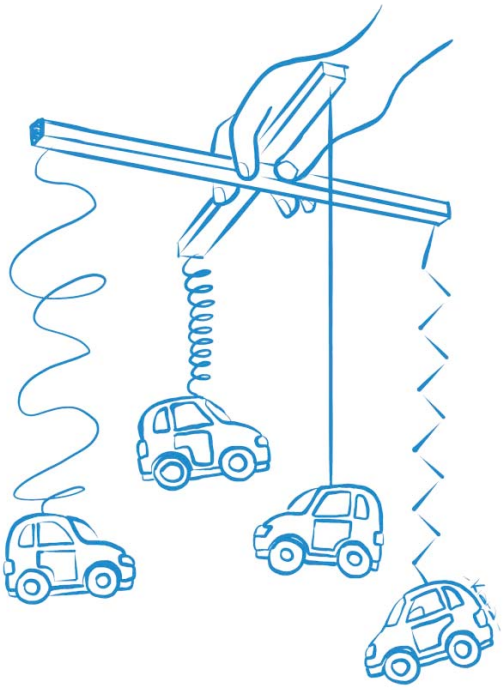
Meaningful Human Control over Automated Driving Systems



Marjan Hagenzieker

Institute of Transport Economics
Norwegian Centre for Transport Research

HFC meeting " Humans in modern transport systems – who is in control?", Oslo, 7 & 8 May 2019

Background

- Educational background in psychology (MSc & PhD Leiden University NL)

- Currently:

- Professor Traffic Safety Delft University of Technology (since 2014)
  &
- Senior researcher TØI Institute of Transport Economics
  (since September 2018)

- Title of presentation refers to title of interdisciplinary research project Delft funded by Dutch Research Council

The concept

- MHC stems from political debate on the use of (semi)autonomous weapon systems in the military, such as armed drones with autonomous capabilities
- Drivers should *at all times* have meaningful human control (MHC) over automated driving systems (ADS). This is different from approaches that focus on specific moments during partial automation when humans have to take back control.
- Fundamental idea MHC is that 'a human in the loop' is *insufficient* for preventing unwanted risks and accountability gaps. What is needed is that *the right human* is *able* and *motivated* to influence the behaviour of the system *in the right way at the right time.*
- MHC has three essential components:

1) human operators will make informed, conscious decisions;
2) human operators will be sufficiently informed about lawfulness of an action and its context;
3) human operators will be properly trained, to ensure effective control over the use of it.

Horowitz & Scharre 2015

Santoni de Sio & van den Hoven 2018

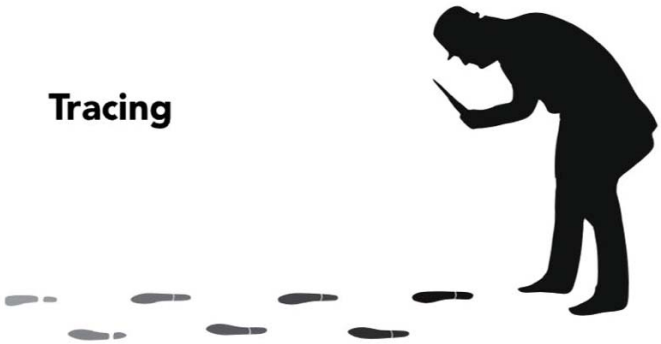
TUDelft

Two criteria for *meaningful human control* over autonomous systems

Tracking

Automated vehicles respond accurately to their controllers' intentions and to everything that is happening around them

Tracing



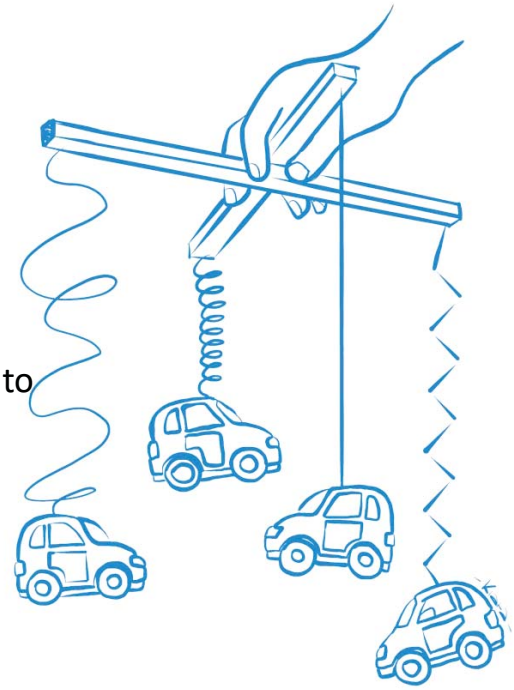
Since automated vehicles cannot be held responsible for accidents, there should always be a human who is expert of the system and who recognizes themselves as morally responsible for the behaviour of the vehicle

TU Delft

(Santoni de Sio & Van den Hoven 2018)

Meaningful Human Control over automated driving systems

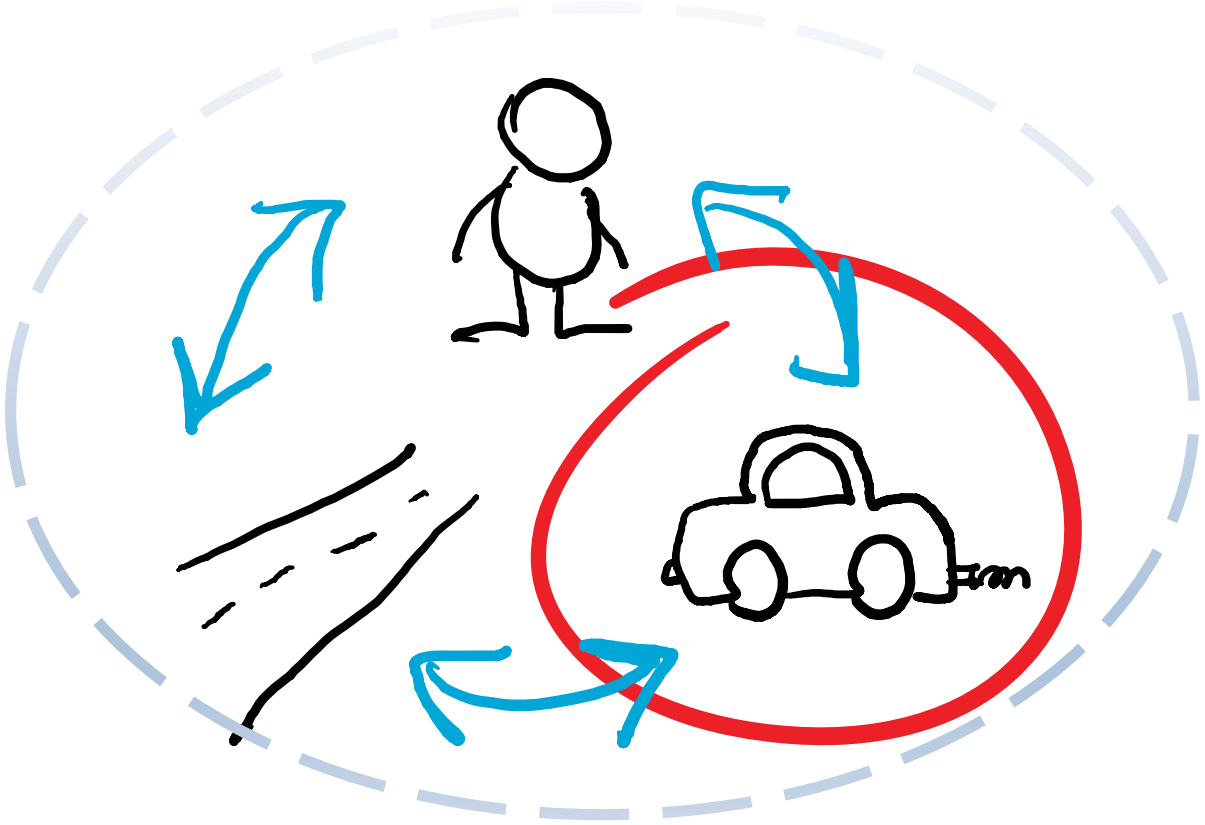
Applies not only to individual drivers of vehicles, also to remote operators, in control room, to designers of HMI, to policy makers, to design of the traffic system as a whole , ..



TU Delft

Drawing Maura Houtenbos

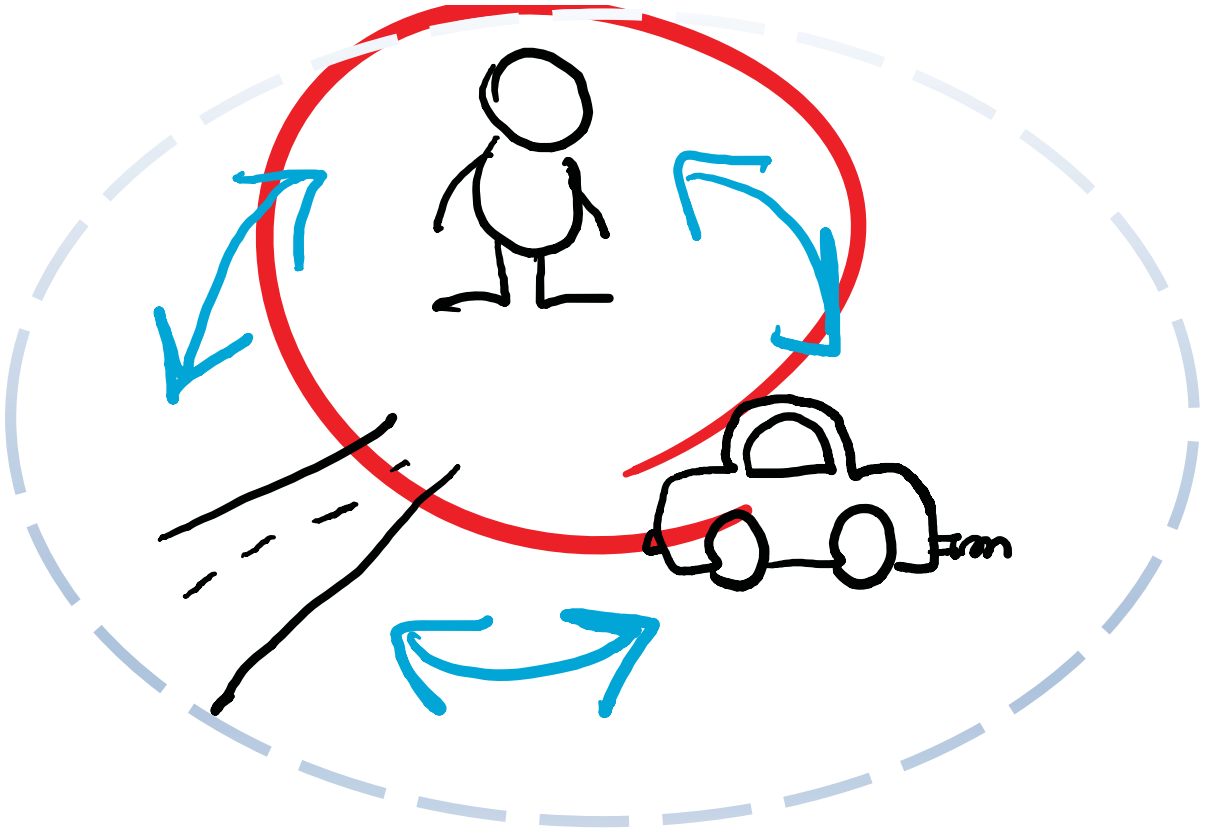
TU Delft



“Technology solves the problem of human error” ?

Drawing Maura Houtenbos

TU Delft



JACK STEWART TRANSPORTATION 03.30.18 10:34 PM

TESLA'S AUTOPILOT WAS INVOLVED IN ANOTHER DEADLY CAR CRASH

Bron: wired.com 30 maart 2018

AUTOMOTIVE

Deadly Tesla Crash Exposes Confusion over Automated Driving

ı federal investigation, ignorance of the technology's limitations comes into focus

Bron: Scientific American 2016

CNN 19 maart 2018

International Transport Forum | CPB Corporate Partnership Board

Safer Roads with Automated Vehicles?

"Claims of a more than 90% reduction in road traffic deaths resulting from automation eliminating crashes linked to human error are untested"

ITF/OECD 2018

NOS | Nieuws | Sport | Uitzendingen | TELETEKST | AEX | 0 km

Many smart and 'safe' gadgets in new cars hardly used

1 nov 2017

Modern cars are full of technical aids, but drivers hardly use them. That is because motorists do not know what their car can do.

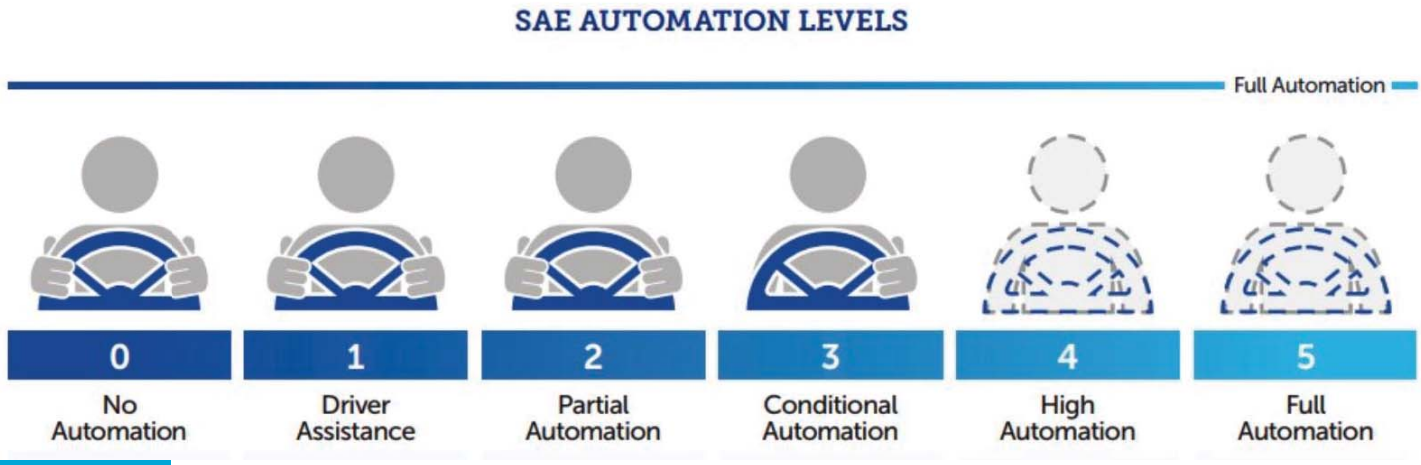
GESCHREVEN DOOR

Achille Prick

Redacteur Economie

NHTSA 2017

The higher the level of automation
... the less important the role of people?



TU Delft

non automated — fully automated

Drawing Maura Houtenbos

TU Delft

(From: Hagenzieker, M. (2015). "That bollard could have been a child". About road safety and behaviour of people in traffic. *Inaugural speech*. Delft University of Technology, 21 October 2015)

From: Heikoop et al. 2019. In *Theoretical Issues in Ergonomics Science*

TUDelft

Human \ Automation	SAE 0 No automation	SAE 1 Driver assistance	SAE 2 Partial automation	SAE 3 Conditional automation	SAE 4 High automation	SAE 5 Full automation
Skill (CIECA - Harmonisation of the Assessment of Driving Test Candidates)	128	127-114	114	114-43	40-0?	39-0?
Rule (Vienna Convention)	254	255-250	250	250-69*-66	51-29?	29-0?
Knowledge (SWOV factsheet, 2009; Lyman & Twisk, 1995; Advanced driver training courses, etc...)	65	65-81	81	81-34?!	0-?!	0?

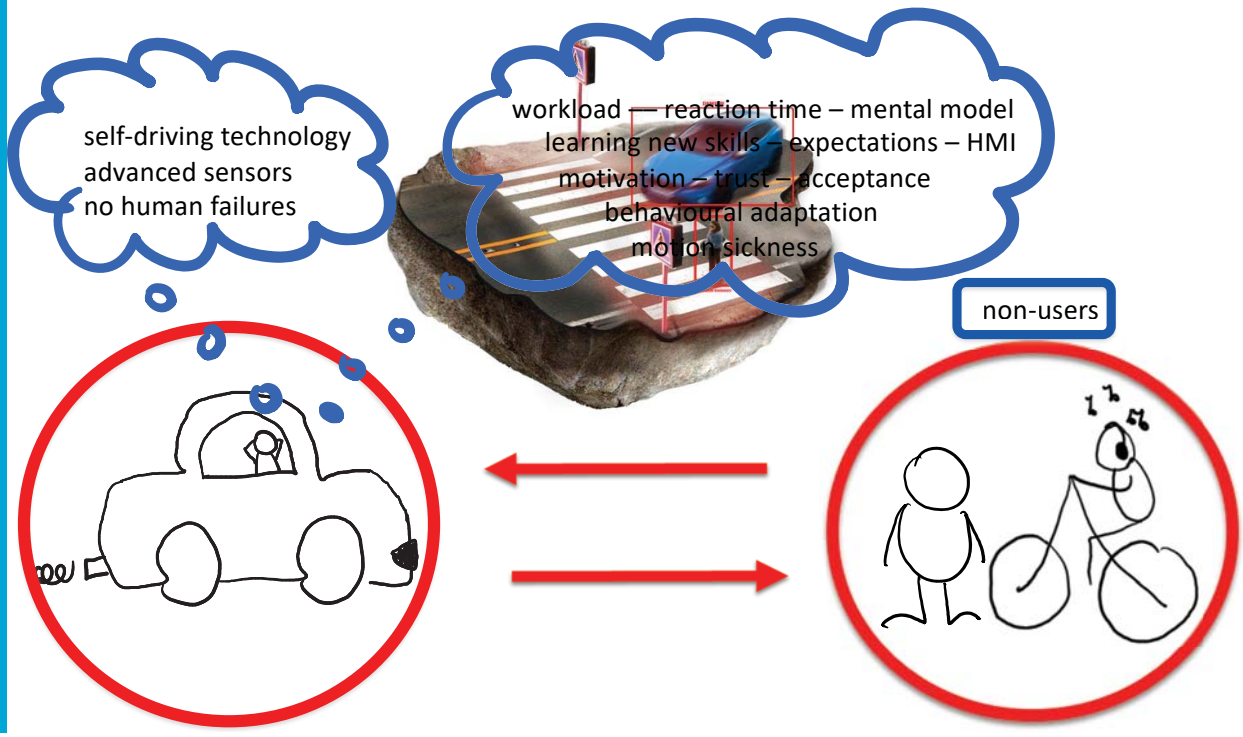
* = in case of an accident

? = Higher levels involve unknown situations and definitions

! = Skill/rule may already deteriorate to knowledge, adding up to drivers' knowledge-behaviour.

NB: Unweighted counts

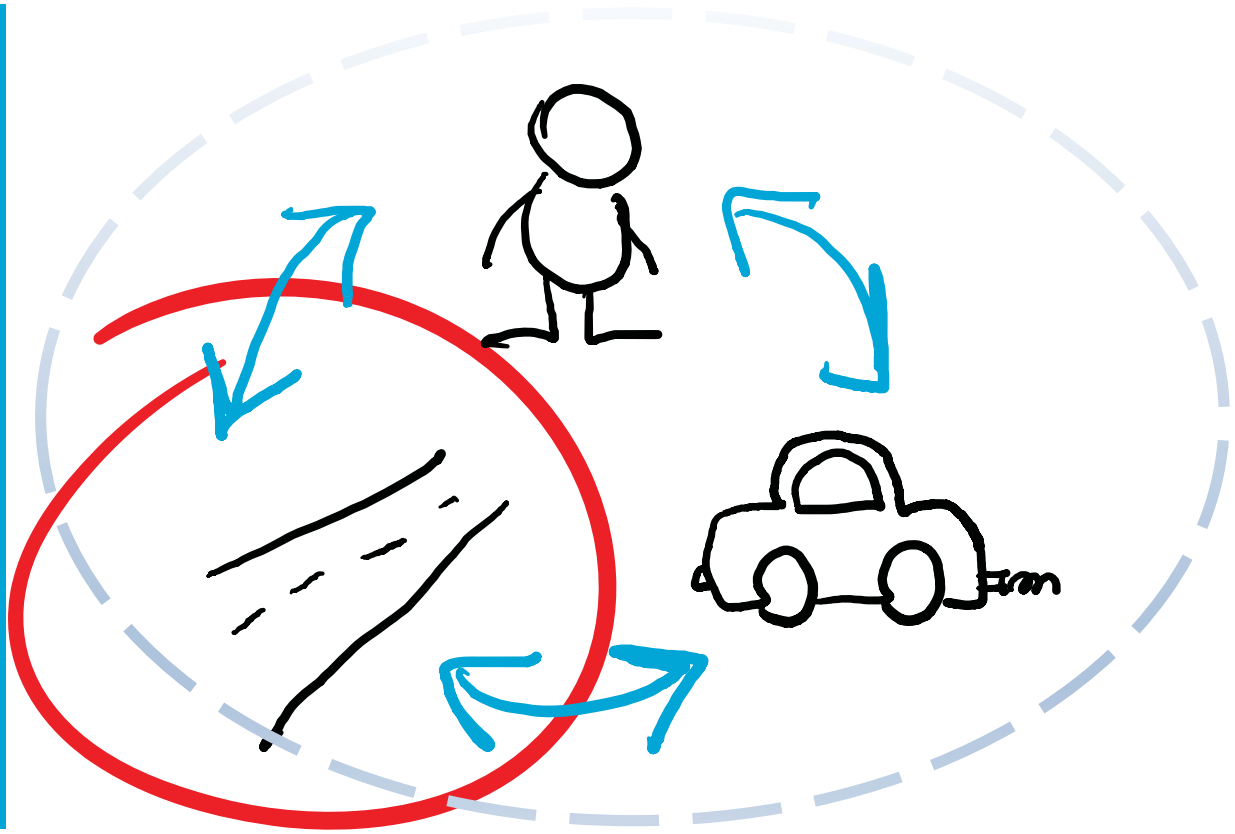
Different worlds



e.g.
Beggiato & Krems 2013
Hagenzieker et al 2019
Heikoop et al 2018
Rodriguez Palmeiro et al 2018

Figure adapted from: Hagenzieker (2015)

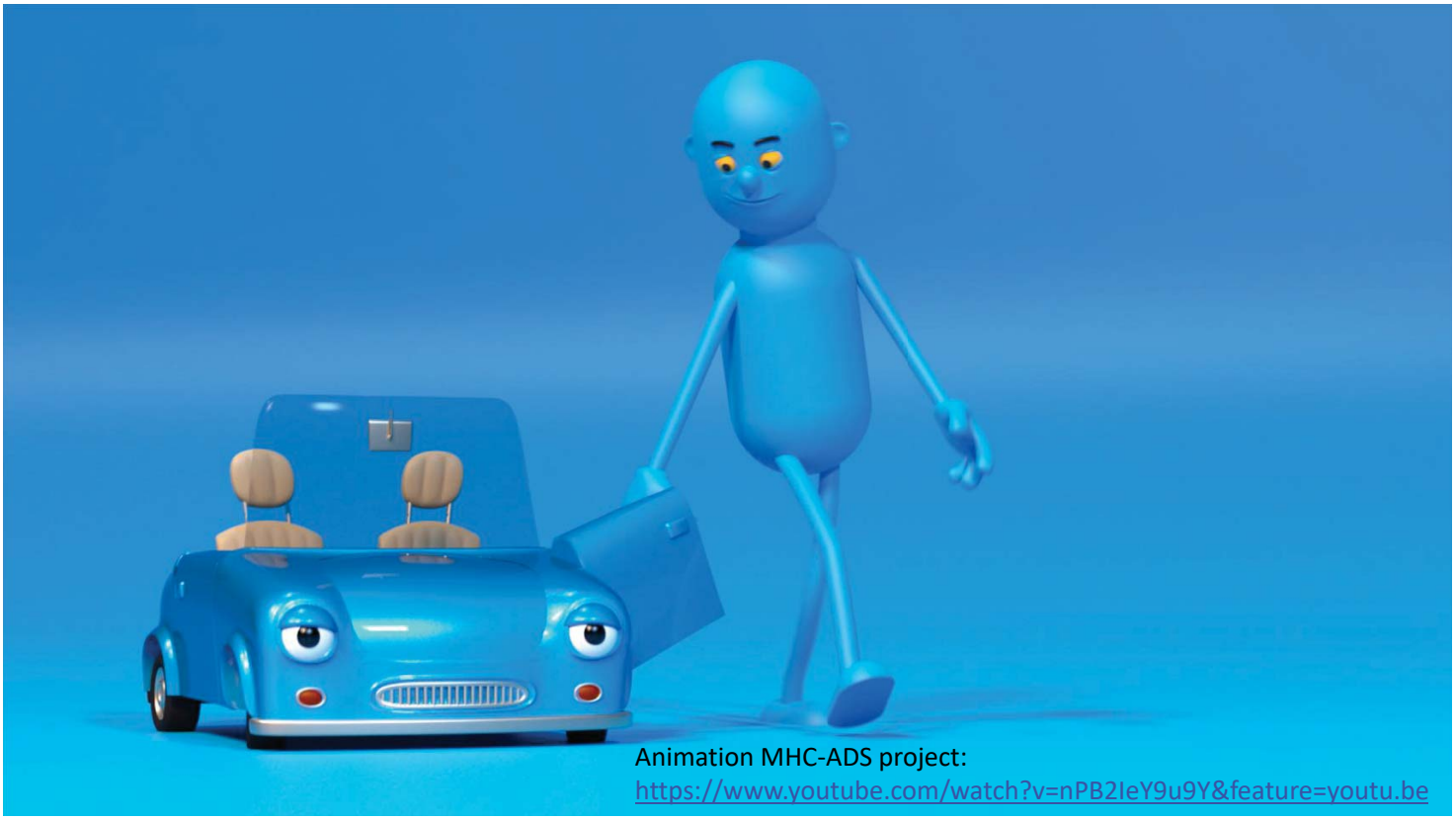
TUDelft



Drawing Maura Houtenbos

TU Delft

Source: Jame Bridle 2017, “Autonomous Trap 001”



Animation MHC-ADS project:
https://www.youtube.com/watch?v=nPB2IeY9u9Y&feature=youtu.be

Conclusions

- Human behaviour and control remain important
- Interdisciplinary approach & meaningful human control
- Complex combined effects on behaviour with increasing automation levels
- Mental models !
- New driving skills needed
- Road user education, training, and testing need to be ‘reinvented’ & tailored
- Inclusive future traffic system = for all road users:
- Take into account perspective of non-automated road users, such as pedestrians and cyclists

TU Delft

References

- Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. Transportation Research Part F: Traffic Psychology and Behaviour, 18, 47–57
- Hagenzieker, M. (2015). That bollard could have been a child". About road safety and behaviour of people in traffic. *Inaugural speech*. Delft: Delft University of Technology.
- Hagenzieker et al. (2019). Interactions between cyclists and automated vehicles: results of a photo experiment. *Journal of Transportation Safety & Security*.
- Heikoop et al. (2019). Human behaviour with automated driving systems: A quantitative framework for meaningful human control. *Theoretical Issues in Ergonomics Science.*
- Heikoop, D.D., de Winter, J.C.F., van Arem, B., Stanton, N.A. (2018). Effects of mental demands on situation awareness during platooning: A driving simulator study. *Transportation Research Part F,* 58, 193–209.
- Horowitz & Scharre (2015). Meaningful Human Control in Weapon Systems: A Primer. Center for New American Security. CNAS Working paper.
- MHC-ADS project https://www.tudelft.nl/citg/over-faculteit/afdelingen/transport-planning/research/projects/mhc-ads/
- Rodriguez Palmeiro et al. (2018). Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment. *Transportation Research Part F: Traffic Psychology and Behaviour, 58*, 1005-1020.
- Santoni de Sio & van den Hoven (2018). Meaningful Human Control over Autonomous Systems: A Philosophical Account. Front. Robot. AI, 28 February 2018.