

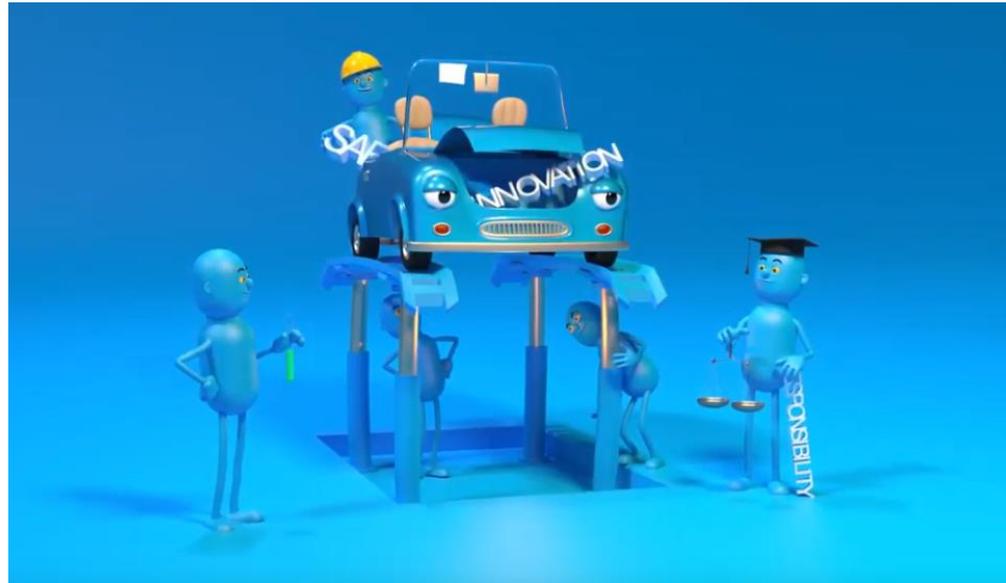
# Meaningful Human Control for Automated Driving Systems

Dr. Ir. S.C. Calvert – AiTech Agora – December 4, 2019

**Meaningful Human Control**  
over automated driving systems



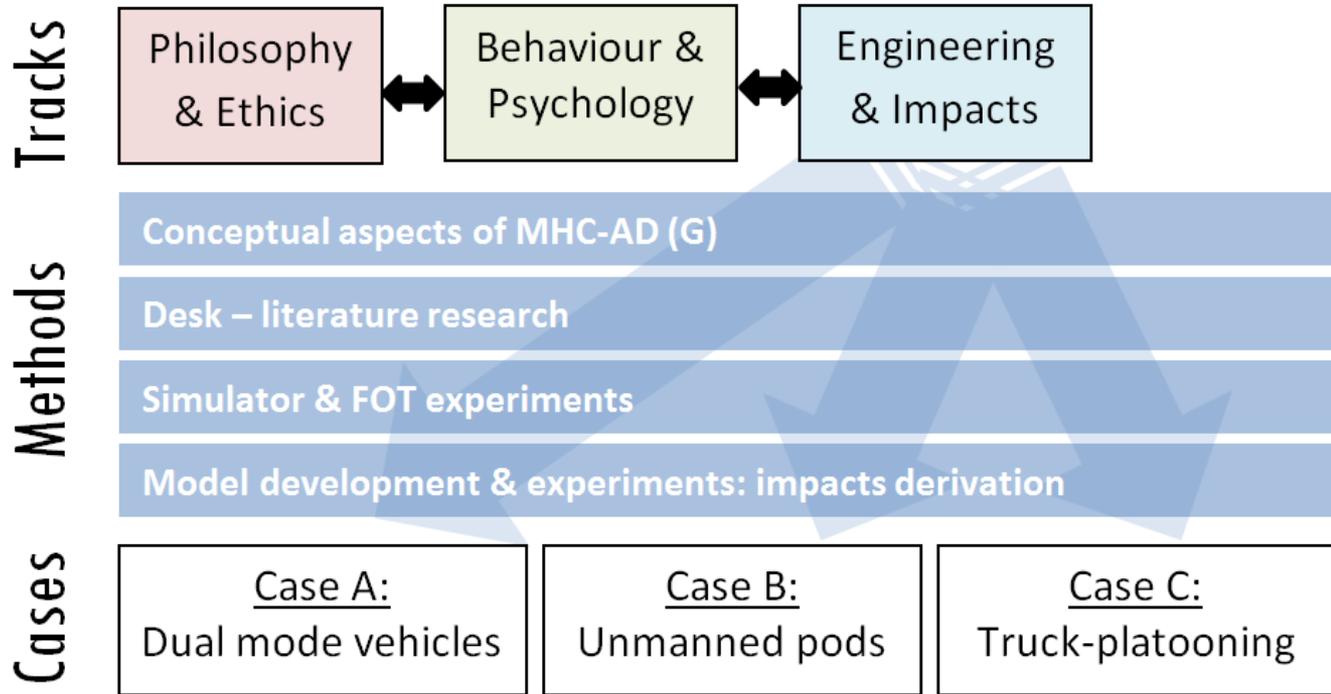
# Operationalisation of MHC



# Outline

- The project: MHC-ADS
- MHC and it's conditions
- Proximity scales
- Operationalisation
- Case studies

# MHC-ADS



# MHC conditions

NEW conditions for (*meaningful*) human control over automated systems

Tracking



The system (human operators, operated devices, infrastructures...) should be able to co-vary its behavior with the relevant reasons of the relevant human agent(s) for carrying out X or omitting X

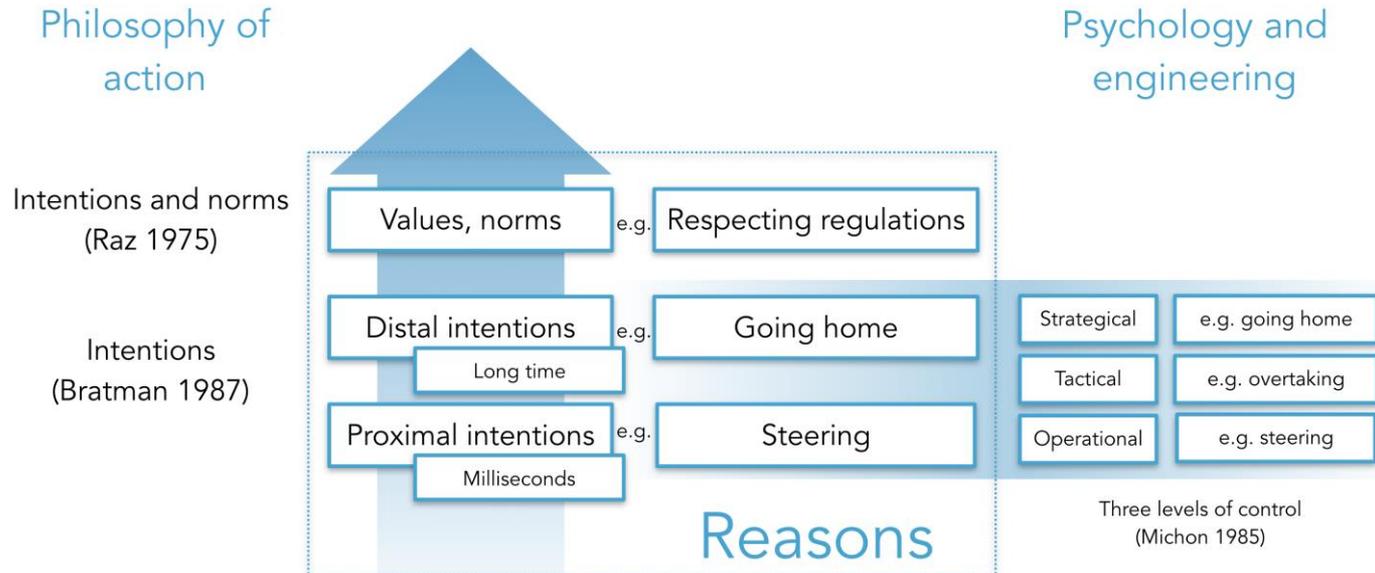
Tracing



There is at least one human agent in the system design history or use context who can appreciate the capabilities of the system and her own role as target of potential moral consequences for the system's behaviour

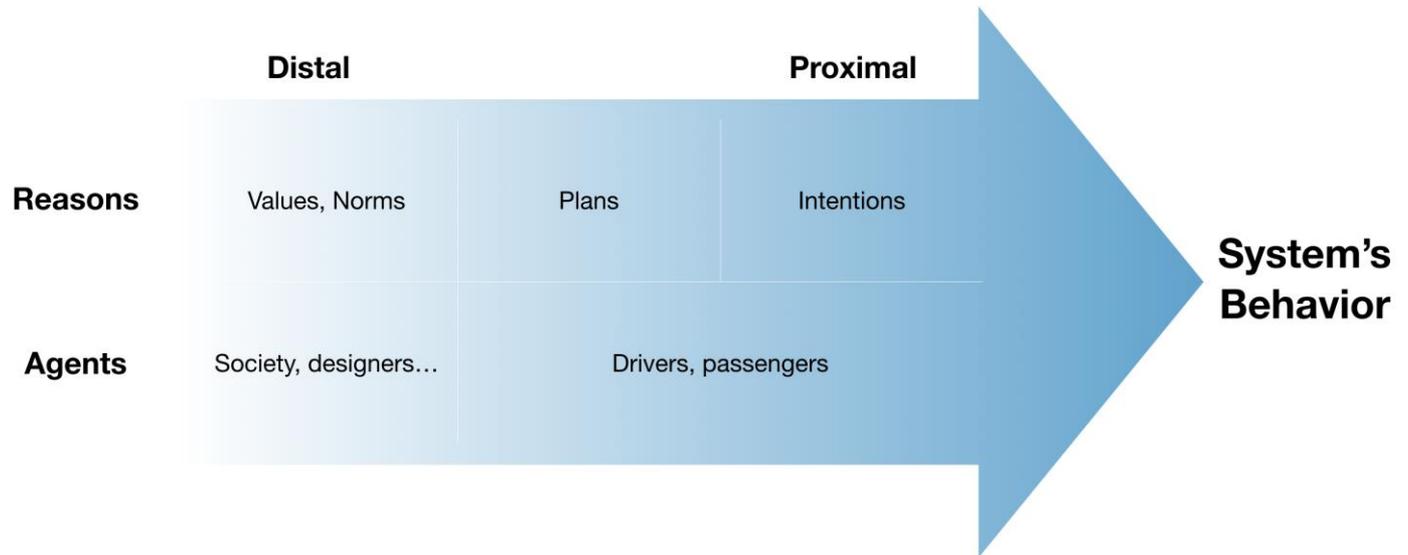
# Control: Connecting philosophy & engineering

Expanding a classic theory of control with intuitions from philosophy of action



G. Mecacci & F. Santoni de Sio (2019). "Meaningful human control as reason-responsiveness: the case of dual-mode vehicles". *Ethics and Information Technology*.

# Proximity scale of reasons



# Challenges towards application

- Very abstract concept  
(in contrast to physical operational control)
- Complex ethical and behavioural interactions
- Collectively subjective
  
- Operationalisation allows application in practice
  1. Definition of core components
  2. (Simulation modelling framework)
  3. Dissection of MHC
  4. Quantification of MHC conditions



# Core components

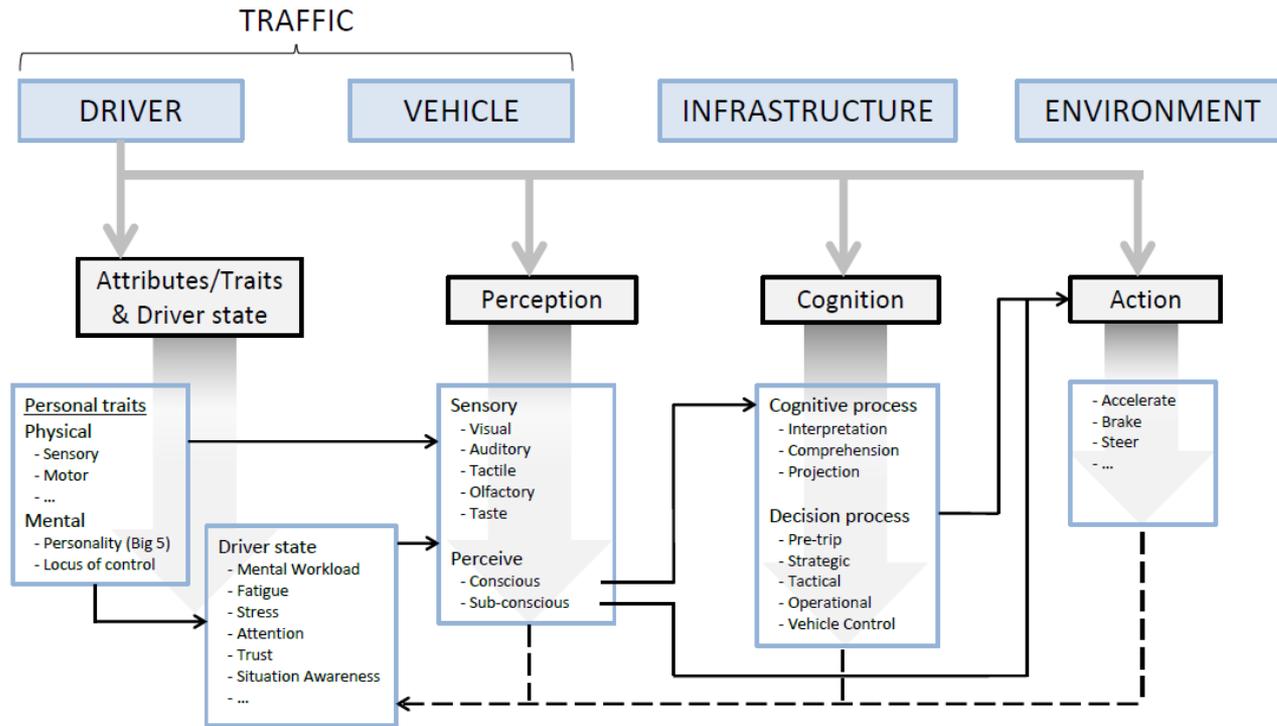


Figure 1. Driver core components of control in ADS

S. C. Calvert, D. D. Heikoop, G. Mecacci & B van Arem (2019), "A human centric framework for the analysis of automated driving systems based on Meaningful Human Control". *Theoretical Issues in Ergonomics Science*.

# Core components

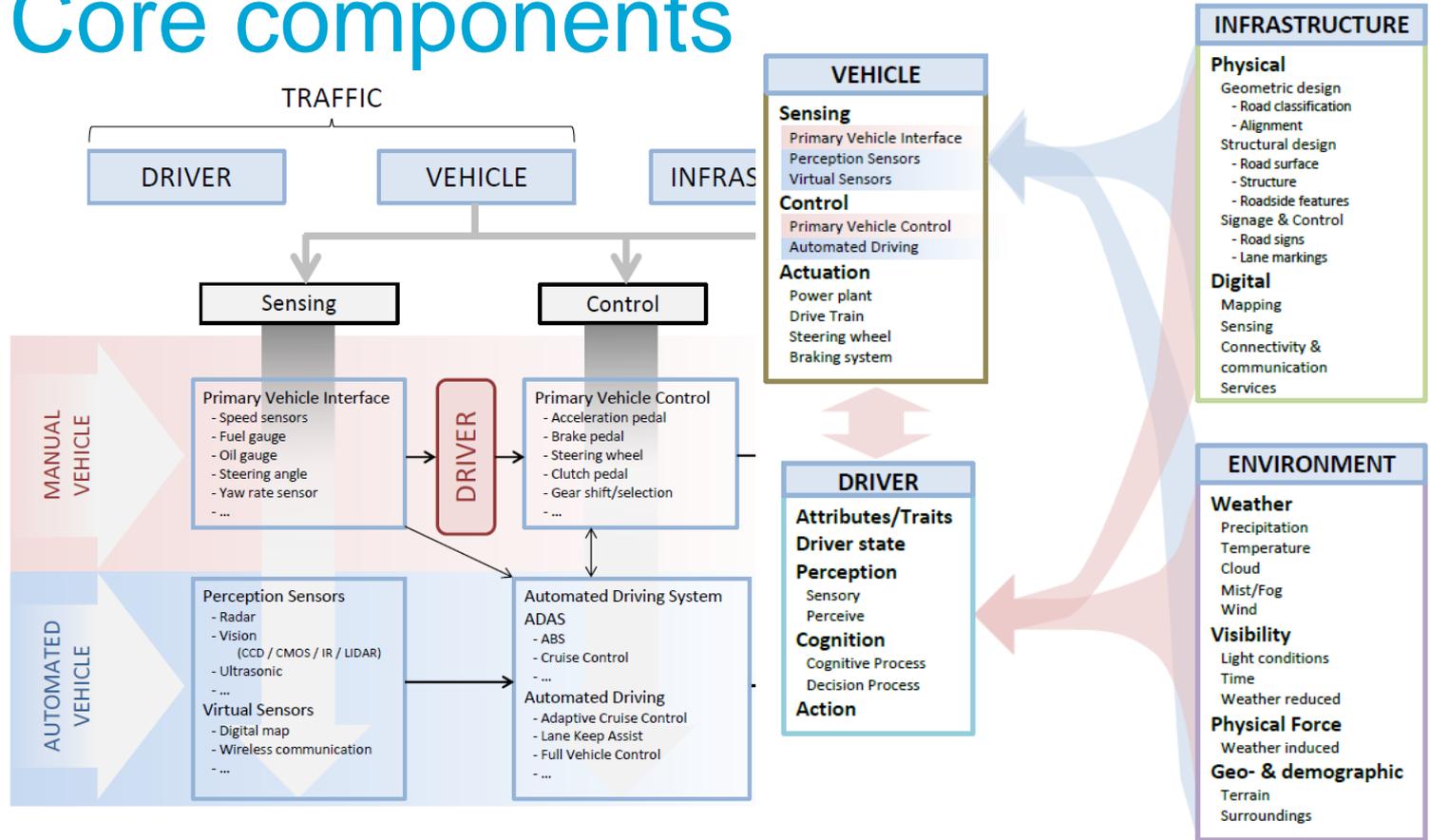


Figure 2. Vehicle core components of control in ADS (red is manual control, blue is automated control),

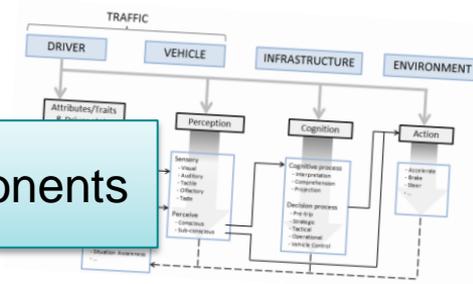
S. C. Calvert, D. D. Heikoop, G. Mecacci & B van Arem (2019), "A human centric framework for the analysis of automated driving systems based on Meaningful Human Control". *Theoretical Issues in Ergonomics Science*.

# Cascade evaluation

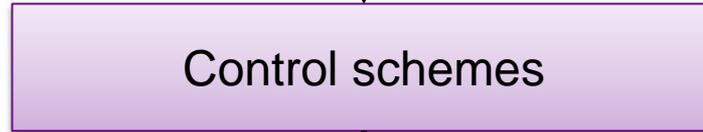
Tracing		Score (0-5)	Critical score
Is operational control exerted by...	Driver	1	5
	ADCS	5	
	Other, namely...	0	
Is there a human agent involved irt...	Driver	5	4
	ADCS	4	
	Other, namely...	NA	
Does or should this person understand the system in the sense that: iii) They have a propositional knowledge of the system's functions (know-that) They have the right capacities to exercise a control task (know-how)	Driver	min(3,2)	2
	ADCS designer	min(5,0)	
	Other, namely...	NA	
Does or should this person understand their own moral responsibility for the consequences of the actions of a system	Driver	4	2
	ADCS designer	2	
	Other, namely...	NA	

# Operationalisation of MHC

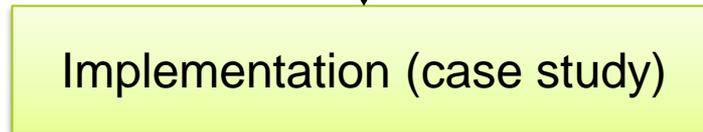
Tracking  
Tracing



Simplified!

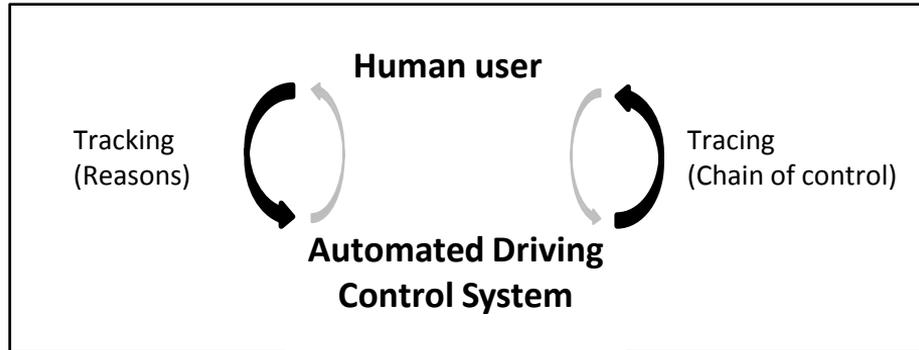


Based on a learning system!

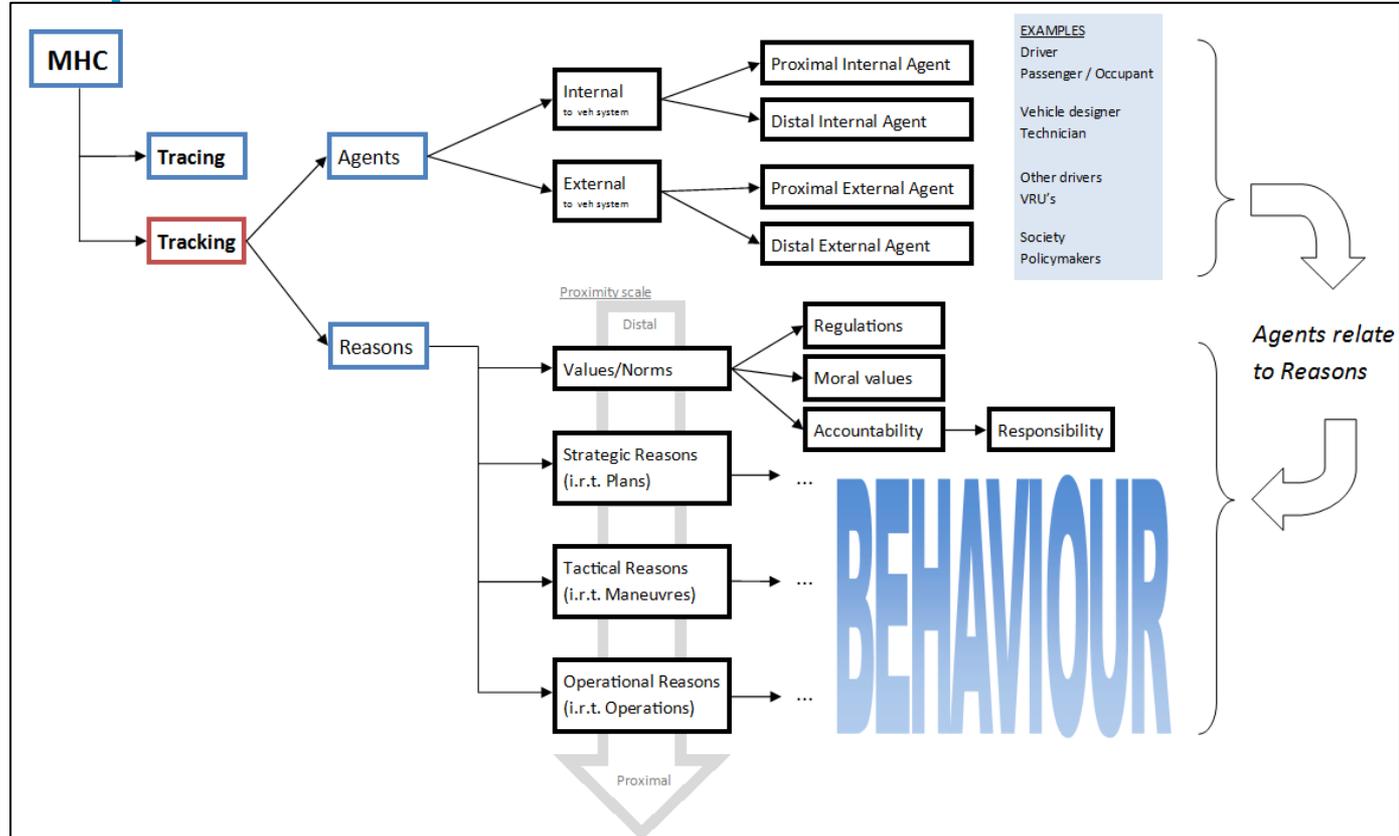


# Operationalisation of MHC

- General connection with MHC

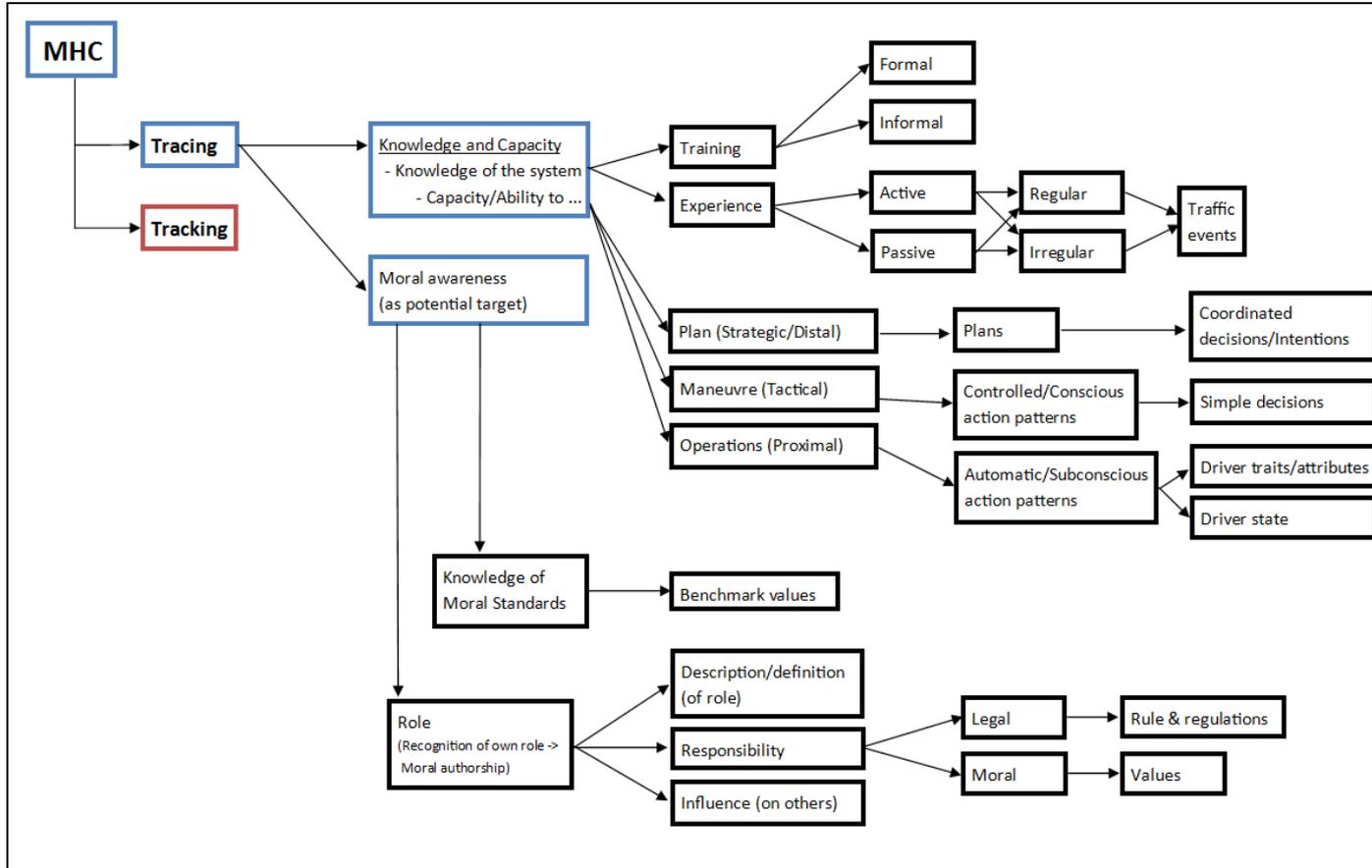


# Operationalisation of MHC

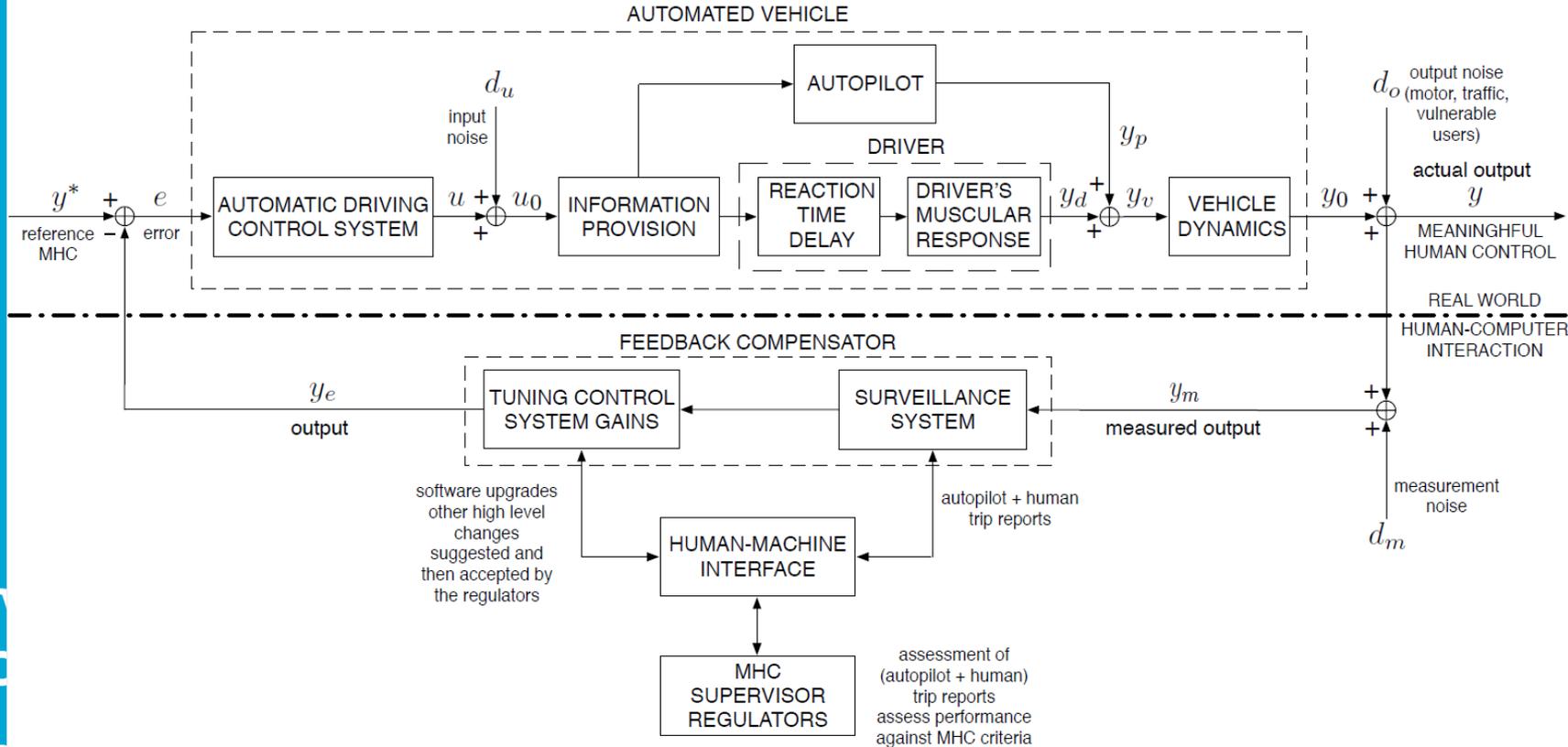


S. C. Calvert, K. Ampountolas & G. Mecacci (2020). "A mathematical system control description of Cooperative and Automated Driving in mixed urban traffic with Meaningful Human Control". Forthcoming

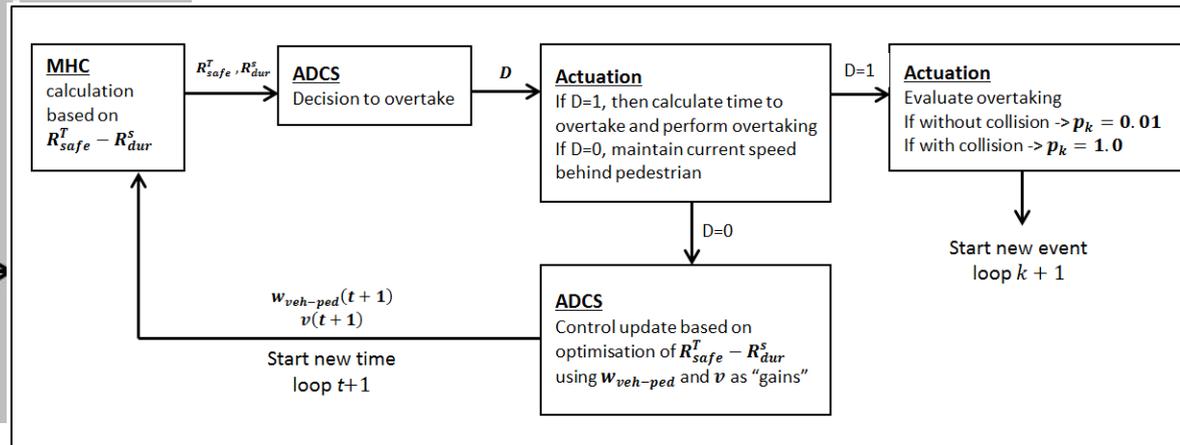
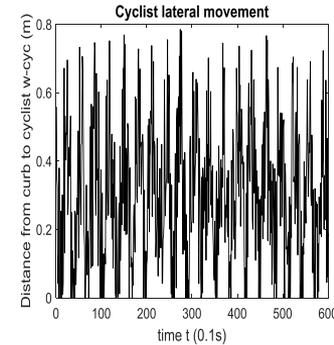
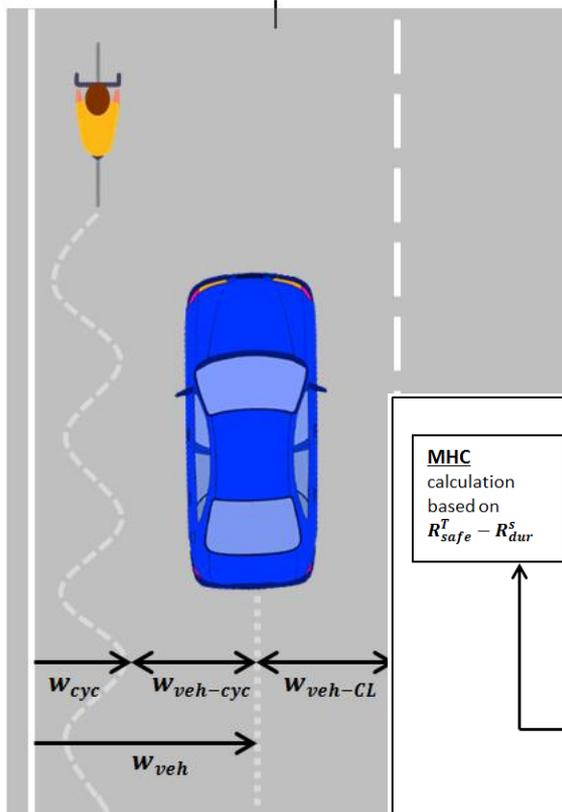
# Operationalisation of MHC



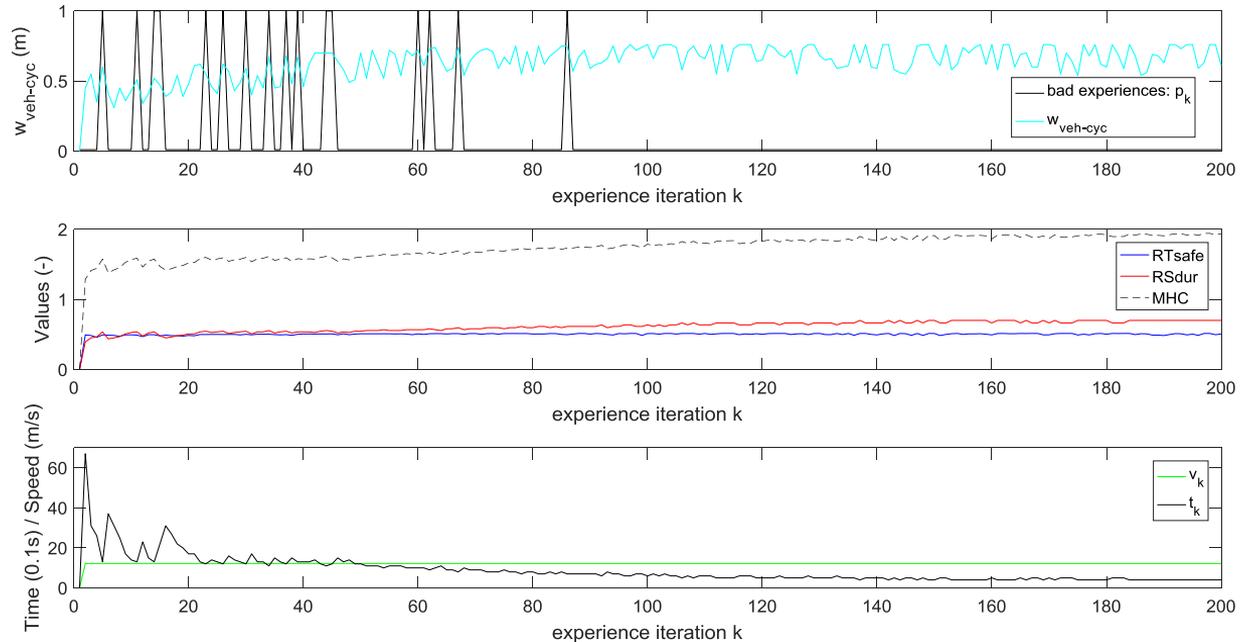
# Operationalisation of MHC



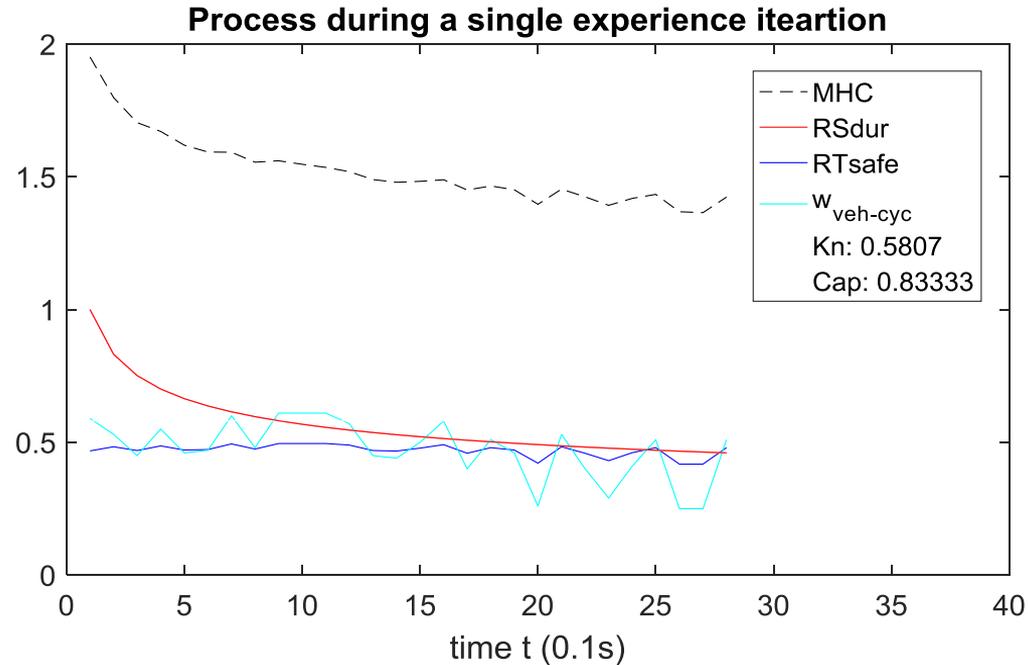
# Operationalisation of MHC



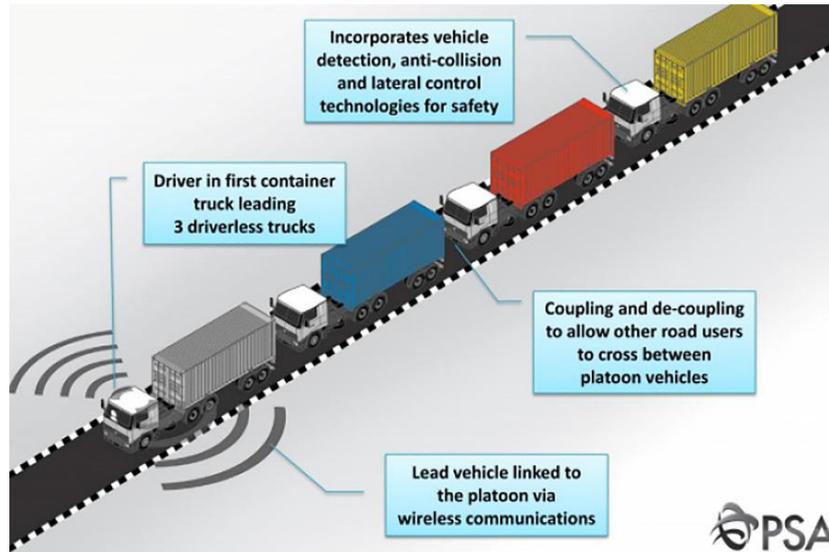
# Operationalisation of MHC



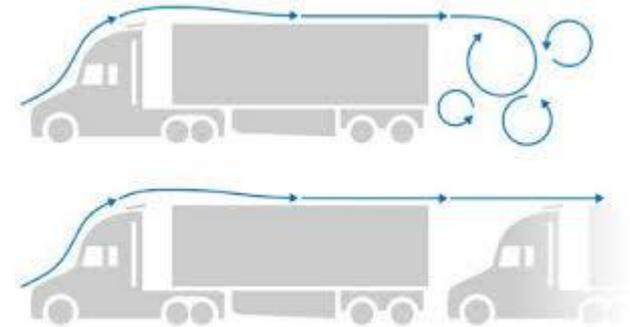
# Operationalisation of MHC



# Case study: Truck platooning



- Cooperative
- (Partially) automated

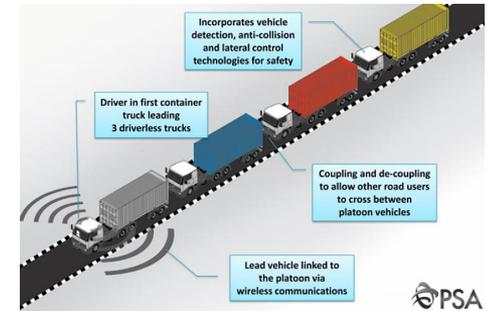


Source: NREL

- Fuel savings
- Lower emissions
- Lower labour costs?
- Traffic efficiency?

# TP control (physical & practical)

- Operational control of platoon:
  - Front driver & (C)ACC-system
  - Human control: Front driver
- Operational control of following trucks:
  - (C)ACC-system
  - Following driver: monitors truck and system
  - Human control: Front driver !?
- Responsibility:
  - Lies with each individual driver (in principle!)
- ‘Effective’ control is potentially confused or incomplete!



# TP control: example

- Situation:
  - TP approaches workzone
  - CACC-system doesn't detect or makes takeover request
  - Front driver is distracted and reacts late
  - Front driver disengages the system by performing an emergency manoeuvre without crashing
  - The following truck drivers cannot react in time and collide with workzone and each other
- Problem
  - The following drivers are held responsible because they were meant to monitor their own situation of their own trucks



# Applying MHC to examples



- Tracking ( 'system' compliance to...):
  - Yes, platoon is en-route and compliant
  - Although in an instance, the system cannot perform an emergency manoeuvre, but that is situational rather than by intention.
- Tracing (actor capable to control system...):
  - Front driver: **Not met**: Performed delayed emergency manoeuvre – also puts too great demands on system
  - Following drivers: **Not met**: unrealistic transition of control demand
  - ADS designer: **No MHC**, outside ODD (hence TO-request made)
- Both system design and driver performance translate to a lack of MHC (even before an accident) for the case

# TP control and challenges

- Operational control and responsibility not aligned
- Misbalance in:
  - Operational control vs ‘effective human’ control
  - Behavioural capabilities of drivers (cognitive)
  - Ethical issues (demands, accountability and responsibility)

# Operationalisation of MHC

Why is operationalisation important?

- Makes abstract concept **applicable** in practice
- Demonstrates ways that MHC can be considered in vehicle **design**
- Demonstrates an approach to **evaluate** the extent of MHC
- Demonstrates potential **policy** influence on MHC