Risks from permitting system-initiated lane changes

Oliver Carsten and Frank Mütze

ETSC

European Transport Safety Council

The precautionary principle

The precautionary principle is an approach to risk management, where, if it is possible that a given policy or action might cause harm to the public or the environment and if there is still no scientific agreement on the issue, the policy or action in question should not be carried out. However, the policy or action may be reviewed when more scientific information becomes available. The principle is set out in Article 191 of the Treaty on the Functioning of the European Union (TFEU).

The concept of the precautionary principle was first set out in a European Commission communication adopted in February 2000, which defined the concept and envisaged how it would be applied.

The precautionary principle may only be invoked if there is a potential risk and may not be used to justify arbitrary decisions.

https://eur-lex.europa.eu/EN/legal-content/glossary/precautionary-principle.html

More

- Recourse to the precautionary principle presupposes that potentially dangerous effects deriving from a phenomenon, product or process have been identified, and that scientific evaluation does not allow the risk to be determined with sufficient certainty
- The implementation of an approach based on the precautionary principle should start with a scientific evaluation, as complete as possible, and where possible, identifying at each stage the degree of scientific uncertainty
- Products should be treated as dangerous unless and until businesses do the scientific work necessary to demonstrate that they are safe

(European Commission, COM (2000) 1 final).

The precautionary principle:
The Wingspread Statement (1997)

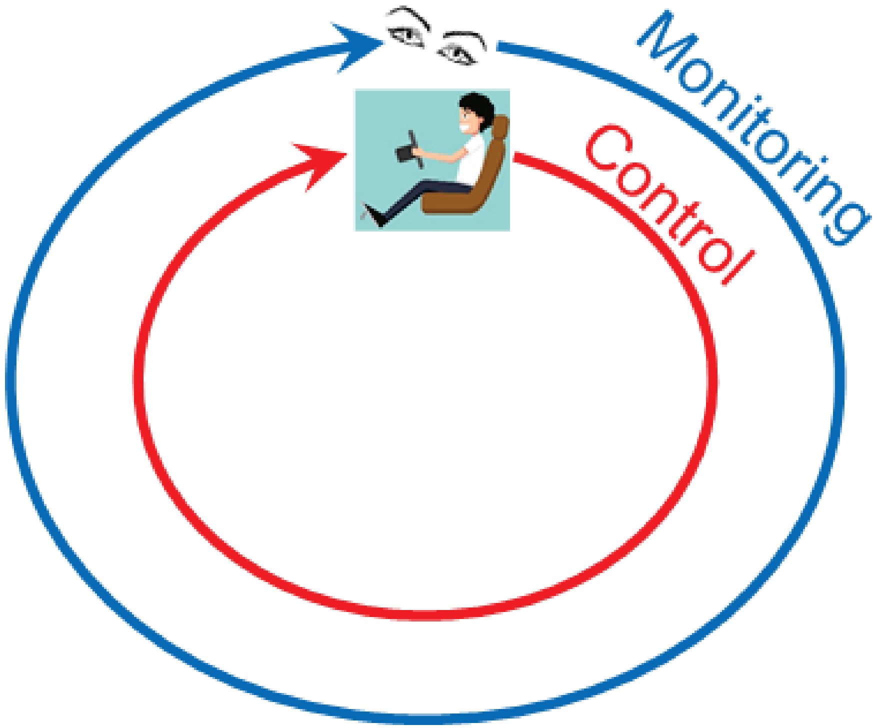
Five key elements to following the PP:

1. Taking anticipatory action to prevent harm in the face of scientific uncertainty.
2. Exploring alternatives, including the alternative of “no action.”
3. Considering the full cost of environmental and health impacts over time.
4. Increasing public participation in decision making.
5. Shifting the responsibility for providing evidence to the proponents of an activity.

https://www.healthandenvironment.org/environmental-health/social-context/history/precautionary-principle-the-wingspread-statement

Potential problems

Out of the loop



The monitoring (decoupled) and control (coupled) loops (Carsten and Martens, 2019

"We know from many 'vigilance' studies… that it is impossible for even a highly motivated human being to maintain effective visual attention towards a source of information on which very little happens, for more than about half an hour. This means that it is humanly impossible to carry out the basic function of monitoring for unlikely abnormalities, which therefore has to be done by an automatic alarm system connected to sound signals."

Bainbridge, L. (1983). Ironies of Automation.

Confusion with ADS (mode confusion)

		Type 1 Mode Awareness	
		Yes	No
Type 2 Mode Awareness	Yes	-	-
	No	Mode Confusion	-

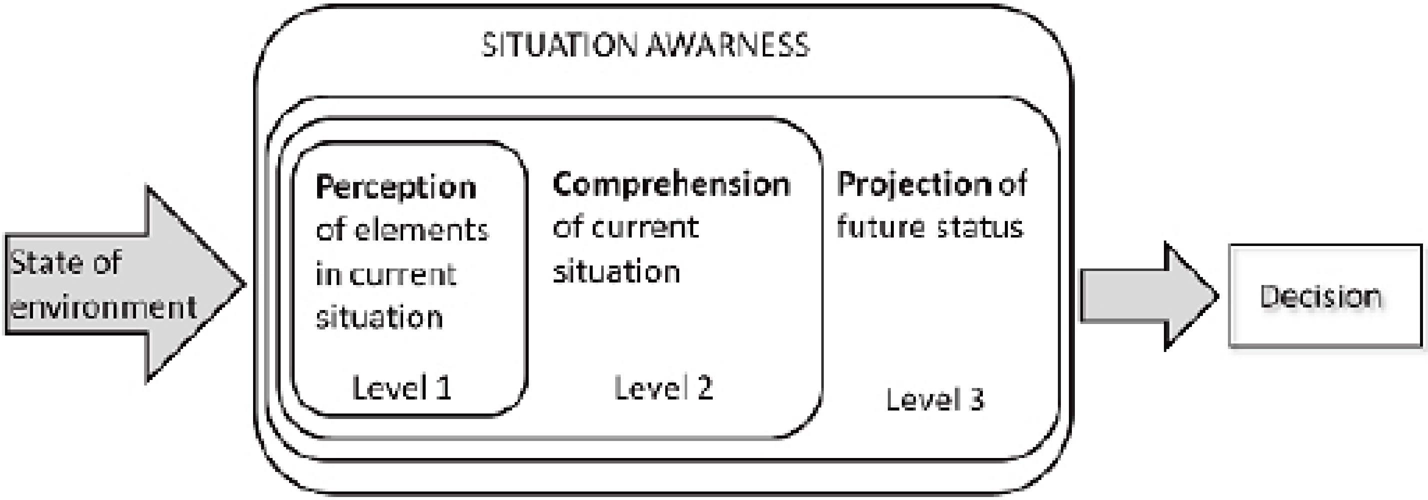
Type 1: General knowledge of the vehicle modes and driver responsibility in each mode

Type 2: Correct perception of the currently active mode

Source: Haghzare et al., 2022

Errors of Type 1 will be more likely to occur the more a DCAS performs like an ADS
Errors of Type 2 will occur if the DCAS changes mode rather silently

Imposition on Situation Awareness

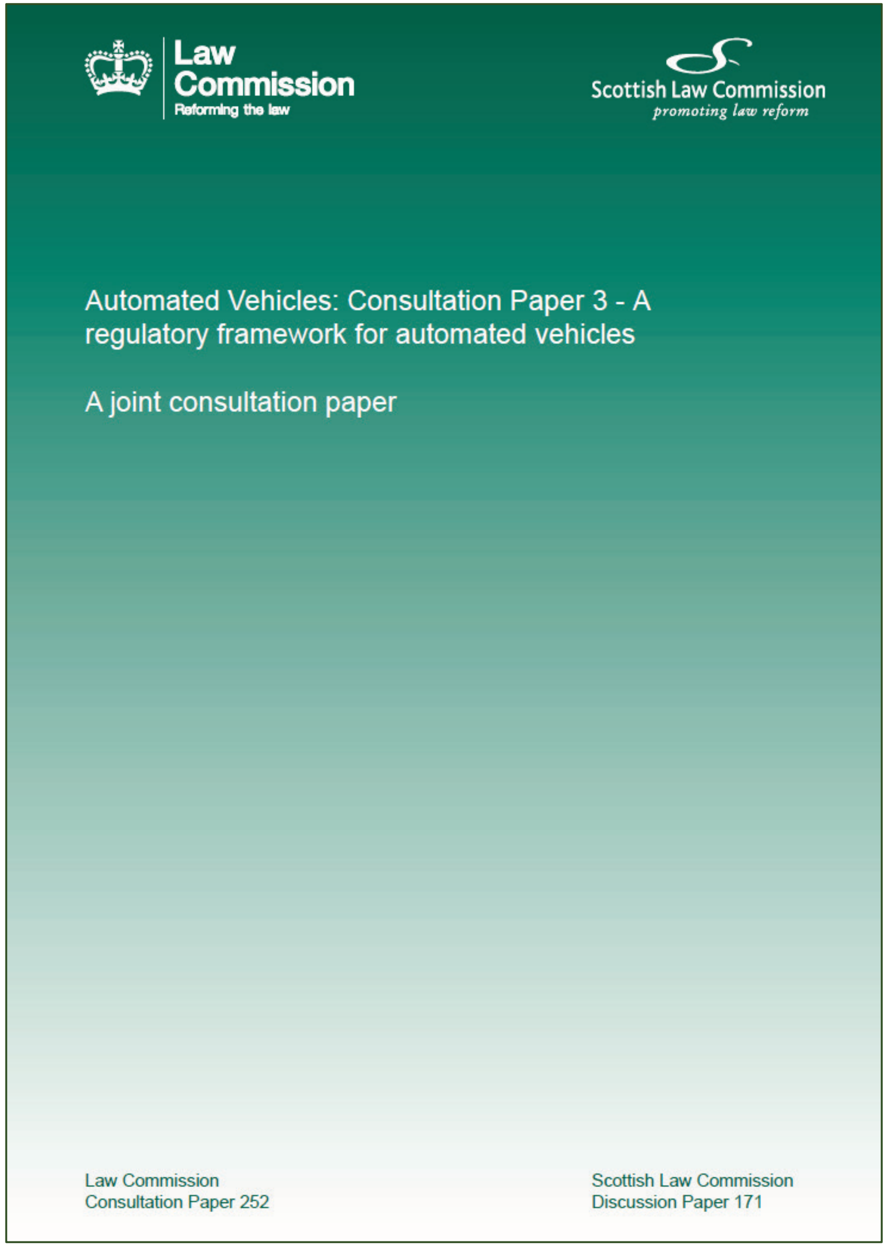


Source: Endsley, 1995

Now we have to add projection of what the ***system*** will do and when it will do it

Complacency

- If it handles several lane changes in a way that the driver deems satisfactory, then drivers will relax and pay less attention in a subsequent SILC
- That effect will extend to more general ***over-reliance*** on the system



Law Commission
Reforming the law

Scottish Law Commission
promoting law reform

Automated Vehicles: Consultation Paper 3 - A regulatory framework for automated vehicles

A joint consultation paper

Law Commission
Consultation Paper 252

Scottish Law Commission
Discussion Paper 171

4.9 As we move incrementally towards automated driving, the problem of over-reliance has come to the fore. Misuse and abuse of driving automation features has already led to dangerous risk-taking. For example, in 2018 a driver moved into the passenger seat on an English motorway, while the driver assistance system propelled the car at 40 miles an hour. When convicted of dangerous driving, the driver commented that he was "the unlucky one who got caught".[61] Unfortunately, there have also been fatalities associated with over-reliance on driver assistance features.[62]

4.10 This means that increases in the sophistication of driving automation do not necessarily lead to increases in safety. Several commentators have pointed to a "dip" in safety between helpful assistance and true self-driving. There comes a point in which the system gives the appearance of self-driving without being good enough to cope with common scenarios (such as weaving motorcycles, inconsistent road markings or parked vehicles partially occupying the lane).

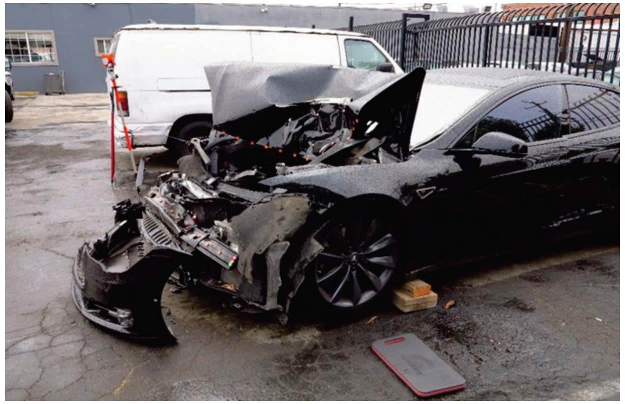
Evidence that complacency and over-reliance is a major problem in the real world

Fatal crash investigations by the U.S. National Transportation Safety Board

Rear-End Collision Between a Car Operating with Advanced Driver Assistance Systems and a Stationary Fire Truck Culver City, California, 22 January 2018

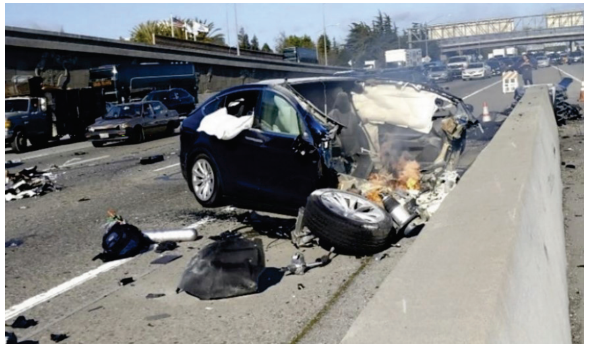
"The National Transportation Safety Board determines that the probable cause of the Culver City, California, rear-end crash was the Tesla driver's lack of response to the stationary fire truck in his travel lane, due to inattention and overreliance on the vehicle's advanced driver assistance system; the Tesla's Autopilot design, which permitted the driver to disengage from the driving task; and the driver's use of the system in ways inconsistent with guidance and warnings from the manufacturer."

NTSB Highway Accident Brief 19/07



Collision Between a Sport Utility Vehicle Operating With Partial Driving Automation and a Crash Attenuator
Mountain View, CA, 23 March 2018

“The National Transportation Safety Board determines that the probable cause of the Mountain View, California, crash was the Tesla Autopilot system steering the sport utility vehicle into a highway gore area due to system limitations, and the driver’s lack of response due to distraction likely from a cell phone game application and overreliance on the Autopilot partial driving automation system. Contributing to the crash was the Tesla vehicle’s ineffective monitoring of driver engagement, which facilitated the driver’s complacency and inattentiveness.”



NTSB Highway Accident Report 20/01

Collision Between Car Operating with Partial Driving Automation and Truck-Tractor Semitrailer
Delray Beach, Florida, 1 March 2019

"The National Transportation Safety Board determines that the probable cause of the Delray Beach, Florida, crash was the truck driver's failure to yield the right of way to the car, combined with the car driver's inattention due to overreliance on automation, which resulted in his failure to react to the presence of the truck."

Moreover:

"Based on system design, in an SAE-defined Level 2 partial automation system such as Autopilot, it is the driver's responsibility to monitor the automation, maintain situational awareness of traffic conditions, understand the limitations of the automation, and be available to intervene and take full control of the vehicle at any time. In practice, however, the NTSB and researchers have found that drivers are poor at monitoring automation and do not perform well on tasks requiring passive vigilance."

NTSB Highway Accident Brief 20/01

More human factors challenges

The “unsafe valley” (Flemisch et al., 2017)

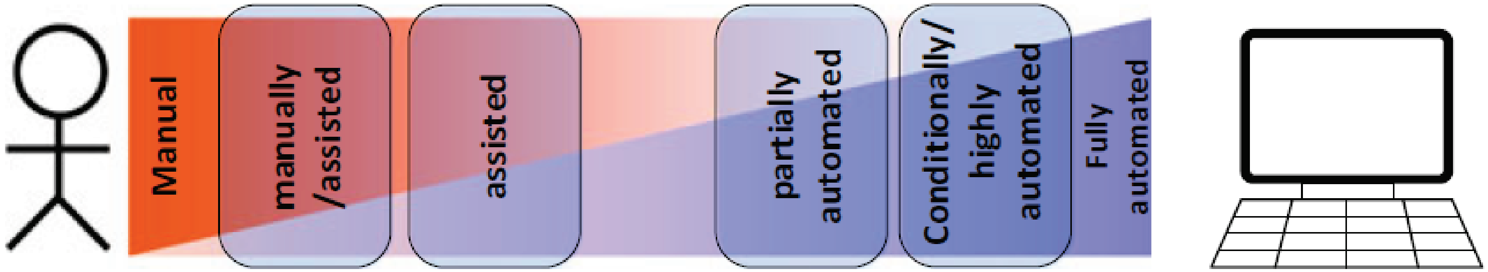


Fig. 1 Control distribution between the human and automation represented as an assistance- and automation-scale, here with explicit automation-levels/modes (inspired by Sheridan 1980; Flemisch et al. 2003, 2008, 2012, 2014, 2015a, b; Gasser et al. 2012b; SAE 2014)

A possible unsafe valley of automation can be found in the right half of the scale between partly- and highly-automated, which could be rather uncanny for the user, and more importantly, rather unsafe, as described further down.

In order to increase or maintain safety, and to harvest many chances of automation, the risks of automation have to be controlled. The systematic mapping and development of safeguarding measures down to safe combinations will require an interdisciplinary research and development, but will hopefully prevent naive and probably too trustful operators, system designers and engineers from falling into the unsafe valley, sometimes even an abyss of automation.

Automation surprise

Carsten and Martens (2019) identify 2 types:

1. Absence of an expected action
2. Presence of an unexpected action

(2) could happen, for example, if a DCAS made a SILC from the slow lane to overtake a slow moving lead vehicle when a driver was intending to take the next exit

Driver attention

- How do we maintain that all the way through a manoeuvre?
- An alarm cannot maintain attention; it can only stimulate resumed attention

Driver realisation that a steering intervention is needed…

- Will lead to delayed reaction
- And over-compensation

Expanding capability to urban roads — which manoeuvres?

- Roundabouts?
- Intersections?
- Zebra crossings?
- Lane choice at traffic lights?

If we have ALKS, why do we need this?

Is there empirical evidence of these theoretical problems leading to user failures?

Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel

Trent W. Victor, **Emma Tivesten**, **Pär Gustavsson**, **Joel Johansson**,
Fredrik Sangberg, and **Mikael Ljung Aust**, Volvo Cars, Gothenburg, Sweden

Human Factors, 2018

Objective: The aim of this study was to understand how to secure driver supervision engagement and conflict intervention performance while using highly reliable (but not perfect) automation.

Method: One hundred six drivers participated in three test-track experiments in which we studied driver intervention response to conflicts after driving highly reliable but supervised automation. After 30 min, a conflict occurred wherein the lead vehicle cut out of lane to reveal a conflict object in the form of either a stationary car or a garbage bag.

Results: Supervision reminders effectively maintained drivers' eyes on path and hands on wheel. However, neither these reminders nor explicit instructions on system limitations and supervision responsibilities prevented 28% (21/76) of drivers from crashing with their eyes on the conflict object (car or bag).

Conclusion: The results uncover the important role of expectation mismatches, showing that a key component of driver engagement is cognitive (understanding the need for action), rather than purely visual (looking at the threat), or having hands on wheel.

Capabilities of SILC

Can it handle…

- Lane drops?
- Roadworks, including contraflows?

If not, we need a Request to Intervene

IIHS and Euro NCAP
Safety Ratings of L2 Systems

First partial driving automation safeguard ratings show industry has work to do

March 12, 2024

The Insurance Institute for Highway Safety is introducing a new ratings program to encourage automakers to incorporate more robust safeguards into their partial driving automation systems. Out of the first 14 systems tested, only one earns an acceptable rating. Two are rated marginal, and 11 are rated poor.

"We evaluated partial automation systems from BMW, Ford, General Motors, Genesis, Lexus, Mercedes-Benz, Nissan, Tesla and Volvo," IIHS President David Harkey said. "Most of them don't include adequate measures to prevent misuse and keep drivers from losing focus on what's happening on the road."

Related

Current partial automation safeguard ratings

More about advanced driver assistance

Media contact

Joe Young
Director of Media Relations
office +1 434 985 9244
mobile +1 504 641 0491

For more information, visit our press room.

IIHS rating protocol

Requirements for a good partial automation safeguard rating

Monitors both the driver's gaze and hand position

Uses multiple types of rapidly escalating alerts to get driver's attention

Fail-safe procedure slows vehicle, notifies manufacturer and keeps automation off limits for remainder of drive

Automated lane changes must be initiated or confirme

Adaptive cruise control does not automatically resume after a lengthy stop or if the driver is not looking at the road

Lane centering does not discourage steering by driver

Automation features cannot be used with seat belt unfastened

Automated lane changes must be initiated or confirmed by the driver

Email from Alexandra Mueller, IIHS, 29 Feb 2024

While the debate still rages about whether vehicle-autonomously-initiated-and-executed maneuvers have a crash risk, no one knows what safeguards are “enough” to keep the driver adequate in the loop, which is what matters at the end of the day. Many people, myself included, think that as long as the system is partially automated it would be a design flaw for the system to do something without the driver’s involvement. What safeguards could be in place with such a functionality that would truly ensure the driver is aware of their responsibility and is prepared enough for the maneuver to intervene if necessary? We just don’t know.

The IIHS has taken the stance that a safeguard for this problem is to place limits on the functionalities that these systems have, which includes automated lane changes. Those limits require the involvement and verification of the driver for the initiation of maneuvers. This is one category of our ratings program on partial driving automation.

Given the risks and harm we have seen with systems that can be easily misused (intentionally or otherwise), we argue that the burden of proof is on the automakers to demonstrate the safety of such an automated lane change feature. As you pointed out, there is limited empirical research on this topic, so IIHS is currently doing a naturalistic observation study on automated lane changes.

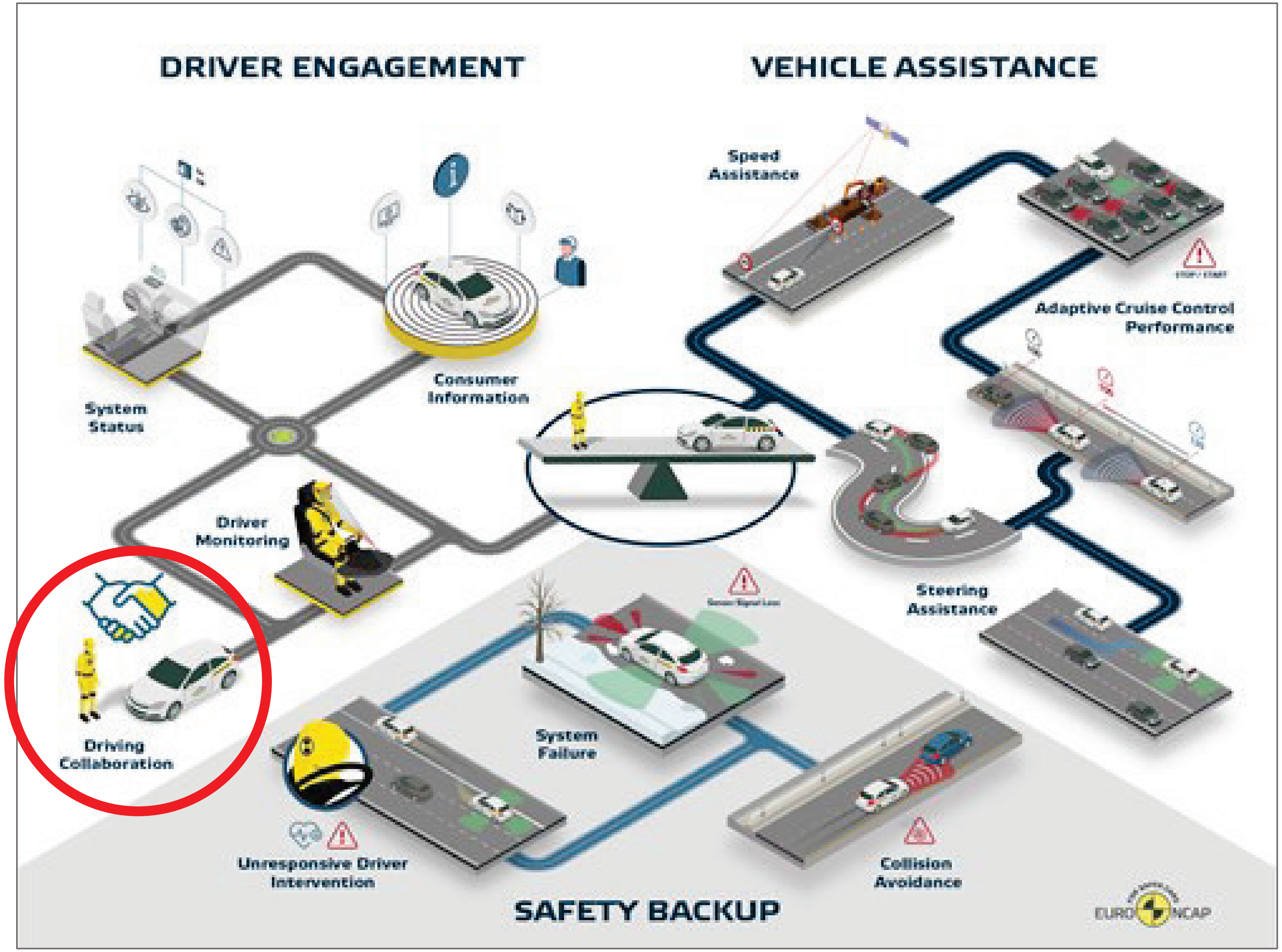
Alex

IIHS scoring of current systems

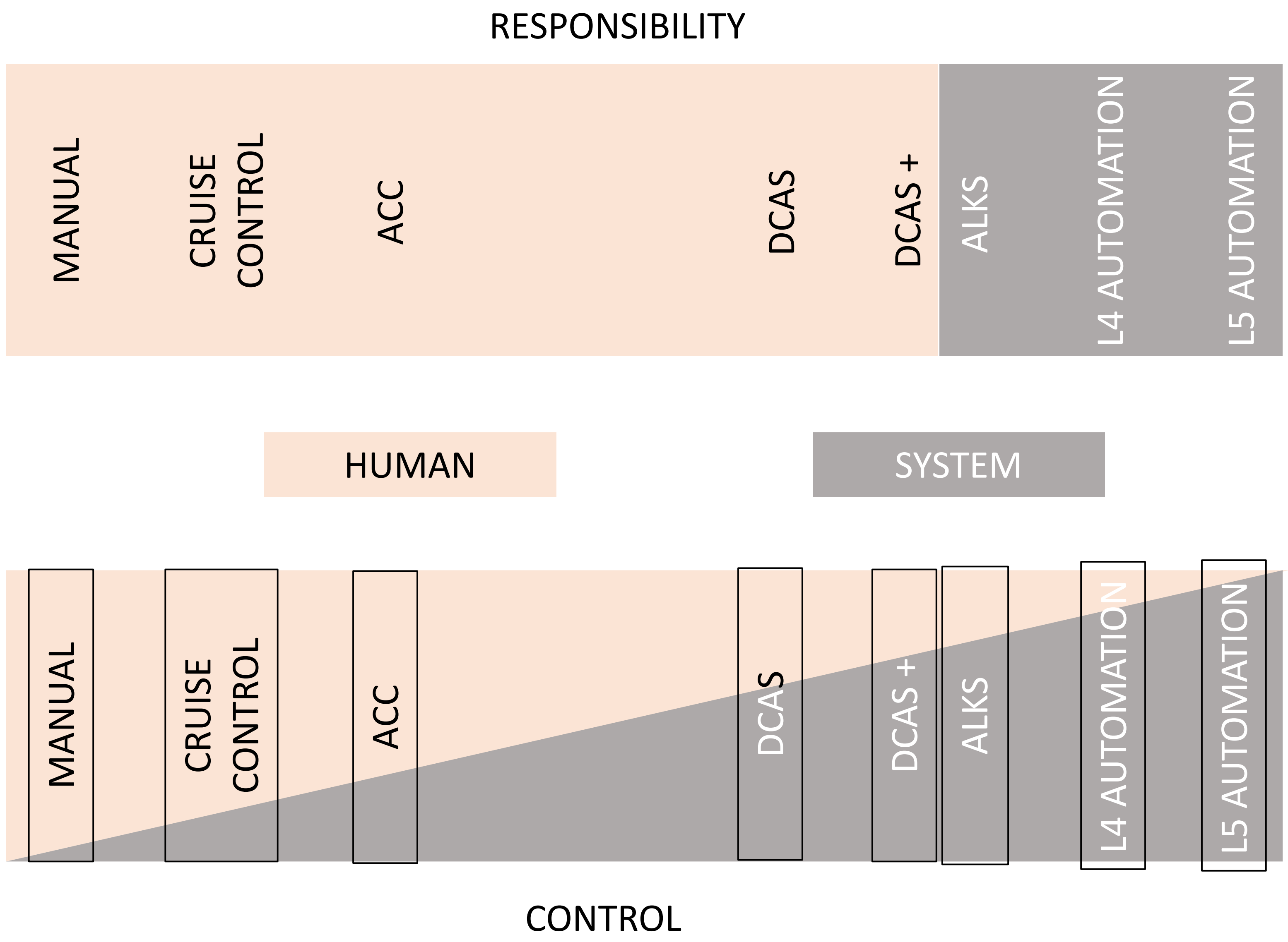
	Overall rating	Driver monitoring	Attention reminders	Emergency procedures	Driver involvement: Lane change	Driver involvement: ACC resume	Driver involvement: Cooperative steering	Safety features
Lexus Teammate with Advanced Drive 2022-24 Lexus LS	A	M	G	A	G	A	G	G
General Motors Super Cruise 2023-24 GMC Sierra	M	P	G	G	P	A	P	G
Nissan ProPILOT Assist with Navi-link 2023-24 Nissan Ariya	M	M	A	M	G	G	G	A
BMW Active Driving Assistant Pro 2023-24 BMW X1	P	M	P	A	G	P	G	A
Ford BlueCruise 2021-24 Ford Mustang Mach-E	P	A	G	M	G	M	G	P
Ford Adaptive Cruise Control with Stop & Go and Lane Centering Assist 2021-24 Ford Mustang Mach-E	P	A	G	M	G	G	G	P

	Overall rating	Driver monitoring	Attention reminders	Emergency procedures	Driver involvement: Lane change	Driver involvement: ACC resume	Driver involvement: Cooperative steering	Safety features
Genesis Highway Driving Assist 2 2023-24 Genesis G90	P	P	P	P	G	M	G	P
Genesis Smart Cruise Control/Lane Following Assist 2023-24 Genesis G90	P	P	P	P	G	G	G	P
Lexus Dynamic Radar Cruise Control with Lane Tracing Assist 2022-24 Lexus LS	P	P	P	P	G	G	G	M
Mercedes-Benz Active Distance Assist DISTRONIC with Active Steering Assist 2022-23 Mercedes-Benz C-Class	P	M	P	A	G	G	G	P
Nissan ProPILOT Assist 2.0 2023-24 Nissan Ariya	P	P	A	M	G	G	G	G
Tesla Autopilot, Version 2023.7.10 2021-23 Tesla Model 3	P	P	P	A	G	P	P	P
Tesla Full Self-Driving (Beta), Version 2023.7.10 2021-23 Tesla Model 3	P	P	A	A	P	P	P	P
Volvo Pilot Assist 2022-24 Volvo S90	P	P	P	M	G	G	G	P

Euro NCAP protocol for assisted driving, 2020



The ethical questions



DCAS+

Is it fair that drivers are responsible, while:

- a majority of the control is with the system?
- to ensure the safety of the vehicle, they have to take on an additional task: supervisor/operator of the system?
- they are not trained to be supervisors?
- we know that humans are not good at prolonged monitoring?

Crossroads

Current approach

- Problem: DCAS systems have shortcomings and limitations
  - Solution: Require the human driver to monitor the system's operation and intervene when needed
    - Problem: Inherent design paradox/flaw
    - Problem: automation of the driving task induces driver engagement in non-driving related tasks
      - Solution: Require driver monitoring systems
        - Problem: Shortcomings in the effectiveness of the DMS (e.g. eyes on the road doesn't equal mind on the road)
          - Solution: ??? How many safeguards more do we need?

Proposed approach

- Problem: DCAS systems have shortcomings and limitations
  - Solution: Follow the precautionary principle and not regulate additional capabilities, but instead focus on delivering safe automated driving

Thank you for your attention!

o.m.j.carsten@its.leeds.ac.uk
frank.mutze@etsc.eu