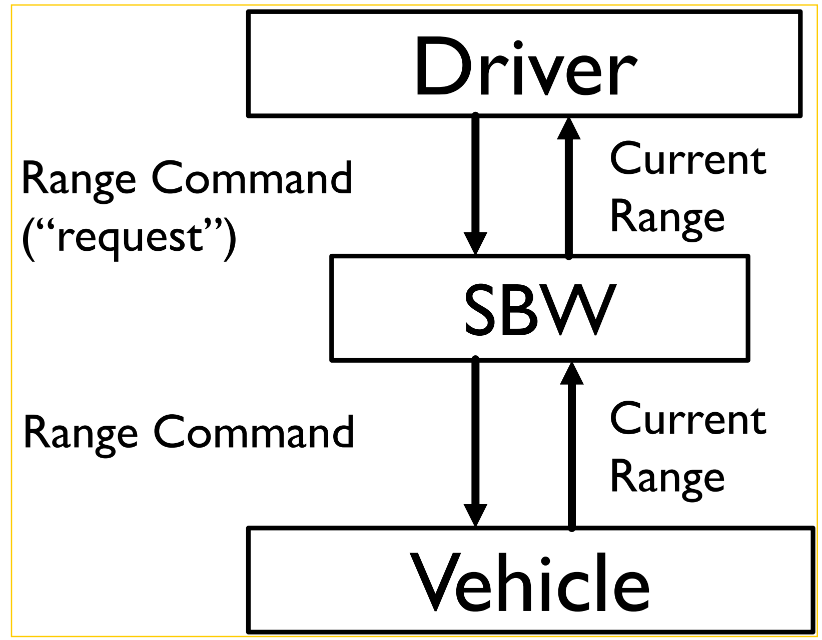
Driver didn't notice vehicle ignored cmd to shift to Drive (stayed in reverse)

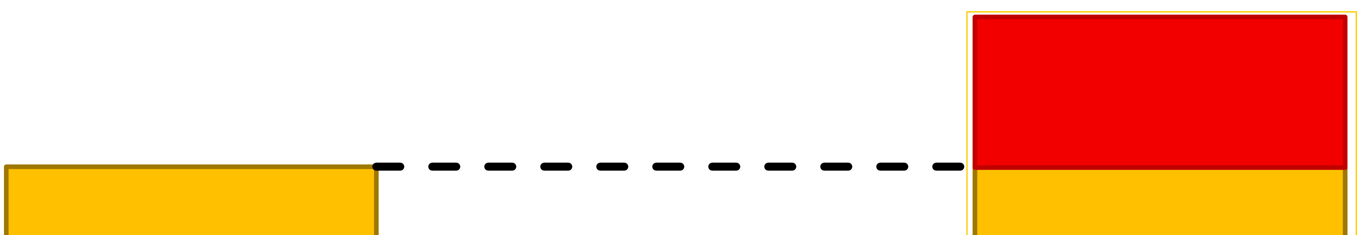


Old System

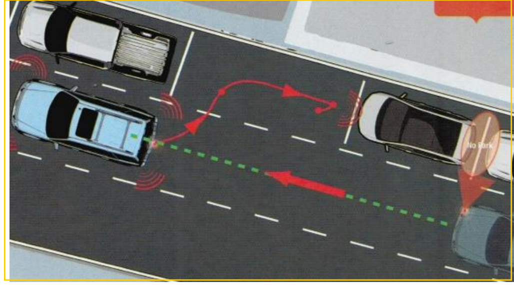
New System







AUTOMATED PARKING



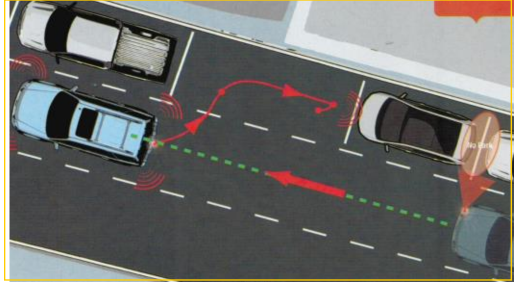
Features of each system considered for this analysis:

	Level 0* No Driving Automation	Level 1 "Driver Assistance"	Level 2a "Partial Automation"	Level 2b "Partial Automation"	Level 3 "Conditional Automation"
Steering	-	✓	✓	✓	✓
Braking	-	-	✓	✓	✓
Shifting and Acceleration	-	-	-	✓	✓
Object and Event Detection and Response	-	-	-	-	✓

*System numbering is consistent with SAE definitions for levels of automation, while "a" and "b" indicate different implementations which are classified within the same SAE level.

Analysis reuse

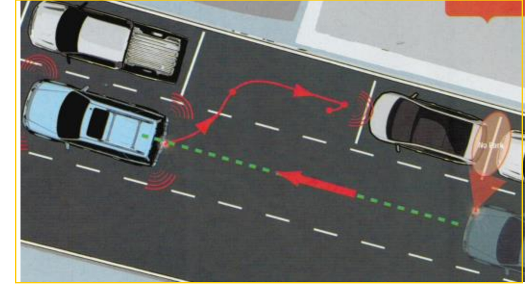
M. France, 2016



AUTOMATED PARKING

	Level 1 "Driver Assistance"	Level 2a "Partial Automation"	Level 2b "Partial Automation"	Level 3 "Conditional Automation"
Driver UCAs	42	41	38	44
APA Computer UCAs	5	13	28	28
Total				

AUTOMATED PARKING



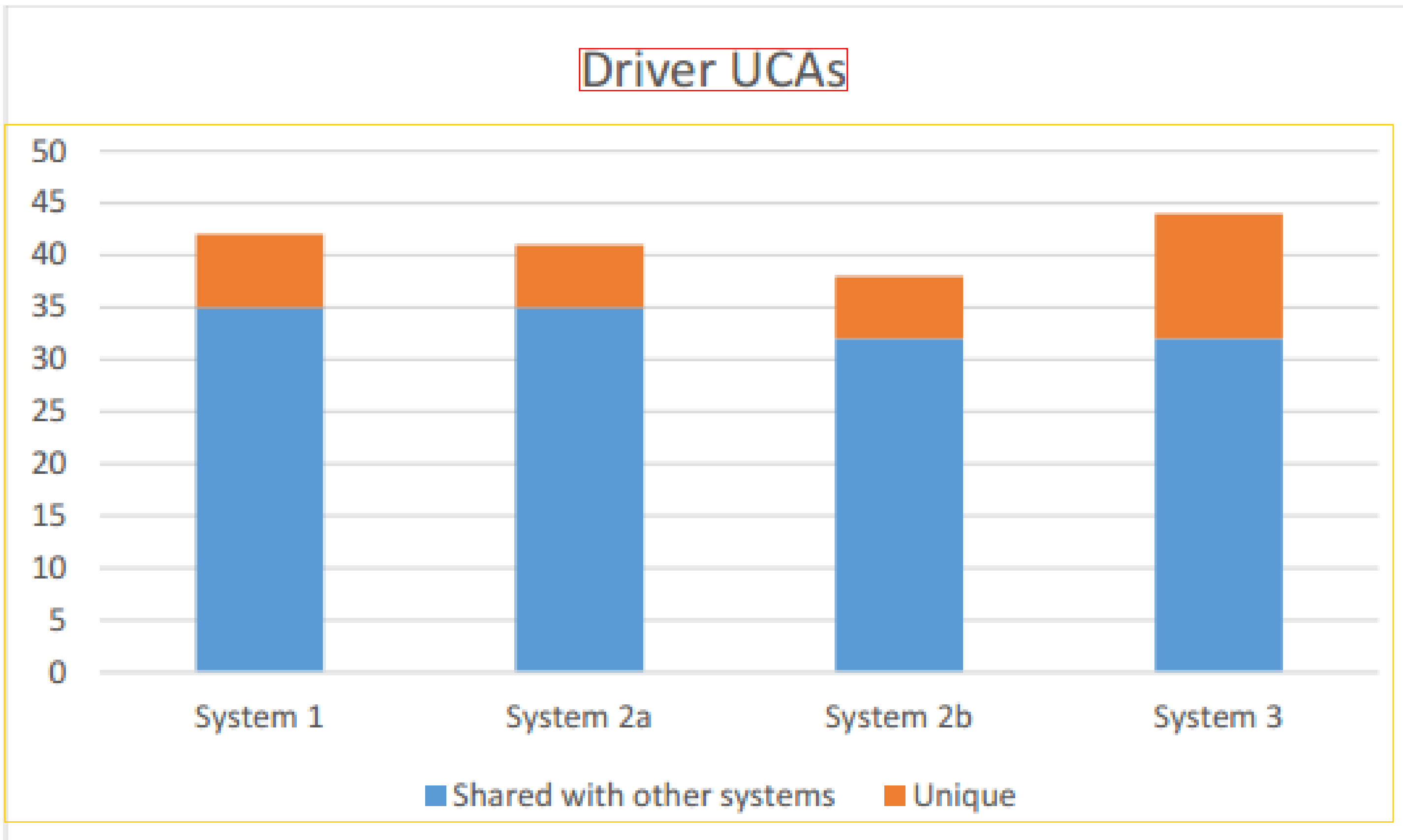
	Level 1 "Driver Assistance"	Level 2a "Partial Automation"	Level 2b "Partial Automation"	Level 3 "Conditional Automation"
Driver UCAs	42	41	38	44
APA Computer UCAs	5	13	28	28
Total	47	54	66	72

Driver UCAs: 35 in common, 32 in common, 30 in common

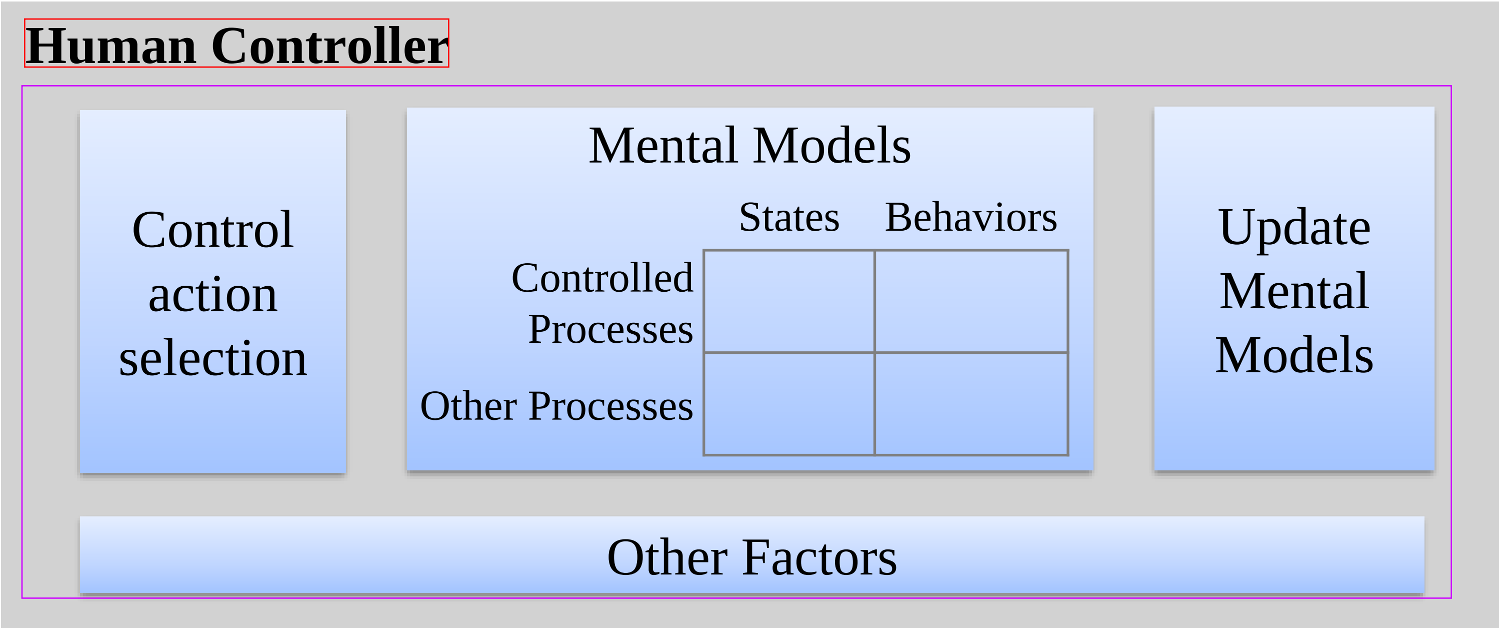
APA Computer UCAs: 5 in common, 28 in common, 13 in common

Total: 40 in common, 60 in common, 43 in common

	Level 1	Level 2a	Level 2b	Level 3
Driver UCAs	42	41	38	44
APA Computer UCAs	5	13	28	28
Total	47	54	66	72



CONCLUSIONS



New human engineering process strengths:

- Easy for engineers to learn, use
- Drive engineering requirements and concepts from the start
- Can be used earlier in design process than detailed simulations or prototypes
- Successful in industry, adoption