

# Automated Vehicle Safety Assurance: A Framework for Automated Driving Systems

A report commissioned by the Department for Transport  
(Non-Abridged Version)

Date: 07/07/2022



# Executive Summary

## Disclaimer

The information contained within this response does not necessarily represent the position of the Department for Transport.

## Introduction

This report details recommendations regarding a safety and security assurance framework for low-speed automated vehicles (LSAVs). It has been prepared by HORIBA MIRA (lead partner), University of York, TRL and Five, as part of a project commissioned by the Department for Transport (DfT). Whilst initially aimed at LSAVs, the report considers how the findings can be extrapolated to other automated vehicle use cases such as higher speed operation.

A prescriptive approach for LSAV approval has *not* been recommended due to the rapidly evolving state of the art and the lack of a standardised safety assurance method. Instead, the submission of a safety case report to the regulator is proposed; this would permit the employment of alternative safety assurance solutions whilst providing an appropriate safety record to enable robust scrutiny.

The safety case report would consist of multiple pieces of evidence to demonstrate safety: evidence to define the nature of the vehicles and their operation; evidence of appropriate analysis of safety; evidence of appropriate safety testing and evaluation (verification and validation); and evidence of appropriate safety management systems. It must also contain the 'safety argument': a structured description of how the evidence is sufficiently complete and comprehensive such that, when all the items are taken together, they support the claims that the overall safety of the LSAV type is acceptable.

In this report, we assume an operating model that consists of a Manufacturer of the LSAV and an Operator. The proposed process would involve a system safety case report being submitted by the Manufacturer, and a deployment safety case report being submitted by the Operator, although it is permissible for the Manufacturer and the Operator to be the same organisation in practice; indeed, it is anticipated that this may be a likely model for early commercial deployments of LSAVs. The safety case report would provide the regulator with all the necessary information, without providing the full information contained within the safety cases as developed and maintained by the Manufacturer and the Operator, which may be impractical to scrutinise.

The complete regulatory lifecycle consists of the phases:

- **Pre-Approval** – the engineering activities to develop the system and to acquire safety evidence prior to the application for approval;
- **Vehicle Type Approval** – the formal process of assessing the safety of the automated vehicle;
- **Deployment Approval** – the formal process of assessing the safety of a vehicle in its expected operational environment;
- **Monitoring** – capturing data while the vehicles are in service to validate safety case assumptions and to identify where remedial action is required;
- **Response** – the implementation of remedial actions;
- **Change** – proposals to adapt or improve the vehicle capability or the service.

## Definition of the System and Deployment

As a precursor to the safety evidence that is collated downstream, the nature of the system and its operation must be robustly defined. In addition to an ODD (Operational Design Domain), this report proposes that a 'TOD' (Target Operating Domain) must also be defined: while the former represents the *design intent*, the latter represents the *deployment reality*; this distinction is important since the two may not be identical.

It is proposed that the TOD should include a definition of the specific location of the deployment route(s) or geofenced area(s), such that the actual location of the deployment within the world is unambiguously defined, rather than merely described in an abstract manner by generic attributes that could apply to many possible, purely conceptual locations. The ODD may optionally be specific to the actual deployment route(s), or it may be generic such that a system developed to operate within it is compatible with multiple specific TODs.

The reason for requiring that a TOD be specific to a defined location is that some elements of the safety case are specifically linked to the deployment location. For example, a review of the operational safety of the route, such as identifying particular segments which may pose a hazard, or a review of the impact on traffic flows, would both be specific to the deployment location. Furthermore, this report argues that a significant proportion of the testing of the full system should be conducted upon the actual deployment route (and potentially upon a representative equivalent such as a 'digital twin' within a simulation). The TOD recommendations are therefore no more restrictive than the implicit development of the safety case itself.

It would not, for example, be permissible to test a system using solely locations within Greenwich and Coventry, and then approve the system as safe for deployment on a route in Milton Keynes upon which it has never been tested; the range of road permutations that exists in the world, and the challenge of identifying and testing the system's response to them, are too great for 'go-anywhere' approvals to be practicable within the foreseeable future.

The definition of the system must include behavioural competencies to state the functionalities that the system is required to perform, and a 'Minimum Equipment List' (MEL) of subsystems that must be in a fault-free state for the system to operate correctly in a given mode. It is anticipated that a system may have multiple definitions of ODD/TOD, behavioural competencies and MELs that can be combined in different ways to facilitate operation in degraded modes (e.g., to accommodate faults or adverse weather).

### Risk Framework and Safety Goals

A high-level safety analysis of road transport in general was undertaken to investigate hazards with the potential to harm road users; safety goals aimed at mitigating each hazard were formulated. A summary of the hazards considered is shown below:

- collision between the automated vehicle and another object (moving or stationary);
- direct harm to passengers from any of,
  - motion of the automated vehicle (e.g., hard braking manoeuvre)
  - a moving mechanism on the automated vehicle (e.g., door mechanism)
  - technological hazards (e.g., electric shock, fire)
  - personal safety concern (e.g., medical emergency, assault on-board)

The aim of the risk framework is to require a level of functionality and performance that can be argued to provide an acceptable level of safety in the context of use. The overarching principles which underpin this risk framework and its objective acceptance criteria are that new automated driving technologies should:

- not expose road users to unreasonable risk; and,
- support the societal goal to make road transport safe for all road users.

The safety of the LSAV should be argued and demonstrated on a case-by-case basis for each system and deployment, rather than by setting a universal quantitative threshold for acceptable risk. Manufacturers may demonstrate achievement by using risk acceptability principles such as ALARP (as low as reasonably practicable), GAMAB (*globalement au moins aussi bon*), and PRB (positive risk balance), and/or by using comparators such as an 'average' driver (as determined, for example, by road traffic accident statistics) or a 'competent and careful' driver.

The following top-level safety goals were formulated, which in turn form the basis for technical performance requirements.

Safety Goal (1)	Do not cause collisions
Safety Goal (2)	Avoid foreseeable collisions
Safety Goal (3)	Protect all persons within and in the vicinity of the vehicle from harm

### Assurance of the Automated Driving System

Functional safety, safety of the intended functionality (SOTIF) and cybersecurity are essential elements in assuring the safety of LSAVs. For each of these disciplines, standards exist that capture industry best practice, together with a regulation in the case of the latter (UNECE Regulation 155). This report therefore focuses upon what is required to adapt existing regulations and standards to the requirements of LSAVs, rather than attempting to duplicate them. Similarly, UNECE Regulation 156 provides a basis for assuring the safety of software updates.

External inputs which support or influence the dynamic driving task, such as GNSS (global navigation satellite system) data or communications from an operations centre, pose a particular threat due to the potential for loss or corruption, either accidentally or via malicious actions, in turn yielding hazardous behaviours. Consequently, it is recommended that if the system makes use of wireless communications data, it must be able to maintain safe operation even in the event that these signals are lost or corrupted. Whilst these signals may act as inputs to support the ADS in its decision making, a cautious approach should be adopted with regards to allowing remote inputs to directly control the vehicle's motion.

Where machine learning (ML) is utilised within the automated driving system (ADS), the report proposes an approach to verification and validation of ML which provides a foundation for addressing this critical aspect of LSAV design. Requirements for the subsystem performance should be specified, and the training data used within the development should be audited to confirm they are appropriate to meet these requirements (including consideration of the relevance, completeness, accuracy and balance of the data). Furthermore, the test data generated to verify the requirements should be similarly audited. Development of an appropriate architecture should be evidenced; this is likely to be an iterative process whereby multiple designs are compared and developed via testing.

The requirements should include quantitative metrics for performance, such as required levels of sensitivity and specificity of object classification, the accuracy of estimates for an object's position and speed, and robustness against variations within the operating environment. The testing should evidence that such performance metrics are satisfied. ML-based subsystems should also be tested once integrated into the full vehicle, to ensure that no undesired emergent behaviours become apparent. In-service monitoring should be used to validate assumptions about the system and its operating environment.

Test evidence derived from multiple test modalities will underpin the safety argument made in the safety case report. Further to traditional requirement verification for the components and (integrated) subsystems, the evidence must also include whole-vehicle test outcomes within realistic and TOD-representative scenarios that the deployed LSAV may be expected to encounter ('scenario-based testing'). Evidence must be provided to support the assertion that a sufficiently large and well-distributed sample has been taken across the space of reasonably foreseeable events and their combinations ('sample size' and 'sample coverage'). This must include justification of a validated test selection and generation methodology.

Assessment of the results of any test should extend beyond crude consideration of whether a collision occurred, to instead include:

- driving context; for example, there will be some scenarios where a collision is unavoidable due to the actions of other road users, but the system should be judged to have performed well if the level of mitigation (e.g., by emergency braking) compares favourably with what a human driver would likely have achieved. Depending upon the overall risk framework approach selected, this comparison could be based on characteristic benchmarks such as a 'competent and careful' driver, or an 'average' driver, derived from existing road traffic datasets.
- potential false-negative leading indicators; test programmes could reasonably be made more efficient at identifying system flaws, compared against cruder methodologies that look just at the 'global' outcomes, by flagging and reporting any *circumstantially inconsequential* failures. For example, the failure by a vehicle to detect a pedestrian, who just so happens not to enter directly into the path of said vehicle, should be recorded and noted as a 'near-miss' style failure which had the potential to yield harm, but transpired on that particular occasion not to do so. This implicitly reduces the test burden since 'fewer stars need to align' for a system flaw to be uncovered.

Reaching a decision upon the overall acceptability of the LSAV requires aggregation of all the evidence generated from individual test cases. This should be a function of: the quality of coverage; the proportion of scenarios in which the performance was deemed to be acceptable; the fidelity of each test modality; and the inherent, residual risk presented by any 'failed' scenarios, which itself derives from both the severity of the consequences and the level of exposure to the identified triggering conditions. This evidence aggregation will ultimately lead to an assertion that the vehicle is shown to be safe or to be unsafe during operation *with a certain level of statistical confidence*; this must then be judged by policy-makers as to whether it is *certain enough*.

### Human Factors

Human factors should be given consideration within both safety case reports, including assurance that all persons who interact with the vehicle have an appropriately clear interface with which to interact safely and confidently. This includes employees such as customer assistants or maintenance staff, as well as passengers and other road users. The safety case report should include consideration of normal operating conditions but also of emergency situations; in the latter, people may be in a state of distress and unable to think clearly, and interactions with emergency service or vehicle recovery personnel will be vital.

Particular consideration should be given where remote assistants play a safety-critical role in providing help to the vehicle when it needs guidance on the next appropriate move; such assistants must be provided with the necessary interface to allow adequate situational awareness and control.

### Operational Safety of the Deployment

A safety case report must be submitted by the Operator to demonstrate that the specific deployment will be acceptably safe. This must include consideration of any hazards that are specific to the route and what measures will be employed to mitigate them, as well as how workshop procedures such as repairs and routine maintenance will be managed so that the vehicle continues to meet the manufacturer's specification.

In the absence of full commercial, large-scale deployments of LSAVs at the time of writing, this report makes use of standards and guidance relating to automated vehicle trials and to other industries such as rail. However, once commercial deployments commence, and particularly once they begin to scale up, it is expected that a more mature state of the art will emerge. Regulators should therefore have a process to collect data on operational safety, and to use them to progressively develop operational safety regulations.

### **Safety Management Systems**

Both the Manufacturer and the Operator should provide evidence within their respective safety case reports that they have an effective safety management system (SMS) in place during the development and operation of the LSAV. The SMS should be bespoke to the system and deployment, and should follow the 'plan, do, check, act' process during its creation and improvement. A safety policy must be established to capture the organisation's commitment to safety and to the SMS implementation, including the responsibility for senior staff to help foster a strong 'safety culture'.

The SMS should define processes through which in-service data will be collected and used to trigger improvements, and set out how staff will be selected, trained, assessed and managed. For any safety-critical roles, this should include consideration of what expectations can reasonably be placed upon humans in the role and how to mitigate human error, such as the implementation of a fatigue risk management system.

Safety objectives and safety performance indicators should be defined to provide benchmarks against which in-service safety data can be compared. Safety risk assessment, safety reporting and employee consultation should also form a key part of a robust SMS and strong organisational safety culture.

# Table of Contents

<b>1</b>	<b>Acronyms and Definitions</b>	<b>9</b>
<b>2</b>	<b>Background</b>	<b>11</b>
2.1	Scope of Safety and Security Scheme	11
2.2	Scope of Work Package 1	12
2.3	Methodology	13
2.4	Ensuring that the Scheme is Flexible and Proportionate	14
2.5	Project Partners	15
2.6	Index into document content	15
<b>3</b>	<b>Approval Mechanism</b>	<b>20</b>
3.1	Approval Process	20
3.1.1	<i>Pre-approval</i>	22
3.1.2	<i>Vehicle type approval</i>	23
3.1.3	<i>Licensing and Deployment</i>	23
3.1.4	<i>Monitoring</i>	24
3.1.5	<i>Response</i>	25
3.1.6	<i>Change</i>	26
3.2	Use of a Safety Case and Safety Argument	27
3.2.1	<i>Nature of the Safety Case and Safety Case Reports</i>	27
3.2.2	<i>Vehicle Safety Case Objectives</i>	28
3.2.3	<i>Machine Learning Safety Case Objectives</i>	31
3.2.4	<i>(Operator) Deployment Safety Case Objectives</i>	33
3.2.5	<i>Vehicle Type Safety Management System Objectives</i>	33
3.2.6	<i>(Vehicle Operator) Deployment Safety Management System Objectives</i>	34
3.2.7	<i>Additional Objectives for Deployment SMS</i>	35
3.2.8	<i>The Balance Between Auditing and Testing</i>	36
3.3	Safety Goals and Risk Framework	37
3.3.1	<i>Explanation and Rationale of each safety goal</i>	44
3.4	Acceptance Criteria	64
<b>4</b>	<b>Definition of the System &amp; Deployment</b>	<b>69</b>
4.1	Operational Design Domain and Target Operating Domain	69
4.1.1	<i>Background</i>	69
4.1.2	<i>Recommendations</i>	85
4.1.3	<i>Supporting Guidance</i>	92
4.1.4	<i>Future Considerations</i>	95
4.2	Validating the Compatibility of the System and its Deployment	97
4.2.1	<i>Background</i>	97
4.2.2	<i>Recommendations</i>	99
4.2.3	<i>Future Consideration</i>	101
4.3	Behavioural Competencies	102
4.3.1	<i>Background</i>	102
4.3.2	<i>Recommendations</i>	109
4.3.3	<i>Future Considerations</i>	111

<b>5 Assurance of System Safety</b>	<b>112</b>
5.1 Functional Safety	112
5.1.1 <i>Adherence to ISO 26262</i>	112
5.1.2 <i>The Challenge of Qualifying AV Simulation Tools</i>	113
5.2 Safety of the Intended Functionality (SOTIF)	116
5.2.1 <i>Nominal Functionality vs Malfunctions</i>	117
5.3 Cybersecurity and Software Updates	119
5.3.1 <i>Cybersecurity</i>	119
5.3.2 <i>Software Updates</i>	120
5.4 Performance requirements	121
5.4.1 <i>Background</i>	121
5.4.2 <i>Proposed Risk Framework</i>	127
5.5 Minimal Risk Manoeuvres and Conditions	133
5.5.1 <i>Proposed Requirements</i>	133
5.6 External Inputs	134
5.6.1 <i>Background</i>	134
5.6.2 <i>Recommendations</i>	138
5.6.3 <i>Future Considerations</i>	140
5.7 Human Factors	141
5.7.1 <i>Methodology</i>	142
5.7.2 <i>Results and recommendations</i>	142
5.7.3 <i>Recommendations</i>	151
5.7.4 <i>Summary</i>	166
5.8 Safety of Machine Learning	167
5.8.1 <i>Inputs to the ML Assurance Process</i>	168
5.8.2 <i>ML Safety Requirements Assurance</i>	169
5.8.3 <i>ML Data Management</i>	171
5.8.4 <i>Model Learning</i>	176
5.8.5 <i>Model Verification</i>	178
5.8.6 <i>Model Deployment</i>	181
5.8.7 <i>Safe Use of ML in low-speed automated vehicles</i>	184
5.9 Test Programmes	194
5.9.1 <i>Background and State of the Art</i>	194
5.9.2 <i>Reporting results</i>	195
5.9.3 <i>Arguing ADS Safety robustly and with confidence</i>	200
5.9.4 <i>Coverage</i>	210
5.9.5 <i>Balance between Manufacturer / Developer Testing and Independent Testing</i>	213
5.10 System Monitoring	215
5.11 System Updates	216
5.12 Proposed Technical Requirements for GB Approval Scheme	217
<b>6 Operational Safety of Deployment</b>	<b>232</b>
6.1 Considerations for Deployment	232
6.1.1 <i>Background</i>	232
6.1.2 <i>Recommendations</i>	237
6.1.3 <i>Future Considerations</i>	244
6.2 Post-Deployment	245
6.2.1 <i>Background</i>	245



6.2.2	<i>Current State of the Art</i>	245
6.2.3	<i>Recommendations</i>	250
6.2.4	<i>Future Considerations</i>	251
<b>7</b>	<b>Other Evidence to be Supplied to the Regulator</b>	<b>252</b>
7.1	Manufacturer Safety Management Systems	252
7.1.1	<i>Problem Summary</i>	252
7.1.2	<i>Introduction to SMS, Safety Culture and Structure</i>	252
7.1.3	<i>Requirements and Recommendations</i>	253
7.1.4	<i>Summary</i>	260
7.2	Collated Systems-Level Administrative Requirements	262
<b>8</b>	<b>Conclusion</b>	<b>267</b>
<b>9</b>	<b>References</b>	<b>268</b>
<b>10</b>	<b>Appendices</b>	<b>278</b>
10.1	Appendix 1: Literature review methodology for 4.7 Human Factors	278
10.2	Appendix 2: Measurement methods for 4.7 Human Factors	279
10.3	Appendix 3: Summary of requirements from 21 <sup>st</sup> FRAV session	280
10.4	Appendix 4: Summary of requirements from European Commission draft (Dec. 2021)	282
10.5	Appendix 5: National Driving Standard – summary of review	288
10.6	Appendix 6: Safety Management Systems	326
10.6.1	<i>Appendix 6 – A: Background to SMS Structure and Processes</i>	326
10.6.2	<i>Appendix 6 – B: Creation and Maintenance of an SMS</i>	327
10.6.3	<i>Appendix 6 – C: Education and Awareness of an SMS</i>	332
10.6.4	<i>Appendix 6 – D: Continuous Monitoring and Improvement</i>	333

# 1 Acronyms and Definitions

## Acronyms

ADS	Automated driving system
ASDE	Authorised self-driving entity (as defined by the Law Commissions)
AV	Automated Vehicle
BEV	Bird's Eye View
BSI	British Standards Institution
CS	Cybersecurity
COD	Current operating domain [new concept explored within the report – the actual surroundings that the vehicle finds itself within at a given moment when deployed – this may not necessarily lie within the bounds of the ODD or TOD]
DDT	Dynamic Driving Task (see SAE J3016 [19] for definition)
DSC	Deployment Safety Case
DSCR	Deployment Safety Case Report
FS	Functional Safety
GB	Great Britain
ISO	International Organization for Standardization
LSAV	low speed automated vehicle
MEL	Minimum Equipment List
MRC	Minimum Risk Condition (see SAE J3016 [19] for definition)
MRM	Minimum Risk Manoeuvre (see SAE J3016 [19] for definition)
NHTSA	National Highway Traffic Safety Administration
NUIC	No user in charge (as defined by the Law Commissions)
TOD	Target Operating domain [new concept explored within the report – the specification of the deployment domain within which the system will operate]
ODD	Operational design domain
OEDR	Object and Event Detection and Response
SAE	Society of Automotive Engineers

SOC	Safe Operating Concept
SPI	Safety Performance Indicator
SOTIF	Safety of the Intended Functionality
UIC	user in charge
UNECE	United Nations Economic Commission for Europe
VSC	Vehicle Safety Case
VSCR	Vehicle Safety Case Report

## Definitions

This report uses the terms and definitions set out in the BSI Connected and Automated Vehicles Vocabulary version 4 (BSI Flex 1890, 2022); the report should therefore be read in conjunction with this vocabulary.

Where appropriate terms to convey a particular meaning do not currently exist within the BSI vocabulary, other means will, by necessity, be taken to provide a clear and concise definition. The following definitions from the draft EU Commission Implementing Act are carried over:

**‘Nominal traffic scenarios’** are reasonably foreseeable situations encountered by the ADS when operating within its ODD. These scenarios, often referred to as “traffic scenarios”, represent the non-critical interactions of the ADS with other traffic participants and generate normal operation of the ADS.

**‘Critical scenarios’** are scenarios related to edge-cases and operational insufficiencies, not limited to traffic conditions but also including environmental conditions, human factors, connectivity, and miscommunication. Critical scenarios lead to emergency operation of the ADS.

Other definitions are provided within the body of the document, where the relevant topic is discussed, to ensure a common understanding is reached. Where the term is a newly-introduced one that has not previously been used outside this project, this is highlighted to the reader.

## 2 Background

### 2.1 Scope of Safety and Security Scheme

This report forms part of a project commissioned by the UK's Department for Transport (DfT) to support the development of an approval scheme for assurance of the safety and security of automated vehicles. This forms part of DfT's wider aims of boosting economic growth, building a One Nation Britain, improving journeys, and ensuring safe, secure and sustainable transport.

Although the scope of this report is limited to low-speed automated vehicles (LSAVs), it is expected that it may help inform future phases that will widen the scope to accommodate a broader range of vehicle categories and operating environments. As such, the report considers how the proposed requirements and supporting guidance could be applicable to applications beyond the current scope, and what further work may be needed in the future to support such extension of the scheme. This includes consideration of expanding the capabilities of the system, expanding the range of operating environments and remaining flexible such that new technology developments can be accommodated; this latter aspect is important in the context of automated vehicles due to rapid technological evolution meaning that the state of the art is not stable.

Furthermore, the scheme must be flexible enough to meet the various needs of all stakeholders, and must therefore:

- Ensure that the burden placed upon manufacturers and operators is proportionate to the level of risk posed, such that safety oversight is balanced against the time and resources required;
- Be practical to implement for both industry and regulators;
- Be appropriate for both small and large organisations involved in the development or deployment of LSAVs;
- Be broadly aligned with international best-practice

The scope of this report is restricted to fully-electric, highly-automated pods and shuttles with no driver present (i.e. with no human inside the vehicle who is responsible for monitoring the driving task, overriding the system, or reacting to takeover requests from the system). These would operate upon one or more fixed routes or within a fixed geographical area, with at least part of this consisting of public roads. The scope is further summarised in Table 1.

Characteristic	Scope
Purpose	Carriage of goods or passengers (maximum 16; seated, standing or mixed)
Level of Automation	Highly automated without a driver present
Powertrain	Fully electric
Maximum Speed	20 mph
Maximum Mass (gross vehicle weight)	3,500 kg
Operating Environment	Roads with a speed limit up to 30 mph with mixed traffic (including Vulnerable Road Users); Areas which may include high density of pedestrians; Dedicated roadways (which may or may not have segregation barriers); Operating on a fixed route or within a fixed geographical area

*Table 1: Summary of the scope.*

The project as a whole consists of five work packages:

- Work Package 1: Safety of the Automated Driving System (ADS)
- Work Package 2: Defining a codified behavioural rule set
- Work Package 3: Scenario generation, selection and coverage
- Work Package 4: Non-ADS vehicle requirements
- Work Package 5: In-use safety monitoring of the ADS

This report details the findings of Work Package 1, and covers all aspects of the Work Package 1 scope (as set out subsequently in Section 2.2). However, there are naturally many areas where the scope of the work packages overlaps or has strong interdependencies; therefore, whilst the other work packages are taking the lead on their areas of scope, this report does, by necessity, touch upon their scope in order to set out the wider picture of how the scheme as a whole should function. To support this, extensive interaction has taken place between the work packages, including regular alignment meetings attended by all work packages and separate meetings featuring a subset of the work packages.

Contributions to the project provided by DfT, the Vehicle Certification Agency (VCA), the Centre for Connected and Autonomous Vehicles (CCAV) and the Law Commissions are gratefully acknowledged.

### Disclaimer

The information contained within this response does not necessarily represent the position of the Department for Transport

## 2.2 Scope of Work Package 1

Work Package 1 examines what is required to assure the safety of the ADS prior to entry into service. The non-ADS functionality of the vehicle is addressed by Work Package 4, although it should be noted that there is an intrinsic relationship between safe functioning of the ADS and the characteristics of the vehicle that it is installed within. This is because the same ADS installed upon different vehicles, with different dimensions, sensor locations, actuator characteristics, vehicle dynamics etc. may be expected to exhibit significantly different performance, highlighting the need for the approval of the ADS behaviour to consider the nature of the vehicle type that the ADS is installed within. Work Package 1 therefore touches upon the non-ADS aspects of the vehicle in order to ensure compatibility, but leaves Work Package 4 to set the recommendations within this area.

For the ADS equipped vehicle to perform the required functionality within the required operating environment, the safety assurance must consider risk arising from:

- Malfunctioning behaviour of the vehicle
- The intended functionality of the vehicle
- Unintentional or intentional (including malicious) misuse
- Other events relating to the vehicle and its environment (e.g. tyre blowout, sudden adverse weather events)

Ensuring safe behaviour of the vehicle type should be a continuous process throughout the deployment lifetime. However, the scope of Work Package 1 is limited to gaining appropriate confidence in the safety *prior to deployment*, including consideration of whether the vehicle is acceptably safe and secure, whether appropriate safety management systems (SMSs) are in place, and whether the ADS meets applicable technical requirements and rules. This report does not, therefore, examine how the system would be monitored during its lifecycle, which falls within the remit of Work Package 5, but it does examine how to ensure that appropriate processes have been put in place prior to approval and deployment, in order to support this monitoring and facilitate any resulting updates to the safety case. This falls under the SMS aspect of Work Package 1's scope.

Similarly, this report does not examine what behavioural rules the vehicle should meet, which falls within the scope of Work Package 2, but it does examine how data from test programmes would be aggregated

to form an overall conclusion about the performance of the vehicle and the coverage provided by the test programme. The acceptability of performance in the individual test cases would be determined by adherence to the behavioural rules (Work Package 2) and the processes used to select, run and analyse scenarios via a database (Work Package 3), potentially further augmented by other assessment metrics as examined within Section 5.9.

The proposals within this report are provided for consideration by DfT in order to support the development of a regulatory process, but should not be interpreted as a description of DfT policy; this report forms an input to DfT's processes and policies, as opposed to an output from them.

## 2.3 Methodology

A key aim of the report is to ensure that it is aligned with:

- International best practice, and;
- The needs of stakeholders within Great Britain (GB).

In order to achieve this, multiple inputs have been used to acquire the necessary information to base proposals upon; these are summarised in the subsequent sections.

### Literature Review

A significant proportion of the resource allocated to this project has been used to undertake comprehensive reviews of the literature relating to the various topics. This has included reviewing:

- Regulations, such as UNECE (United Nations Economic Commission for Europe) regulations relating to automated functionality within road vehicles, or GB domestic regulations relating to road traffic.
- Regulatory proposals, such as those presented within the UNECE FRAV (Functional Requirements for Automated and Autonomous Vehicles) and VMAD (Validation Method for Automated Driving) working groups.
- Standards, including international and British standards and including those aimed at automated vehicle system safety, automated vehicle trial operational safety and ADAS (advanced driver assistance systems) for production vehicles.
- Academic papers such as journal papers, conference papers and 'grey literature'/ white papers published by projects or institutions themselves.
- Relevant websites such as those relating to official safety processes (e.g. for road or rail accident investigation).
- Other sources such as guidance documents produced by or for testbeds, or presentations that have been shared online

### Stakeholder Reviews

In order to gain an understanding of the needs of relevant stakeholders, the work package 1 consortium has undertaken two rounds of consultation:

- An initial phase of consultation, which was undertaken in the form of interviews via a virtual meeting. Participants were asked relatively open questions relating to key topics to gain a baseline understanding of their needs. This phase was undertaken early in the project, as a precursor to any proposals being drafted.
- A second phase of consultation, which was in written form; participants were given questions directly relating to key areas of the draft version of this document, with background information preceding the question where appropriate. This was undertaken late in the project, in order to gain targeted feedback on areas of the report that could prove contentious.

Where stakeholder feedback was obtained in relation to a report section, the findings are summarised within that section itself.

### Discussion Within the Project Team

The personnel involved within the project, both for work package 1 and for the other work packages, have significant experience within relevant fields such as automated vehicle trials, ADAS, safety engineering, machine learning and type approval regulations. This was further supplemented by significant contributions from representatives of DfT and VCA.

To maximise the use of this experience and to ensure alignment between work packages, regular 'Cross Work Package' meetings were held to update on progress and debate key issues. These were further supplemented by separate meetings between relevant work packages to address particular topics, and by additional meetings with representatives from closely related projects such as the consultation undertaken by the Law Commissions (2022), and the Safe MRX and CertiCAV projects led by Connected Places Catapult.

Furthermore, draft proposals and a draft of this document have been reviewed by DfT and VCA, resulting in extensive feedback that has helped inform the final version.

## 2.4 Ensuring that the Scheme is Flexible and Proportionate

This report details recommendations for how such a scheme should operate and what safety evidence regulators should look for. Where possible, the report proposes unambiguous requirements, but where the diverse and rapidly-evolving nature of the technology precludes such a prescriptive approach, guidance is provided to assist manufacturers, operators and regulators to reach informed decisions. Such an approach allows flexibility to adapt to different use cases and solutions whilst supporting fair and proportionate assessment.

Furthermore, the adoption of a safety case report submitted to the regulator for approval, as opposed to the setting of prescriptive assessments and tests, allows flexibility in terms of:

- The level of regulatory burden that is appropriate - this allows a balance to be made between protecting against undue risk and ensuring that the time and cost involved is reasonable.
- The pieces of safety evidence included to provide assurance, and the nature of the safety argument that justifies the completeness and relevance of the evidence. This may be expected to vary significantly according to the technology used and the nature of the deployment, but could also vary as a function of the preferences of the organisation creating the safety case; in the absence of a settled and agreed state of the art for LSAV safety arguments, it is not possible to prescribe one single way to structure a safety argument and to rule out all others.
- The format and content used within each individual piece of safety evidence – again, there is, as yet, no established norm for how evidence should be presented. Therefore, flexibility must be maintained for each organisation to put forward, with justification, an approach that works for that particular application

The nature of the safety case, and the safety case report that is submitted to the regulator, is examined further in Section 3.2.1.

It should be noted that automated vehicles are a technology that presents an arguably unprecedented level of complexity with regards to assuring safety and security, as a result of the complexity of the systems, of the behaviours they have to perform, and of the often chaotic and unpredictable environments they have to operate within. Given this unprecedented complexity, combined with the novelty of a technology that is as yet unproven with regards to full commercial deployments, it would not be unreasonable to expect the volume of testing and analysis required to produce sufficient evidence of acceptable safety to be similarly unprecedented.

## 2.5 Project Partners

The following organisations and individuals contributed to the production of this document; their efforts are gratefully appreciated.

- HORIBA MIRA Ltd. (Lead Partner)
  - Richard Hillman
  - Edith Holland
  - Michael Orgill
  - Clive Carter
- TRL
  - Chris Fordham
  - Georgina Abram
  - Gareth Slocombe
- University of York
  - John McDermid
  - Mike Parsons
  - Richard Hawkins
  - Jennifer Dick
- Five
  - Iain Whiteside

## 2.6 Index into document content

To facilitate navigation within the document, an index and explanation of the different report sections is given here. The document contains recommendations concerning both the attributes of the ADS as part of a low-speed automated vehicle itself, as well as processes and ways of working when developing these vehicles. A representation of the low-speed automated vehicle and its elements in the road transport ecosystem is shown in Figure 1, with Table 2 indicating relevant sections addressing their elements.



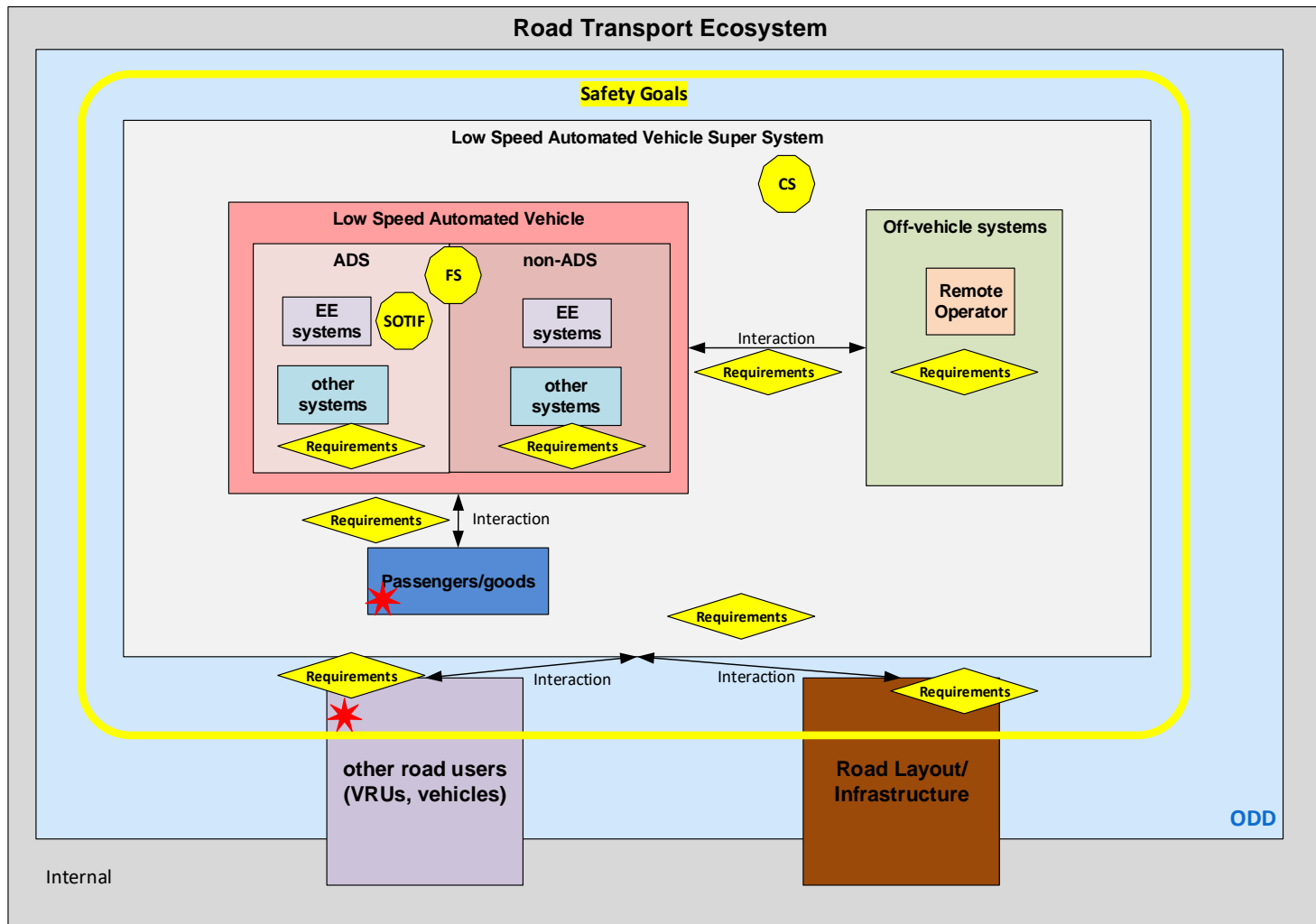


Figure 1: Low-speed automated vehicle and road ecosystem representation.

Similarly, Figure 2 shows an indicative process flow for the development, deployment and operation of an LSAV; the sections which treat each phase in detail are listed in Table 2.

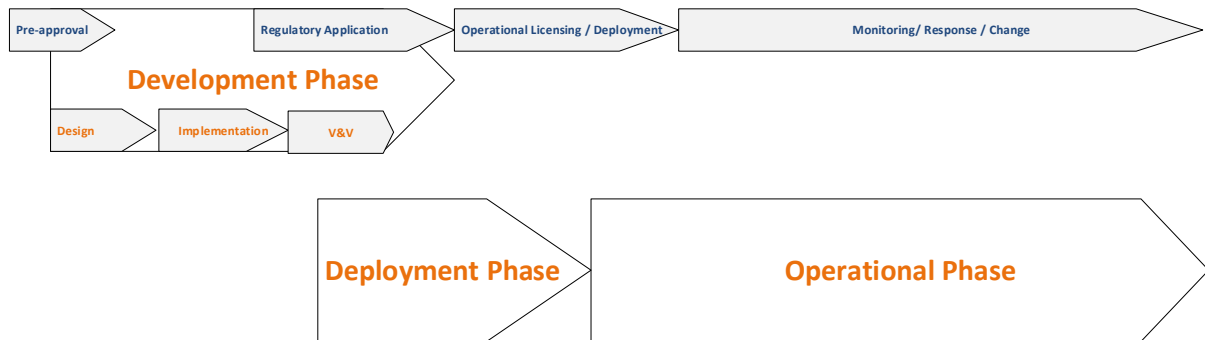


Figure 2: Indicative process flow for development, deployment and operation.

Figure Reference (LSAV element, phase or organisation involved)	Specific topic	Section
Whole vehicle	Approval mechanism	Section 3 Section 5.4 Section 5.12
	Functional Safety	Section 5.1
	Safety of the intended functionality	Section 5.2
	Cybersecurity	Section 5.3
ADS	Behaviour/Competencies	Section 3.3 Section 4.3 Section 5.4 Section 5.12
	Minimum Risk Manoeuvres/ Conditions	Section 5.5
	EE system	Section 5.1 Section 5.2 Section 5.3 Section 5.4
	Machine Learning	Section 5.8
	Other systems	Section 5.4 Section 5.10
Non-ADS	General	Section 5.10

	EE system	Section 5.1 [WP 4]
	Other systems	[WP 4]
Goods		Section 3.3 Section 3.4
Passengers		Section 3.3 Section 3.4 Section 5.7 [WP4]
Road Ecosystem	Operational Design Domain incl. other road users, road infrastructure and environmental conditions	Section 3.3 Section 3.4 Section 4.1
	Target Operating Domain	Section 4.1 Section 4.2
Off-vehicle systems	external systems	Section 5.6
	Remote operator	Section 5.3 Section 5.6 Section 5.7
Pre-Approval		Section 3.1.1
Development Phase	general	Section 7.1
	Design	Section 5.1 Section 7.1
	Implementation	Section 5.1 Section 5.8 Section 7.1
	V&V	Section 5.9
Regulatory Application		Section 3.1 Section 7.2
Operational Licensing		Section 3.1.3 Section 6.1
Deployment Phase		Section 4.2 Section 6.1
Operational Phase	General	Section 6.2

	Monitoring	Section 3.1.4 Section 6.2 [WP5]
	Response	Section 3.1.5 Section 5.11
	Change	Section 3.1.6 Section 6.2 Section 7.1 [WP5]
Manufacturer	Manufacturer Safety Management System	Section 7.1

Table 2: Section References. Entries in square brackets denote where the topic is addressed within the outputs of other work packages within the wider project.

Overall, the document could be summarised as consisting of the following areas:

- A description of the overall process flow for how the scheme would work and what would be submitted within the safety cases provided by the relevant organisations (Sections 3.1 and 3.2)
- A description of the high-level safety goals that all vehicle types should be expected to meet (Sections 3.3 and 3.4). These are, by nature, abstract, and would need to be supported by lower-level requirements that are bespoke to a vehicle type (addressed later in the document).
- A description of how to define the key attributes of the system and deployment, including what behaviours the system is required to provide and in what operating environment. Such documentation provides no assurance of safety in its own right, but is a vital intermediate step to allow safety evidence that is appropriate to the particular vehicle type and deployment to be developed downstream (Section 4)
- A description of what steps will be required to assure the safety of the ADS-equipped vehicle type itself, including consideration of aspects that are already covered, or partially covered, by existing regulations and standards (e.g. functional safety, cybersecurity) and areas that need bespoke solutions for LSAVs (e.g. scenario-based testing, safety of machine learning). An approach to developing performance requirements that are specific to the vehicle type but fulfil the more abstract safety goals described above is also included (Section 5)
- A description of the operational safety measures that should be put in place around the vehicle to mitigate risks during the deployment (Section 6)
- A description of other documentation that should be provided to regulators to assure safety, including evidence of appropriate safety management systems. This section also includes a summary of the key items of evidence that will need to be produced within the safety and security assurance process (Section 7)

### 3 Approval Mechanism

This section describes the overall approval mechanism for development of low-speed automated vehicles (LSAVs) and for their operation. It provides an overview of the entire approval process, defining the key artefacts that support regulatory approval and setting out the safety objectives and criteria on which approval will be based. A summary of the content of safety case reports and safety management systems is also provided here; further detail on these aspects is provided in the subsequent section, which provide in-depth analysis topic-by-topic.

#### 3.1 Approval Process

An overview of the approval process for LSAVs is shown in Figure 3 (overall view) and Figure 4 and Figure 5 (expanded views). The process is divided into three “swim lanes” flowing left to right, representing the key stakeholders in the approval process. These are:

1. Operator - the organisation providing services with the vehicles.
2. Manufacturer - the developer of the vehicle who submits it for approval.
3. Regulator - the body or bodies responsible for approval of the vehicle and approval of its operation.

These are top-level stakeholders, and it is recognised that the overall approval ecosystem will be more complex. For example, it is likely that, rather than the manufacturer producing every subsystem and component, there will be an extensive supply chain involved in developing the vehicle. However, from an approval point of view, the key stakeholder is the manufacturer, who will submit the vehicle for approval and act as it’s representative to the regulator. Similarly, it is possible that there will be more than one regulator, as the Law Commissions (2022) have recommended the establishment of an “approval authority” and an “in-use regulator”. Also, it is possible that the manufacturer and operator will be the same entity; however, they are shown separately here as the roles and responsibilities are different. Finally, from an approvals perspective, the most important artefacts are those at the boundaries between the stakeholders, e.g., safety case reports.

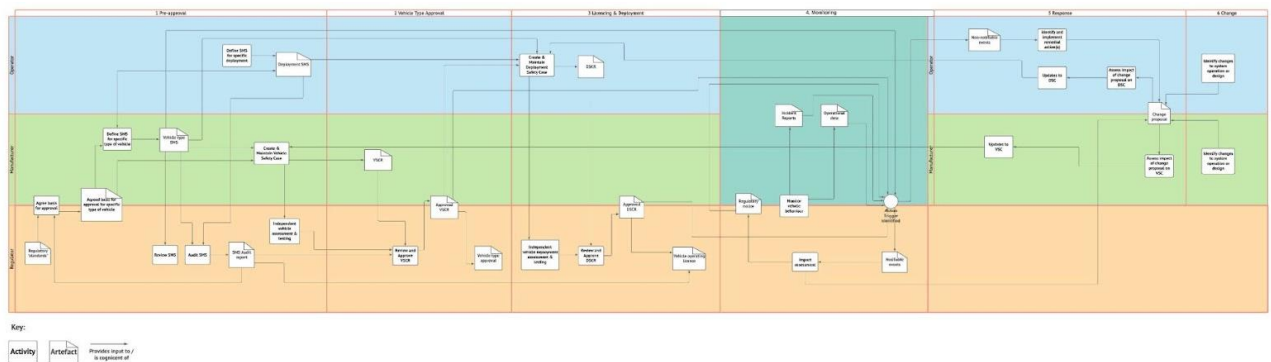


Figure 3: Overview of a potential approval process for AVs (see also embedded file).

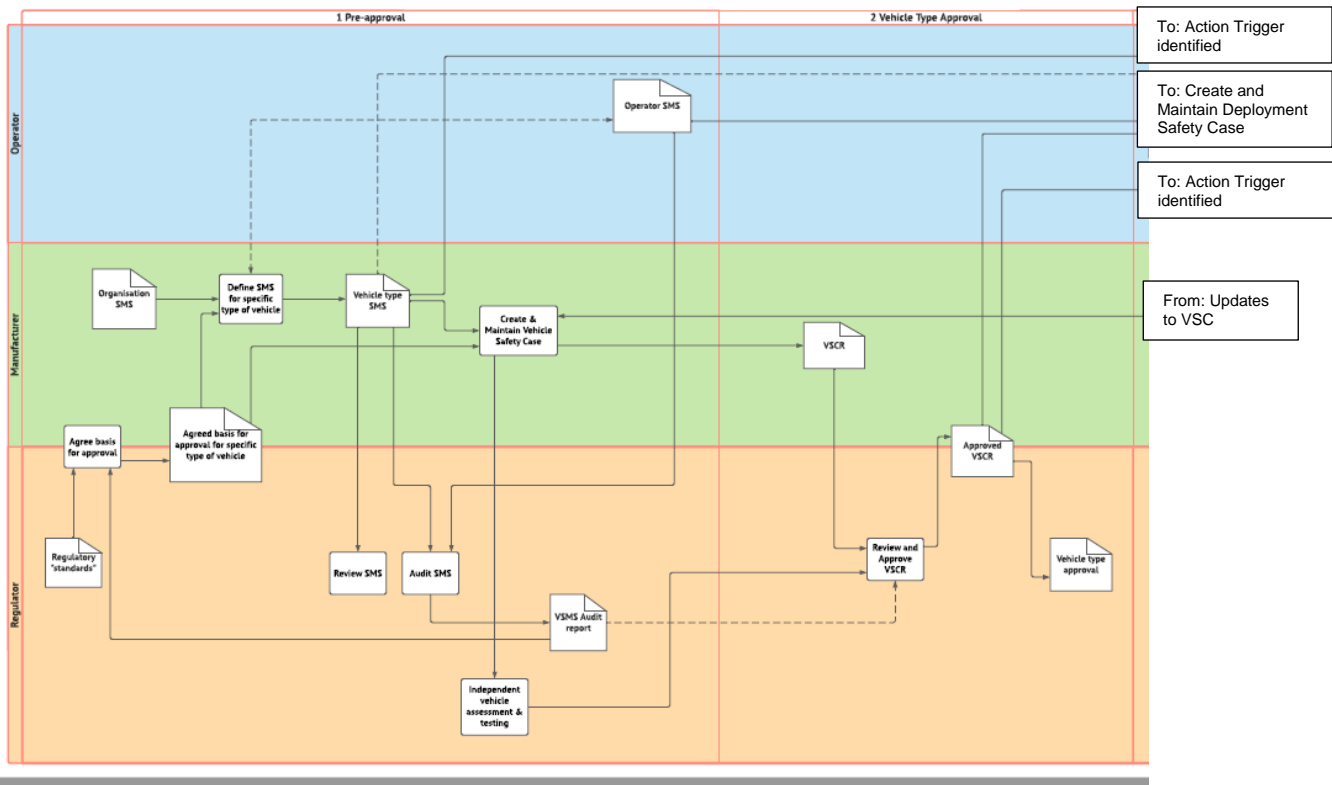


Figure 4: Detailed view of the process from commencement (1) Pre-approval, up to (3) Regulatory Application.

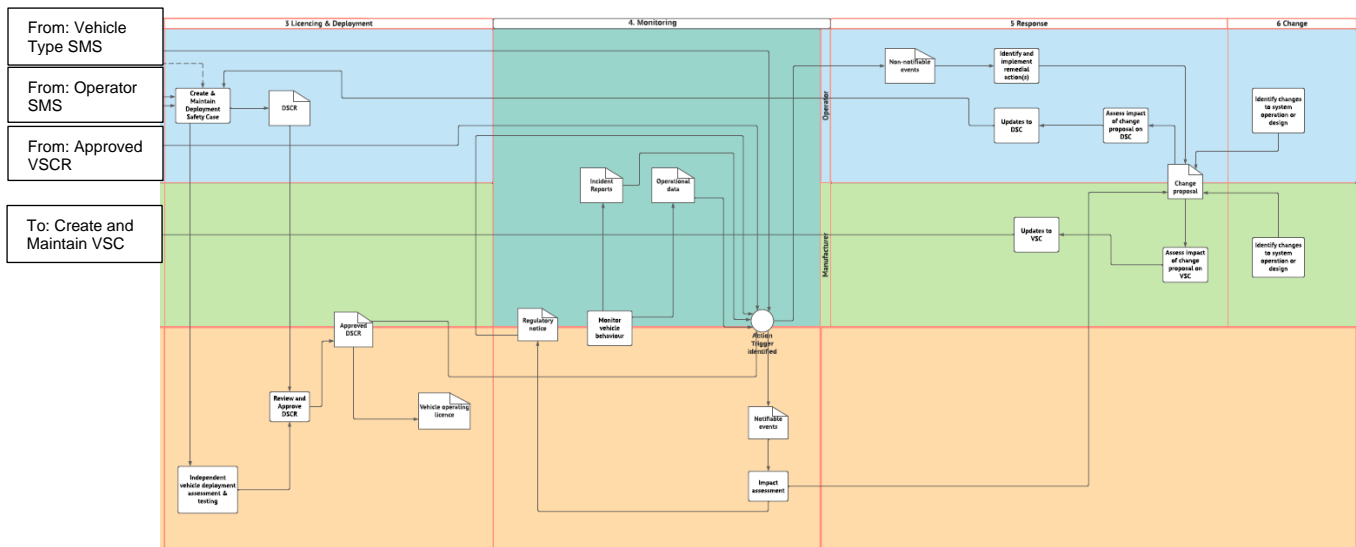


Figure 5: Detailed view of the process (cont'd) from (4) Deployment, up to (7) Change.

The approval process is also divided into phases reflecting the complexity of safely managing LSAVs. The phases are as follows:

1. **Pre-Approval** – the engineering activities that produce a vehicle and a safety case report to support an application for approval.
2. **Vehicle Type Approval** – the formal process of assessing a vehicle and the safety case report and granting or withholding approval.
3. **Licensing and Deployment** – the formal process of authorising (or withholding authorisation) for deployment of a vehicle, followed by vehicle deployment.

4. **Monitoring** – observing vehicle behaviour and incidents in operation and deciding whether remedial action is required.
5. **Response** – identification of engineering or operational change to implement remedial action or in response to desired changes to improve capability (actual changes are in other parts of the flow, e.g., pre-approval for engineering changes to the vehicle).
6. **Change** – proposals for change to improve vehicle capability or service.

The full engineering and support process for a vehicle is very extensive, and the approval process shows only those that are particularly pertinent to approvals. Details of the key activities and artefacts in the approval process are described phase-by-phase in the remainder of this section. Due to their central role, the safety case reports and SMS are described in more detail; please refer to section 3.2.

The approval process is shown as if it is a “flow” largely from left to right. In practice, there will be iteration. There are some explicit feedback links, e.g., from the **Response** phase into **Pre-approval** to deal with a recognised need to address operational shortcomings. There will also be other forms of iteration, e.g., between **Pre-approval** and **Regulatory Type Approval** as and when vehicle capabilities are significantly modified, but the potential iterative loops are elided to avoid cluttering the diagram. Furthermore, it may be desirable in some cases, particularly where the manufacturer and operator are the same entity, for the ‘vehicle type approval’ and ‘licensing and deployment’ activities to run concurrently. As such, the process flow diagram should be seen as broadly indicative, for informative purposes, but should not form an absolute requirement.

### 3.1.1 Pre-approval

The key *interface* is between the *manufacturer* and *regulator*, and involves the activity *Agree basis for approval*, resulting in the artefact *Agreed basis for approval for specific type of vehicle*. This process and its result would identify the type of evidence required in support of the safety case, e.g., what proportion of test evidence should be deployment site specific rather than generic (see Section 4.1 for further detail on ‘specific’ and ‘generic’ testing in the context of this report), how testing is balanced against analysis, or identification of applicable regulations of standards (such as UNECE regulations). The agreed basis would also reflect the content of the manufacturer’s SMS and, for example, identify the way in which change is handled, including when the *VSCR* (vehicle safety case report) would be re-issued. The intent of this activity and the artefact is to minimise the risk of disagreement on scope, contents, etc. whenever a vehicle safety case report is submitted for approval.

The *manufacturer* will, of course, design and produce the vehicle in conjunction with the wider automotive supply chain. From the approval perspective, however, the focus is on the manufacturer’s SMS and safety case. It is expected that the manufacturer will have to produce the *Vehicle type SMS (Define SMS for specific type of vehicle)*, perhaps based upon an Organisation SMS which is tailored for the specific vehicle. This tailoring is likely to be needed to reflect the way safety is managed in development, including on-road testing, how quality assurance is done in the supply chain, and so on. The manufacturer will *Create & Maintain (the) Vehicle Safety Case*, for the specific vehicle type, reflecting the *Agreed basis for approval for specific type of vehicle*. The *VSCR* is supplied to the regulator. The whole safety case is also made available on an audit sampling basis; see Section 3.1.2 on vehicle type approval.

The *regulator* will identify *Regulatory Documents* with which vehicles must conform, e.g., UNECE regulations software updates, and also applicable standards (e.g. ISO 26262). Key activities include *Review SMS* to support defining the *Agreed basis for approval for specific type of vehicle* and *Audit SMS* producing a *SMS Audit Report*, which will support the approval process. Visibility of the SMS in this way will enable the regulator to better judge the depth of evidence required for the *VSCR*. The regulator may also carry out *Independent Vehicle Assessment & Testing* or arrange for it to be carried out. This is linked to the activity to *Create & Maintain (the) Vehicle Safety Case* and will inform the regulatory decision-making process.

For the *operator*, the only pre-approval artefact is the *Deployment SMS*; this relates to the manufacturer’s SMS as it will, for example, need to show how necessary inspection and maintenance activities for the vehicle are supported across the deployed fleet. It will be informed by *Define SMS for specific type of vehicle*, and may even be produced within the same step, in the case where the *manufacturer* is also the vehicle *operator*. It should be noted that the identity of the operator may be

unknown at the pre-approval stage; should this be the case, it is permissible for the Deployment SMS to be provided later in the process, although it is advised that this should be undertaken as early as possible.

More detail on safety case reports and SMS is found in Sections 3.2 and 7.1.

Whilst this is referred to as a **Pre-approval** phase, the activities will be ongoing as the vehicle capabilities are updated. Further, it is anticipated that the audit of the *Vehicle Type SMS* and *Deployment SMS* will be conducted periodically during the operational life of the vehicle/service.

### 3.1.2 Vehicle type approval

The key interface between the *manufacturer* and the *regulator* is an *Approved VSCR* that documents the *regulator's* approval of the *VSCR* submitted by the *manufacturer*.

The *manufacturer* will produce the *VSCR*, extracting information from the overall vehicle safety case as documented in the *Agreed basis for approval for specific type of vehicle*. See Section 3.2 for details of the scope of the safety case and the *VSCR*.

The regulator will *Review and Approve the VSCR*, taking the submitted *VSCR* and the *Independent Vehicle Assessment & Testing* as inputs. The evaluation of the *VSCR* will be based on the Safety Goals and Risk Framework, and on the Acceptance Criteria (see Sections 3.3 and 3.43.4 respectively). The *Approved VSCR* will also be used to support the *Vehicle Type Approval*. It is anticipated that the type approval process for the physical vehicle will remain largely unchanged, and that the *Approved VSCR* for the ADS capabilities of the vehicle will be treated as supplementary material (see also the discussion of administrative requirements in Section 7.2). It should be noted that the 'approval' of the safety case may take a different form in terms of the processes and legal framework in comparison to 'approvals' granted within the existing type approval mechanisms.

It is expected that there will be iteration between the phases, as noted above; iteration between **Pre-approval** and **Vehicle Type Approval** could occur for several reasons. First, an initial application for approval might not be accepted, and this is likely to result in the production of an updated *VSCR* which is then re-submitted<sup>1</sup>. Second, there will sometimes be iteration when vehicle modifications are made. Some modifications are likely to be judged significant enough to warrant further safety engineering work and a re-issue of the *VSCR* for regulatory approval before the modifications can be (permitted to be) deployed; others will be able to be accepted based on the approved processes in the SMS without an immediate update to the *VSCR*. The Safety Goals and Risk Model, and the Acceptance Criteria, will provide the basis for determining whether a change will be deemed significant enough to lead to an update to the *VSCR*; ideally, the agreed interpretation of these factors will be included in the *Agreed basis for approval for specific type of vehicle*.

### 3.1.3 Licensing and Deployment

The key interface between the *operator* and the *regulator* is an *Approved DSCR* that documents the *regulator's* approval of the *DSCR* submitted by the *operator*, and thus gives approval to the *operator* to provide the planned service. In general, there is likely to be a licensing process which addresses economic considerations etc. as well as safety concerns. The approval of the *DSCR* forms the safety component of this wider licensing process. Note that some of the wider issues relating to licensing are addressed in the report by the Law Commissions (2022), including consideration of non-safety issues that lie outside the scope of this report. The term 'approved' is intended to be interpreted flexibly here, and the approval may be very different in nature to approvals within existing type approval, depending on the administrative and legal frameworks adopted.

---

<sup>1</sup> The initial review may also result in further development of the ADS which would result in an updated safety case



The operator will *Create & Maintain (the) Deployment Safety Case*, for the specific vehicle type, in the intended deployment domain. As with the vehicle, a deployment safety case report (*DSCR*) will be supplied and the whole safety case can be sampled by audit. Details of the scope and content of the *DSCR* can be found in Section 3.1.3, but it should be noted that a key component of the *DSCR* will be the compatibility of the *TOD* (target operating domain, i.e. the domain in which the system will be deployed) with the *ODD* (assessment of such compatibility is further detailed in Section 4.2).

The *regulator* is responsible for *Independent Vehicle Deployment Assessment & Testing*, which will include driving on particular routes in the intended deployment domain. The *regulator* will *Review and Approve the DSCR*, taking the submitted *DSCR* and the *Independent Vehicle Deployment Assessment & Testing* as inputs. The approval of the *DSCR* will be guided by the Safety Goals and Risk Model and the Acceptance Criteria, although it is anticipated that these factors will be more significant in the initial vehicle approval.

It is expected that there will be iterations of the **Licensing & Deployment** phase. As well as iterations prompted by vehicle changes including responses to operational shortcomings, it is likely that there will be independent operational changes, e.g., addition of new driven routes for an autonomous bus service, or changes in training for maintenance staff. As with the vehicle, some of these changes will warrant re-licencing whereas others will be acceptable based on the *SMS* (which is audited in the **Pre-approval** phase).

Note that, if the vehicle *manufacturer* and *operator* are the same organisation, some of the activities in the **Pre-approval**, **Vehicle Type Approval** and **Licensing & Deployment** phases might be merged for efficiency. However, it should be noted that the type approval and licencing are very different in nature. Type approval is judging whether or not a vehicle is fit to be on the roads, in some defined operating environment; it is vehicle-centric. On the other hand, licencing is organisation- centric; it is saying that arrangements including training of operators, maintenance processes, and protocols for incident management, etc. are sufficient to preserve safety in operation. Thus, the **Vehicle Type Approval** and **Licensing & Deployment** phases have intrinsically different scope, even though they will relate to the same vehicle type.

### 3.1.4 Monitoring

For the **Monitoring** phase, the ‘swim lanes’ for the *manufacturer* and *operator* are combined. This is not because they might be the same organisation, although they indeed may, but because it is unclear which organisation would carry out particular activities and produce particular artefacts in practice.

For example, it might be that all vehicle telemetry goes to the vehicle *manufacturer*, who then notifies the *operator* of issues that require attention, e.g., maintenance to replace a faulty sensor. Alternatively, telemetry might go to the *operator* who then notifies the *manufacturer* when appropriate, e.g., a sensor is routinely going out of calibration. Where such responsibilities are discharged should be determined by the commercial arrangements between the *operator* and *manufacturer*; the precise allocation of responsibility is immaterial for the safety assurance provided that all responsibilities are fulfilled, and all parties understand who is responsible for each obligation. This is further examined in Section 6.2 (Post-Deployment).

There are three items at the regulatory interface: *Regulatory Notices*, *Monitor Vehicle Behaviour* and *Action Trigger Identified*. Before describing these, it is worth amplifying the concept of *Notifiable Events*. Vehicles will suffer failures and malfunctions, e.g., a camera not being properly aligned after a windscreen is replaced. Many will only require “local” action, e.g., a maintenance activity to return the vehicle to a state consistent with the *Vehicle Type Approval* (and similarly with regard to licence conditions). However, some events will be significant enough to require remedial action; for example, failure to correctly identify certain signs. The *Notifiable Events* are those that are significant enough to warrant further investigation and potentially remedial action. As shown in the approval process, the *Notifiable Events* are actual events that occurred; there is likely to be merit in defining the type of event in the *Agreed basis for approval for specific type of vehicle* or in other regulatory documentation, e.g., the *VCSR* or *DCSR*. However, it is unlikely that it will be possible to have full foresight in defining what is notifiable, hence these events are an output of a collaborative process at the interface between the *regulator* and *operator/ manufacturer*. The need for a reporting process to trigger updates is again further examined in Section 6.2.

The *Monitor Vehicle Behaviour* activity is shared, although the onus is likely to be on the *operator/manufacturer*. The regulator will have a role both in terms of observing operations as vehicles are deployed to ascertain the safety of operations, but also as they may be the recipient of reports from the public (not shown in the diagram for simplicity). However, this activity would mostly depend on the vehicle and operational monitoring (see below) to identify events, including driving incidents and accidents, requiring further action.

The *Action Trigger Identified* is the one action in the approval process. It is triggered by *Incident Reports* identified as coming from *operator/manufacturer* analysis of vehicle behaviour, but which could also come from police, etc. (again elided for simplicity). Analysis of (trends in) *Operational Data* could also be action triggers - for example where it is clear that there is a high rate of 'false positives' leading to unwanted emergency braking. The analysis of these "dynamic inputs" will also be shaped by criteria and processes in the *Operator SMS*, the *Approved VSCR* and the *Approved DSCR* as, for example, they will give a baseline against which to judge whether or not a trend in data is statistically significant.

The *Notifiable Events* will then be the subject of *Impact Assessment* which will assess the severity and immediacy of the issue and determine the form of response. This might include issuing a *Regulatory Notice* limiting operation - this could reduce the ODD or deployment domain physically, e.g., avoiding routes with level crossings, or in other ways, e.g., not operating after dusk. The intent is that the *Regulatory Notice* would have legal force, qualifying the *Vehicle Type Approval* or *Vehicle Operating Licence* such that the *operator/manufacturer* would have to limit operations until remedial action had taken place (and been re-approved, as necessary). Note that there is a 'loop round' through the *Approved VSCR* and the *Approved DSCR* meaning that updates to these documents could lead to the rescinding of a *Regulatory Notice*. The **Impact Assessment** also informs the response, see 3.1.5 below.

As with other phases of the approval process, there will be iteration. Some iterations may need to be fast, to issue a *Regulatory Notice* as soon as possible after a serious *Notifiable Event* has been identified. On the other hand, remedial action might be quite slow, and it may be several months before a suitable engineering change can be designed, implemented and assessed. This would lead to an update of the *Approved VSCR* and the *Approved DSCR*, allowing the previously-licensed operation to recommence (or to operate without restrictions under a *Regulatory Notice*). Analogies can be seen in the aerospace industry, most recently with the problems related to the Boeing 737 Max.

### 3.1.5 Response

The **Response** phase is concerned with change to the system, in terms of responding to identified problems and implementing desired improvements (see the **Change** phase, Section 3.1.6). The activities and artefacts are all the responsibility of the operator and manufacturer, as regulatory approval is achieved by iterating through the **Pre-approval**, **Vehicle Type Approval** and/or the **Licensing and Deployment** phases.

The key interface item is the *Change Proposal*, whether it is prompted by *manufacturer* activity or by *operator* activity. This definition will contain a set of coordinated changes in vehicle design and in operation of the vehicle; in some cases, vehicle changes might not need an operational change and vice versa. However, this is viewed as a key interface between the *operator* and the *manufacturer*, as it is essential that the operator is made aware of design changes to assess their impact. Whilst it is less obvious, the same is true for operational changes as they may run counter to (perhaps undocumented) assumptions made by the *manufacturer* when designing and demonstrating safety of the vehicle.

The *manufacturer* needs to analyse change proposals to *Assess impact of change proposal* on the VSC and make the corresponding *Update to VSC* (and implicitly the *VSCR* as necessary). (One of the important aspects of the analysis would be to determine the extent to which the change does need to be reflected in the *VSCR*; see the discussion in 3.1 above and the discussion of Safety Goals and Risk Model, and the Acceptance Criteria, in Sections 3.3 and 3.4.3.4 respectively). There may also need to be engineering work undertaken by the manufacturer to implement changes to the vehicle, to carry out safety analysis etc., also requiring changes to the VSC. As indicated earlier, such engineering activities are not made explicit in the approval process, as the focus is on regulatory approval.

The operator needs to analyse change proposals to *Assess impact of change proposal on the DSC* and make the corresponding *Update to DSC* (and implicitly the *DSCR* as necessary). Like the *manufacturer*, the *operator* also needs to be made aware of the associated changes - but in this case they will be to

aspects of *Operator SMS*, e.g., altering monitoring criteria or maintenance regimes. The swim lane for the operator also shows *Non-notifiable Events* and the associated *Identify and implement remedial action(s)*. Based on analysis of operational data, the operator may decide that they want to make operational changes and/or ask for modifications to the vehicle to provide better services, to improve the economic return etc. Such changes might have unintended consequences on safety, so they need to be recorded in a *Change Proposal* and assessed.

### 3.1.6 Change

The *manufacturer* and *operator* might identify the need for change to the vehicle or its design (*Identify changes to system operation or design*), for upgrades including those to pre-emptively fix issues identified by them, and very likely to improve capabilities of the vehicle or to offer new services using the vehicle, etc. These identified changes will be treated in the same way as responses to the need to change arising from operational shortcomings, see Section 3.1.5.

## 3.2 Use of a Safety Case and Safety Argument

This section is concerned with the key artefacts in the approval process - the safety case - and the SMS which support the safety case. The safety case reports give snapshots in time of the arguments and evidence which are intended to demonstrate the safety of the vehicle and of the operation, respectively. The SMS complements the safety case both by showing how evidence to support the safety argument is produced and how safety will be managed on an ongoing basis during operation. The two are complementary, and serve as the bedrock of the approval process.

This section details the objectives for the vehicle safety case, deployment safety case, and the safety management systems used to support development and deployment. The role of these artefacts in the approval process is described in Sections 3.2.2, 3.2.4 and 3.2.7. For each of these artefacts, evidence shall be provided to show that each of the objectives is met. The description of the safety case and safety management system is provided in a form proposed requirements that would be suitable for transcription into secondary legislation supporting the recommendations of the Law Commissions requiring a safety case. It uses the terms:

Shall = Must be achieved to demonstrate compliance

Should = Recommended to be achieved. Alternate means are acceptable.

Part of the role of the *Agreed basis for approval for specific type of vehicle* is to set out the means by which these objectives will be met, i.e. the nature and extent of evidence to be provided. It is also expected that this artefact will define what is delivered in the safety case reports (*VSCR* and *DSCR*) and what evidence remains with the *manufacturer* or *regulator* but is available to be inspected by the *regulator*; this is a similar philosophy to that adopted in various UNECE regulations, e.g. for ALKS (2021).

### 3.2.1 Nature of the Safety Case and Safety Case Reports

A safety case is a widely-used method to provide evidence that something has achieved an acceptable level of safety. Within the automotive industry, a safety case is required in order to evidence the functional safety (ISO 26262, 2018) and the 'safety of the intended functionality' (ISO/PAS 21448, 2022), but safety cases are also widely used in other industries, with the advantage that they allow the flexibility for an organisation to develop, and argue, its own approach to evidencing safety. Whilst the relative stability of established vehicle safety technology allows more prescriptive 'type approval' tests to be applied for aspects such as passive safety, newer technologies that haven't yet reached such a point require the greater adaptability provided by a safety case.

As such, the information contained within a safety case will vary significantly depending upon the nature and complexity of the system and deployment. Furthermore, those creating the safety case may choose different approaches to subdivide this information into separate documents. In theory, a safety case could even provide all the necessary information within a single document, although this could make it difficult to engage with; for example, updates to one aspect would require the change management process to be applied to the whole safety case. More typically, the safety case would consist of multiple documents, each covering a particular aspect (Zenzic, 2021).

As described within Section 3.1, this report proposes that, rather than the manufacturer or operator submitting their full safety case, they should instead submit a 'safety case report'. This does not absolve them from the responsibility to maintain a complete safety case as part of their own safety management practice and as a means to provide detailed evidence to support incident investigations where necessary, but allows the information that is provided to the regulator to be focussed upon the pertinent facts; given the complexity of automated vehicles and their operating environments, it is anticipated that it may be impractical for regulators to scrutinise the full safety case. However, where the regulator identifies a need for further information beyond that provided within the safety case report, they should be able to require access to the full safety case, or to the relevant components of it.

The level of detail provided within the safety case report should be proportionate to the complexity and level of risk posed by the system and deployment. As there is no objective means to specify what level of detail would be appropriate for any given application, it is recommended that discussions on this between the regulator and the organisation creating the safety case report should commence as early

as possible such that an agreement on a suitable methodology can be reached. Factors that should be considered to determine the level of detail required include:

- The complexity of the deployment environment (the 'ODD' and 'TOD', as defined in Section 4.1) – for example, a system intended for use within busy city streets may need a more detailed consideration of the hazards presented by other road users than one intended to operate upon its own segregated lane.
- The operational safety measures (see Section 6) that will be placed around the vehicle to manage the hazards that it may face – for example, a system that has an onboard customer assistant may need less consideration of how the system would ensure safety of passengers.
- The complexity of the functionality that the system is able to provide (the 'Behavioural Competencies, as defined in Section 4.3) – for example, a system that operates upon a continuous loop with no ability to navigate junctions may be expected to need less analysis and testing of its behaviours than a system that can negotiate a range of complex urban routes.
- The level of exposure to risk – for example, for a small pilot scheme using a small number of vehicles on a short route with low exposure to other road users, a less detailed analysis may be proportionate in comparison to a deployment with more vehicles over a longer and busier route.

An essential component of any safety case or safety case report will be the safety argument; this provides a justification for how all the evidence presented within the safety case, when taken together, supports the overall goal of acceptable safety. Without a safety argument, the safety case would merely be a mass of disjointed information, with no means to understand how it fits together, and no means to identify any gaps where there is insufficient evidence.

Goal Structuring Notation, or GSN, is a widely used format for presenting a safety argument, allowing it to be displayed graphically with the overall safety goal at the top of the hierarchy and other sub-goals arranged beneath to support it. These are progressively broken down until sufficient granularity is reached where specific pieces of evidence ('solutions') can be provided in support, with additional syntax included in the standard to allow diagrams to be annotated with key aspects such as assumptions and context (GSN, 2018).

It should not be mandatory to use GSN, as the safety argument could be conveyed by other means, including descriptive text. Nevertheless, it is important for the safety case to include some means to explain how the evidence fits together to form a complete and cohesive safety argument (Zenzic, 2021).

The remainder of Section 3.2 summarises some of the key aspects that should be covered within the safety case reports produced by the manufacturer and the operator; more detailed guidance on each of these areas is then provided within the subsequent sections, in order to set out a framework for what should be included in the information submitted to the regulator.

### 3.2.2 Vehicle Safety Case Objectives

A safety case shall be created to demonstrate that the vehicle is sufficiently safe to operate throughout its entire operational life. An argument shall be created and supported by providing evidence to show that each of the following objectives is met. The rigour of the evidence generated shall be proportionate to the overall assessed safety risk and complexity of the vehicle operation (as was described in Section 3.2.1) and as defined in the *Agreed basis for approval for specific type of vehicle*. The evidence shall be documented as part of a safety case and the agreed subset delivered to the regulator as part of the VSCR. The top-level objectives apply regardless of implementation technology, and more detailed requirements for Machine Learning (ML) components are provided in Section 3.2.3. In some cases, the objectives link back to other activities in the approval process, e.g. the **Monitoring** phase, rather than the **Pre-Approval** phase.

*Note: As described in Section 4.1, this report draws no conclusion as to whether the scenario-based testing evidence should be provided within the vehicle safety case report or the deployment safety case report. Section 4.1 identifies that the scenario-based testing would need to include significant evidence that is specific to the actual deployment routes(s), and that it therefore requires the scope to be defined by a 'TOD' (Target Operating Domain), which describes the specific deployment route(s) rather than an*

*'ODD' (Operational Design Domain), which can optionally be described on a more generic basis. If the decision is made that it is more practicable for the scenario-based testing to be assessed via the vehicle safety case report, it would therefore be necessary for both the ODD and the TOD to be provided here; thus, if this approach is chosen, references to "ODD" within the below requirements should be read as "TOD" where the clause relates exclusively to scenario-based testing, or as "ODD and TOD" where the clause covers to multiple aspects of vehicle safety assurance that would include, inter-alia, scenario-based testing. However, if the decision is for the scenario-based testing to form part of the deployment safety case, the below requirements should be interpreted as written.*

1. The automated driving capabilities provided by the vehicle shall be defined and documented.
2. The intended Operational Design Domain (ODD) shall be defined and documented.
  - 2.1. The ODD definitions shall include all relevant features defined to an appropriate level of detail (see further guidance in Section 4.1).
  - 2.2. The ODD definitions shall be validated to determine that it is sufficiently complete.
3. The behavioural competencies (i.e. the behaviours that the vehicle will be required to perform) relevant to the vehicle operating within the defined ODD shall be defined.
  - 3.1. The behavioural competencies shall be sufficiently comprehensive and complete.
  - 3.2. The scenarios that the system encounters within the test programme and in service will be a function of the behavioural competencies that the system must perform and the operating environment in which it must perform them.
4. Hazards relating to the automated driving capabilities and the operating scenarios within the defined ODD shall be identified and documented. The hazards here may include contributions from various systems e.g. failure of an automated driving system (ADS) or failure of the braking system.
5. Efforts shall be made to identify hazards resulting from operation outside the ODD boundary, including consideration of transitions across the ODD boundary. Mitigations shall be documented to ensure operation outside the ODD is minimised as far as is reasonably practicable and that any remaining excursions are acceptably safe.
6. A safe operating concept (SOC) shall be specified that is sufficient to mitigate the identified hazards within the ODD, and provide an appropriate response should the vehicle exit the ODD.
  - 6.1. The SOC shall define safety requirements that are sufficient to mitigate the hazards associated with performing the defined behavioural competencies within the defined ODD.
  - 6.2. The SOC should include any conditions requiring restricted operation within the ODD, with justification.
  - 6.3. Justification shall be provided that the defined safety requirements are sufficient to mitigate the hazards within the ODD.
  - 6.4. Justification shall be provided that the behaviour is acceptably safe in circumstances where the vehicle unavoidably exits the conditions permitted within the ODD definition such that behaviour when outside the ODD, and when transitioning across the ODD boundary, is appropriate and, so far as is reasonably practicable, minimises risk.
    - 6.4.1. Mechanisms shall be put in place to (i) interpret the ODD boundary in a workable manner and to detect (ii) when the vehicle approaches the boundaries of the ODD and (iii) when the vehicle leaves or returns to the defined ODD.
    - 6.4.2. The behaviour shall be appropriate when it is outside its ODD, however briefly. This may for example involve transition to a degraded operating state that features a broader ODD, or the achievement of a Minimal Risk Condition (MRC) via a Minimal Risk Manoeuvre (MRM).

- 6.4.3. It shall be demonstrated that the vehicle remains sufficiently safe during the transitions from inside to outside the defined ODD and back again.
7. Justification shall be provided that the safety requirements specified at each level of the vehicle design adequately capture the intent of the safety requirements from which they derive. This should include consideration of the vehicle subsystems and architecture as well as its items and components.
8. Justification shall be provided that all design decisions taken are appropriate to ensure that the defined safety requirements, and thus SOC, can be met by the vehicle. This should include decisions taken in the design of the vehicle subsystems and architecture as well as its items and components.
  - 8.1. The design shall take account of potential hazardous failures that are identified.
  - 8.2. Where pre-existing components are used, their sufficiency with respect to the defined safety requirements, ODD and behavioural competencies shall be demonstrated.
9. Potentially hazardous failures relating to automated operation shall be identified throughout the lifecycle of the vehicle, including development, operation, maintenance and management of incidents. This should include consideration of those that may arise in the vehicle subsystems and from the architecture, as well as its items and components.
  - 9.1. A justification shall be provided for the sufficiency of the identification of potentially hazardous failures.
  - 9.2. Safety requirements shall be derived where necessary to address the identified potentially hazardous failures.
  - 9.3. A justification shall be provided to explain how the derived safety requirements are sufficient to address the identified potentially hazardous failures.
10. Verification evidence shall be provided that demonstrates that each of the safety requirements is satisfied.
  - 10.1. Justification should be provided that the verification performed is sufficient to demonstrate the satisfaction of the safety requirements.
  - 10.2. For evidence that is not collected directly from the vehicle in situations representative of the operating environment (such as simulations or lab tests), the representativeness of the evidence should be justified.
11. Validation evidence shall be provided to demonstrate acceptable vehicle behaviour and performance irrespective of the safety requirements and other system requirements.
  - 11.1. This may, for example, take the form of extended 'mileage accumulation' testing upon the trial route(s) or of scenario-based testing that provides coverage of the range of possible scenarios that could occur in the real world.
  - 11.2. Justification for the acceptability of the level of coverage shall be provided.
12. A sufficient safety management system (SMS) for the development and through-life support of the vehicle shall be defined and followed.
  - 12.1. Periodic audits of the SMS shall be undertaken throughout the lifecycle of the vehicle.

### 3.2.3 Machine Learning Safety Case Objectives

Where the vehicle includes components developed using Machine Learning (ML), the following objectives apply<sup>2</sup>. More detailed requirements and guidance relating to the use of ML are included in Section 5.8.

#### 3.2.3.1 ML Development Objectives

1. The system safety requirements allocated to the ML component shall be correctly identified and documented.
  - 1.1. The safety requirements should be generated from a system safety assessment process that covers hazard identification and risk analysis and determines the contribution that the output of the ML component makes to potential system hazards.
2. The context into which the ML component will be deployed shall be defined, including:
  - A description of the system into which the ML component will be deployed.
  - The ODD of the system.
  - The behavioural competencies performed by the system.
3. ML safety requirements shall be defined that capture the intent of the system safety requirements allocated to the ML component.
  - 3.1. The ML safety requirements shall be amenable to ML implementation and verification.
  - 3.2. A justification shall explain how the ML safety requirements were derived from the allocated system safety requirements.
  - 3.3. The defined ML safety requirements shall be validated with respect to the intent of the allocated system safety requirements and the results shall be documented.
4. ML data requirements shall be defined that enable the development of a machine learnt model that satisfies the ML Safety Requirements.
  - 4.1. The ML data requirements should include consideration of the relevance, completeness, accuracy and balance of the data sets.
  - 4.2. A justification shall be provided for the sufficiency of the ML data requirements.
5. Data shall satisfy the defined ML data requirements. The following data sets shall be defined:
  - Development data used for creating learned models
  - Internal test data used for testing learned models
  - Verification data used for verification of the learned models
  - 5.1. A data generation log should be maintained that captures and justifies decisions made during data generation.
  - 5.2. The generated data sets shall be validated to ensure they are sufficient to meet the ML data requirements; the validation activities and results shall be documented.
  - 5.3. The verification data set shall be generated independently of the development and internal test data sets.

---

<sup>2</sup> These objectives are derived from the AMLAS (2022) guidance which could be used as part of a means of compliance to these objectives.



6. The model learning process used and the type of model created shall be appropriate for the defined ML safety requirements and the task undertaken by the ML component.
  - 6.1. The process used in creating the model should be documented in a model development log along with a justification for all key decisions made during the learning process.
  - 6.2. Models created during the learning process should be evaluated using the internal test data to check they satisfy the ML safety requirements. The results of the internal testing should be documented.
7. The ML model shall be verified to demonstrate it satisfies the defined ML safety requirements.
  - 7.1. The verification of the ML component should be carried out independently from the development of the component.
  - 7.2. The verification results shall be documented.
  - 7.3. The verification process used and its rationale should be documented in a verification log.
  - 7.4. The testing environment for the ML component shall be sufficiently representative of the operational platform to which the component is deployed.
  - 7.5. Formal models used for verification shall be sufficiently representative of the target system and its environment.

### 3.2.3.2 ML Component Deployment Objectives

1. Measures shall be put in place to ensure the system safety requirements allocated to the ML component continue to be satisfied throughout operational life.
  - 1.1. Measures shall be put in place to monitor properties of the system that may affect the behaviour of the ML component, including:
    - External inputs to the ML component.
    - Outputs generated by the ML component.
    - Validity of assumptions regarding the system and its environment.
  - 1.2. Mitigations shall be put in place to address the risk posed to the system by any deviations in the monitored properties.
  - 1.3. The nature and characteristics of the predicted deviations in the monitored properties and their mitigations should be documented in a deployment log.
2. The integration of the ML component within the deployed system shall be tested in order to demonstrate the system safety requirements allocated to the ML component are met in the system context.
  - 2.1. The integration test results should be documented.
  - 2.2. The integration testing process used and its rationale should be documented in a deployment log.

### 3.2.3.3 ML Through-Life Objectives

1. The performance of the ML component shall be analysed throughout the operational life of the system.
  - 1.1. Measures shall be put in place to compare the performance of the ML component against the predicted performance.
  - 1.2. The analysis shall consider specific events observed during operation as well as analysing performance trends over time.

- 1.3. Where applicable, the analysis should consider the performance of ML components across multiple instances (for example the whole fleet of vehicles).
2. Where analysis identifies unsafe behaviour of the ML component, measures shall be taken to mitigate the risk to the system.
3. Where changes are required to an ML component, the safety impact of the changes shall be assessed.
  - 3.1. Changes should be agreed with the approval authority.
4. Where the safety impact of the change is determined to be significant, the changed ML component shall be considered as a new component and objectives shall be re-demonstrated.

### 3.2.4 (Operator) Deployment Safety Case Objectives

1. The target operating domain (TOD) for the deployment shall be documented in sufficient detail to enable assessment of the compatibility of the vehicle with the specific deployment route(s)/area(s).
2. The compatibility of the TOD(s) with the ODD shall be assessed.
  - 2.1. The basis for the compatibility shall be documented, including any interpretations and approximations used
  - 2.2. Any discrepancies shall be corrected (e.g. through updating the TOD), be mitigated (e.g. through restrictions of use), or be justified
3. Specific verification and validation testing shall be performed within the deployment domain(s)
  - 3.1. Verification and validation testing shall include road testing on representative routes in the deployment domain. This may include the actual route(s) or geofenced area(s), their 'digital twins' within simulations, mock-ups upon a proving ground, or potentially other test environments that can be argued to be representative of the scenario permutations and scenery that the system will encounter in service.
  - 3.2. Verification and validation testing shall cover specific features of the deployment location(s) (i.e. TODs) and operating scenarios not encountered in previous deployments. This should ensure that unique features of a particular deployment domain receive particular attention during verification.
4. Sufficient procedures and infrastructure shall be in place to support safe deployment of the vehicle.
5. A sufficient safety management system (SMS) for the deployment of the vehicle shall be defined and followed by the operator.
  - 5.1. Periodic audits of this SMS shall be undertaken throughout deployment.
6. Testing of the vehicle upon public roads or private facilities prior to approval of the system shall be conducted according to a safety case created specifically to cover such testing. Such a safety case shall be constructed in accordance with industry best practice with respect to such testing of automated vehicles, including conformance with the Code of Practice for Automated Vehicle Trialling (DfT, 2019).

### 3.2.5 Vehicle Type Safety Management System Objectives

- 1 The SMS shall define the competencies, processes, procedures, infrastructure and facilities needed to develop the vehicle for the intended deployments in a manner that is consistent with the vehicle and deployment safety cases. This should include:

- a. A process for the safety assessment of the design, verification of the design, and design change management relating to the vehicle, covering software, hardware, subsystems and data.
  - b. Procedures and mechanisms for responding to test failures, incidents, accidents and hazardous failures
  - c. Processes, procedures, competencies, certifications and training for vehicle design, manufacture, maintenance and upgrade activities
  - d. Processes for responding to directives from regulators, including making design changes and communicating to users/operators of the vehicles
  - e. Processes for updating the safety documentation to allow for regular review and re-issue as appropriate
- 2 The manufacturer shall put in place support mechanisms as part of the SMS. This should include monitoring of the vehicles in-service to ensure the continued validity of the vehicle safety case, including:
- a. Short-term monitoring (e.g. incident and accident monitoring)
  - b. Long-term reports (safety performance trends over time)
  - c. Comparison against analyses, models and predictions, with discrepancies explained
  - d. Where appropriate, analysis should consider multiple instances (e.g. across a fleet of vehicles).
  - e. Where appropriate, fleet and individual vehicle performance safety data shall be used to (i) issue alerts, (ii) initiate changes (e.g. to rectify a problem), (iii) update the safety case(s)
  - f. The safety analysis of the monitoring data shall be used to update the safety cases

### 3.2.6 (Vehicle Operator) Deployment Safety Management System Objectives

1. The SMS shall define the processes, procedures, infrastructure and facilities needed to operate the vehicle within the deployment domain consistent with the vehicle and deployment safety cases.
  - 1.1. This should include:
    - a. A process for safety assessment of changes relating to the vehicle, deployment routes and infrastructure
    - b. Procedures and mechanisms for responding to incidents, accidents and hazardous failures.
    - c. A process for management of specific restrictions, deviations and waivers covering the vehicle, infrastructure and routes.
    - d. Processes, procedures and training for vehicle maintenance and upgrades.
2. The operator shall put in place support mechanisms for the SMS.
  - 2.1. This should include:
    - a) Monitoring of the deployment routes to ensure they remain compatible with the TOD.
      - a. Any incompatibilities identified shall be appropriately addressed.
      - b. Arrangements with any authorities responsible for maintaining or altering roads within the deployment route(s) or area(s), such that infrastructure changes are notified or consulted upon in advance, shall be documented.

- b) Monitoring of the vehicle to ensure the continued validity of the vehicle safety case. Including:
  - a. Short-term monitoring (e.g. incident and accident monitoring)
  - b. Long-term reports (safety performance trends over time)
  - c. Comparison against analyses, models and predictions with discrepancies explained
  - d. Where applicable, analysis should consider multiple instances (e.g. across a fleet of vehicles).
  - e. Where applicable, fleet and individual vehicle performance safety data shall be used to (i) issue alerts, (ii) initiate changes (e.g. to rectify a problem), (iii) update the safety case(s)
  - f. The safety analysis information shall be used to update the evidence in the vehicle and deployment safety cases
3. Facilities and infrastructure required for vehicle communications, navigation and support shall be in place
  - 3.1. There shall be facilities for vehicle data reception and data storage
4. Where applicable, vehicle remote assistance and recovery procedures shall be in place

### 3.2.7 Additional Objectives for Deployment SMS

1. Communications and data related to the vehicle and its deployment shall be stored and retained.
  - 1.1. Information relating to the configuration of vehicle hardware, software and data shall be retained (e.g. software updates, sensor replacements)
  - 1.2. Data relevant to safety shall be verified and retained for a specified duration. This is to facilitate long-term monitoring and support accident investigations.
  - 1.3. Where applicable, vehicle connectivity to remote monitoring systems shall be assured
  - 1.4. Where applicable, data retention and storage integrity at remote sites shall be assured
  - 1.5. Where connectivity is needed to support a safety function (e.g. emergency services call after accident), then this shall be assured in the context of the deployment TOD (e.g. in tunnels on routes)
  - 1.6. Location and positioning functionality shall be assured in the deployment TOD (e.g. GPS positioning around tall buildings)
2. The safety impact of any changes due to maintenance, reconfiguration, tailoring, upgrades or operating choices shall be assessed. Note these could affect hardware, software, data, training, maintenance or operating location.
  - 2.1. Changes shall be assigned a safety status (e.g. none, minor, major). Changes may (i) enable new functionality, (ii) disable existing functionality and (iii) change behaviours of the vehicle.
  - 2.2. Significant safety changes shall be agreed with the approval authority in advance of operation
  - 2.3. Operators and maintainers of the vehicles shall be informed of the impact of the change in advance of operation
  - 2.4. Additional training, guidance or warnings shall be in place in advance of operation
  - 2.5. Changes shall be applied in a controlled manner according to a defined process

- 2.6. Changes shall be monitored for adverse effects
- 2.7. Changes shall be able to be reversed by reverting to an earlier state (software, data)
3. Incident/accident management and response shall be according to a defined set of process and procedures
  - 3.1. The vehicle occupants, operator and emergency services shall be able to disable a vehicle at an accident scene (e.g. via a 'kill switch')
  - 3.2. Coordination with recovery operators and emergency services shall be in place. All those who come in contact with a recovered vehicle shall be aware of the automated functionality.
  - 3.3. The operator shall be able to (i) issue alerts, (ii) remotely disable particular functions, or (iii) disable a vehicle or whole fleet
  - 3.4. The operator and vehicle manufacturer shall support independent accident investigation bodies in their duties
  - 3.5. The operator and vehicle manufacturer shall respond to any recommendations from incident and accident reports
4. If applicable, remote assistance operation shall be assured
  - 4.1. If the vehicle has remote assistance functionality, then this shall be assured to work safely with the automated functionality
5. Management of system deviations and waivers shall be assured
  - 5.1. The set of agreed system deviations, restrictions and workarounds shall be managed
  - 5.2. The regulator, operator and vehicle manufacturer shall be aware of the status of deviations
  - 5.3. Formal waivers shall be issued by the regulator where required
6. Vehicle maintenance shall be carried out according to manufacturer specifications
  - 6.1. Staff used to perform maintenance shall be appropriately trained and identified as being suitably qualified and experienced personnel within a skills matrix or other such competency management system
  - 6.2. Parts shall be approved to confirm the correct parts are used, that they are used in accordance with their specification, and that they meet the required quality standards
  - 6.3. Re-calibration and re-testing shall be to manufacturer specifications
7. Procedures for ensuring the vehicle is in a safe state (e.g. in relation to park brake status or powertrain voltages) and for vehicle recovery shall be according to manufacturer specifications and local environmental and safety regulations
8. Vehicles shall be decommissioned and disposed of according to manufacturer specifications and local environmental regulations
9. A process shall be agreed with the authorities responsible for the routes to ensure that the routes are monitored and maintained according to appropriate codes of practice, rules and regulations, and changes will be assessed to ensure they remain within the ODD and TOD definitions.
  - 9.1. Changes will be notified to the operator

### 3.2.8 The Balance Between Auditing and Testing

It is envisaged that the overall regulatory assurance will be provided by a combination of the following approaches in order to arrive at an acceptable level of confidence that the system is appropriately safe and secure:

- Auditing of safety measures in place at the time of approval (e.g. redundant subsystems able to act as a fallback, operational safety mitigations in place on the deployment route).
- Auditing of safety management processes to support the updating of safety measures over time. This shall include:
  - Confirming that suitable processes were in place during the development and safety assurance of the system such that it is reasonable to trust that the evidence presented is acceptably complete and accurate.
  - Confirming that suitable processes are in place to identify the need for changes relating to safety and security over time, and to make those changes in an appropriate manner.
- Auditing of test evidence acquired by the manufacturer and provided to the regulator, including tests of individual subsystems/ components and also tests of the full vehicle within a representative operating environment.
- Regulatory testing, which is either independently carried out on behalf of the regulator, or witnessed by a representative of the regulator, to provide direct assurance of acceptable performance. This shall again include testing of individual subsystems/ components and also testing of the full vehicle within a representative operating environment.

Regarding the latter two points relating to testing of the system, both are able to provide coverage of the range of situations that could be faced by the vehicle, and its subsystems and components, within service, and as such there is no analytical means to determine what the relative proportion of each should be in order to gain a particular level of confidence that the system is acceptably safe. As such, this should be seen as a regulatory and political decision as opposed to a technical one.

Therefore, whilst this report recognises the need for regulators to determine an appropriate proportion of independent/ witnessed testing in order to provide suitable confidence to themselves and to the general public, this report does not attempt to propose any such proportions. However, it is advised that the precautionary principle should be applied such that AVs, and indeed any advanced AV test methods such as scenario-based testing within simulation, are required to be robustly assured through a significant level of testing and assessment prior to being trusted. This is important given the arguably unprecedented complexity of the systems and the environments that they operate in, the novel and unproven nature of many of the technology solutions, and the potential pressure to get systems to market before they are ready as a result of overly-ambitious promises and financial commitments. The safety and security scheme should therefore include significant volumes testing of the complete vehicle type within the physical world, conducted or witnessed by regulators.

### 3.3 Safety Goals and Risk Framework

The aim of this section is to outline a Safety Framework for the Low-speed automated vehicle approval scheme.

Overall, the aim of Automated driving technology is to improve road safety, and hence harm should be prevented wherever practicable, but the fact remains that accidents will still continue to occur. Therefore, the scheme needs to accommodate a level of imperfection, whilst setting criteria that ensure that the proposed low-speed automated vehicles are safe enough for use on public roads.

This “safe enough” criterion cannot consist of a single target number for fatal or severe accidents per distance driven or operating duration, as risk involved in road transport overall is comprised of a large number of different constituents and varies depending on, amongst others, factors such as:

- types of and intentions of vehicles and other road users participating in road traffic
- speed of participating vehicles
- road conditions (layout, safety features, surface conditions)
- weather
- duration that each individual is taking part in road transport.

Furthermore, road transport risk as an overall concept is also viewed differently by each individual taking part within it, with individual participants, to the degree that they can influence it, being able to accept

different levels of risk by either making adjustments to their behaviour or through decisions regarding how they participate in road transport. Individuals can make decisions around a number of existing safety measures that exist for road transport, which aim to achieve a risk level that is broadly accepted by society in general. This includes training for road transport participants, rules and processes for building and maintaining roads, type approval regulations to ensure vehicles are safe at their introduction and remain safe through regular mandatory checks.

These measures combine together to make the interaction of all parts and all participating actors, and hence the overall road transport ecosystem, acceptably safe for the general public. This is despite current accident rates being high compared to other transport modes such as air or rail, resulting in a continuous effort over time to reduce accident rates in road transport through a number of improvement campaigns. Using new technologies to reduce the occurrence or severity of accidents has been one of the measures used, and the introduction of automated driving is seen as having the potential to continue this trend.

Whilst a safety framework purely based on a top-level safety target like “twice as good as a human driver”, is easy to state, they cannot be quantified entirely (despite the existence of information such as the STATS-19 data) and assessed pre-deployment. Indeed, no examples could be found from any industry sector within the research for this report of any new transport system or vehicle type being required to, or able to, identify statistics for fatality and injury rates prior to commercial deployment at scale; as such, it must be concluded that comparing overall accident statistics for an LSAV type against a baseline would not be a practicable means to reach a decision on whether to approve or reject the safety of such a system. Furthermore, the lack of accumulated statistical data relating to AV use cases makes objective comparison of the safety of AVs in general against traditional manually-driven vehicles impractical.

Instead, a different approach is being proposed that sets out a number of high-level safety goals with the aim of avoiding harm, which combine to set the overall framework of how safe a LSAV should be

- (1) at the point of type approval
- (2) during its operation.

For type approval, the assessment is proposed to be against criteria that set out the objectives for safe behaviour in nominal, fault and threat condition, the evidence required to be shown and the processes to be followed to achieve the objectives. These criteria are set out in Section 3.4.

During operation, the achievement of the safety goals is measured against criteria that monitor the occurrence of safety goal violations, or indicate potential violations, and give confidence of the appropriateness of the safety goals.

The potential harm to be avoided that is addressed by the safety goals should include focus on collisions that can occur as a result of the kinetic energy of a moving vehicle and their impact with other objects. However, safety in non-collision situations, through other immediate or subsequent sources of harm that may arise due to the replacement of the driver with an ADS, and considerations of personal safety should also be included. Examples of this are occupants being harmed as a result of unnecessarily harsh control inputs, or nearby road users being harmed whilst taking avoiding action in response to erratic vehicle behaviour.

There is the potential for harm to the vehicle occupants due to a thermal event/fire as a result of the EV technology that is proposed for the LSAV vehicles. In this case the ADS is required to contribute to an equivalent level of safety to a non-ADS vehicle.

Subsequent sources of harm are considered as a result of responsibilities and actions that would traditionally have been taken by a driver for the safety of the passenger and are not part of the DDT (Dynamic Driving Task), e.g.

- by ensuring that passengers remain safety within the cabin while the vehicle is moving
- by not allowing embarkation of ‘unsafe’ passengers (e.g., drunk or violent people)
- by taking appropriate action in case of on-board emergencies (e.g., of a medical nature)
- by responding appropriately after incidents (e.g., by getting people to a safe location)

Table 3 lists the hazards considered and addressed with the proposed safety framework. The headings of this table are further elaborated as follows:

Hazard Category	This categorises the source of harm, whether it is due to collisions or other risk events. Some of these hazards are related to the DDT, but non-DDT hazards are also included.
Hazardous Event	Using an ISO 26262 concept, this column further defines the situation that results in harm.
Collision Type	Describes the accident type (by impacted area of the LSAV) and the objects involved in the collision.
Who can be harmed?	Linked to the collision type, this describes the “at-risk” persons.
Applicability	Looks at whether the risk applies both to the goods and people carrier applications proposed for this LSAV scheme.



Hazard Category		Hazardous Event	Collision Type		Who can be harmed	Applicability
Harm due to collision	DDT-related (collision)	LSAV collides with other road user(s)	Multiple road users	Front collision	LSAV vehicle occupant(s) other road users: - vulnerable road users and - other vehicles' occupants	LSAV passenger vehicle LSAV goods vehicle
				Rear collision		
				Side collision		
		LSAV collides with stationary object/road infrastructure	Single car	Front collision	LSAV vehicle occupant(s) Other persons in vicinity of stationary object	LSAV passenger vehicle LSAV goods vehicle
				Rear collision		
				Side collision		
	Other road user collides with LSAV	Multiple road users	Front collision	Other vehicles' occupants LSAV vehicle occupants	LSAV passenger vehicle LSAV goods vehicle	
			Rear collision			
			Side collision			
	Non-DDT related	Other road users collide with passengers after debarkation at unsafe location	Other road user (not LSAV) collides with VRU	VRU collision	LSAV vehicle occupants	LSAV passenger vehicle

Hazard Category		Hazardous Event	Collision Type		Who can be harmed	Applicability
Harm through unexpected movement of vehicle resulting in a fall of a vehicle occupant	DDT-related (motion-related)	LSAV swerves, jerks or brakes sharply/ unexpectedly causing injury to occupants within LSAV	Non-collision	Sharp acceleration, sharp deceleration, sharp cornering, or vehicle tripping/rolling over	LSAV vehicle occupant(s) Other road users	LSAV passenger vehicle
Harm from falling load	Non-DDT related (although note that motion of vehicle could be an influence)	Loss of load due to incorrect fixing or harsh movement	Non-collision	n/a	Other road users vulnerable road users and other vehicles' occupant(s)	LSAV passenger vehicle LSAV goods vehicle
Harm from fall from vehicle (passengers)	Non-DDT related	Passenger fall from moving LSAV	Non-collision	n/a	Passengers	LSAV passenger vehicle
Harm from moving mechanisms	Non-DDT related	Entrapment through moving parts (doors/windows, seats)	Non-collision	n/a	Vehicle occupants	LSAV passenger vehicle
Harm through thermal event/gas	Non-DDT related	LSAV releases smoke/ noxious chemicals and/ or suffers fire	Non-collision	n/a	Vehicle occupants or persons nearby	LSAV passenger vehicle
Harm through electric Shock	Non-DDT related	Person comes into contact with live cables or other surface due to vehicle fault or misuse	Non-collision (although risk may be increased after a collision)	n/a	Vehicle occupants or persons nearby	LSAV passenger vehicle LSAV goods vehicle

Hazard Category		Hazardous Event	Collision Type		Who can be harmed	Applicability
Personal Safety	Non-DDT related	Vehicle occupant is assaulted while travelling on LSAV	non- collision	n/a	Vehicle occupant	LSAV passenger vehicle
	Non-DDT related	Medical emergency with a vehicle occupant	non- collision	n/a	Vehicle occupant	LSAV passenger vehicle
	Non-DDT related	Vehicle occupants are trapped in vehicle or vehicle is stranded in busy traffic situation	non- collision	n/a	Vehicle occupant	LSAV passenger vehicle

Table 3: Example Hazard List.

The aim of the set of safety goals is to arrive at an abstract high-level definition for acceptable safe behaviour with the overall aim to avoid harm due to collisions. They will be further refined with more objective definitions that set out what acceptable safe behaviour means, in terms of acceptance criteria that can be assessed by a type approval organisation, before a manufacturer can register a vehicle with an ADS.

These safety goals aim, on the whole, to avoid harm, without stating any overall quantitative target, whose achievement currently cannot be demonstrated at the type approval stage (prior to deployment). The safety goals could be viewed either as a list, a hierarchy or cascade of requirements, with some requirements supporting the achievement of one or more higher-level requirements.

Additional safety goals are proposed to address the non-collision related hazards. Those are presented later in Section 3.3.1. Table 4 contains the proposed top-level safety goals supporting the achievement of avoidance of harm due to collision-related hazards:

Safety Goal (1)	Do not cause collisions
Safety Goal (2)	Avoid collisions
Safety Goal (3)	Protect all persons within and in the vicinity of the vehicle from harm

Table 4: Top level Safety Goals

A combination of all the safety principles that the Law Commissions’ (2022) consultation contains have been considered when formulating the top-level safety goals for the GB LSAV Safety and Security Scheme. It can be seen that the first top level safety goal is broadly aligned to the safety principle <Does not cause at fault accidents> while the second top level safety goal extends this towards ensuring the LSAV’s behaviour is aligned to being <As safe as a competent and careful driver>. Measures for whether Safety Goal (2) has been met to an acceptable level could include:

- ALARP (As low as reasonably practicable)
 

ALARP is a risk framework that is established in case law in the UK and forms the basis of the HSE’s (Health and Safety Executive’s) decision making. It sets goals for duty holders and requires exercising judgment whether a risk is acceptable by weighing it against the resources required to mitigate it.
- GAMAB / MSG *Globalement au moins aussi bon / Mindestens gleiche Sicherheit*

This threshold translates as ‘globally, at least as good’, and aims to ensure that any new technology achieves a level of safety that is at least that of its predecessor or existing state of the art. It could be viewed that the current, accepted level of safety is set by an appropriate performance of safety-relevant systems on the vehicle combined with ensuring an acceptable level of driving skills by all drivers equivalent to what the next safety principle below considers. Note that this concept has been developed in other countries as has no direct applicability in UK law, which instead uses ALARP and the similar SFAIRP (‘so far as is reasonably practicable’).
- As safe as a competent and careful driver
 

This safety principle considers a competent driver to contribute an acceptable level of risk to the risk involved in road transport overall. If an AV can achieve the same behaviour, then it could be considered acceptable, despite the fact that even competent and careful drivers will be involved in collisions. ‘Does not cause at fault accidents’ applies the test *if a human driver had acted in this way, would the driver be held liable for causing the accident in the law of negligence*. Work is ongoing to define ‘safety envelopes’ that aim to eliminate collisions but as per the previous approach, residual risk due to situations not covered by the safety envelopes will remain.

Feedback from the first stakeholder consultation showed strong support that the expectation of the performance is to arrive at better than an “average driver’s performance level.

➤ Positive Risk Balance

This framework sets out that a new technology must overall improve on the current level of safety while accepting that, for certain aspects of the functionality or people at risk, there might be an increase of risk, which is at least compensated by improvements in other areas. This could be seen as stricter than GAMAB, as any permutation where the overall risk of the new system matches that of the predecessor used as the benchmark would meet the criteria of being 'at least as good' but would not result in a 'positive risk balance'. The concept is being proposed in ISO/TS 5083, potentially as a combination with ALARP; note that, as per GAMAB, PRB does not have a basis within UK law.

### 3.3.1 Explanation and Rationale of each safety goal

#### **Safety Goal (1) Do not cause collisions:**

The prevention of collisions is an obvious requirement to avoid harm, which is why it has been chosen as the first top-level requirement, while appreciating that it requires expansion and further explanation on how to claim its achievement; this will be described in further requirements.

This safety goal corresponds to CertiCAV's Driving Performance Criterion 1.

#### **Safety Goal (2) Avoid collisions:**

This second safety goal takes into account that collisions can also be caused by other road users, which the safe behaviour of LSAV must be able to address. The two underlying principles behind this safety goal are that:

- (1) a LSAV should have collision avoidance mechanisms implemented as well as normal driving functionality; and,
- (2) the collision-avoiding behaviour of a LSAV does not cause unsafe behaviour of other road users.

These principles will also be described with additional requirements.

This safety goal corresponds to CertiCAV's Driving Performance Criterion 2.

#### **Safety Goal (3) Protect all persons within and in the vicinity of the vehicle from harm:**

This requirement aims to reduce or eliminate harm both in cases of actual collisions but also harm that might occur without a collision, e.g., by occupants falling due to unexpected or sharp movement of the vehicles, or due to nearby pedestrians scrambling to avoid an erratic vehicle. It is anticipated that this safety goal will be partially supported by requirements that are already covered in existing regulations, e.g., for passive safety systems and requirements for cabin design.

This safety goal relates to CertiCAV's Driving Performance Criterion 3.

#### **Further Safety Goals**

To expand the above top level safety goals, further safety goals are proposed. Before presenting the safety goals, the process that led to their definition is described.

These safety goals have been developed from a high-level analysis of causes of collisions, which have been analysed in the context of what an automated vehicle needs to be capable of during a journey. Journeys are broken down into "chunks" of functionality, which in different projects and literature are described using different terminology. In this section, we are using "vehicle behavioural competency", while NHTSA considers "vehicle manoeuvres" and work from the University of Waterloo's WISE department talks about "vehicle tasks". As a result, some of the additional safety goals are specific to particular road layouts or vehicle capabilities. It is envisaged that, at least initially, LSAVs might not be designed to perform all these capabilities or include all operating environments, and the requirements should therefore allow flexibility to reflect the needs of each system and environment.

Table 5 shows a list of vehicle capabilities that have been considered. These have been derived from work published by NHTSA (2018), SAE J3016 (2021) and Czarnecki, K. (2018a & 2018b). A more detailed analysis of behavioural competencies, and how they form an important aspect of the system that needs to be defined to support safety assurance, is provided within Section 4.3.

Behavioural Competence (operational/control level)	Behavioural Competence (tactical/manoeuvring level)																			
	Enhance Vehicle Conspicuousness	Follow other vehicle	Maintain lat/long position in lane	Lane change	Overtake (double lane change)	Merge/Exit	Avoid Obstacle	Navigate Roundabouts	Navigate T-junctions (left)	Navigate T-junctions (right)	Navigate crossroads	Join/exit traffic	Navigate Pedestrian Crossing	Navigate Rail Crossing	Navigate bridge	Navigate tunnel	Make u-turn	Park Vehicle	Navigate Roadworks	
Hold vehicle stationary	x		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Pull away from standstill	x		x	x	x			x	x	x	x	x	x	x	x	x	x	x	x	x
Perform lateral steering control (in forwards and reverse)		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Perform longitudinal acceleration control (in forwards and reverse)		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Perform longitudinal deceleration control (in forwards and reverse)		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Maintain speed		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Vehicle stability/skid control/roll control		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Table 5: Behavioural Competencies.

Table 5 contains a list of *primary* manoeuvres that comprise the fundamental vehicle control functionality, which can themselves be combined and aggregated into more complex *secondary* manoeuvres to manage particular traffic situations. Each manoeuvre introduces its own requirement that relevant ODD elements be considered in the execution of the driving function; some of the key dependencies are shown in Table 6. The placeholder *Safety Target* in Table 6 will form the basis for the definition of the technical performance requirements for perception, planning and actuation functions.

In addition, it is proposed to use the approach as defined in the ALKS (2021) regulation to develop the definitions of the “minimum threshold for acceptably safe behaviour” for particular scenarios that are envisaged to be within the scope of the low-speed driving application covered by the proposed requirements.

Potential Accident	Hazardous Behaviour (ego vehicle)	Vehicle Factors	Environmental Factors	Behaviour by other object(s) that must be accommodated	Safety Target	Behavioural Competency
Front collision* Side collision* Rear collision* <i>* dependent on direction of movement/rolling</i>	Does not hold vehicle stationary and rolls into - path of other vehicle or - into VRU or - stationary infrastructure object	Brake performance/condition, Tyre condition	Road friction, road incline	n/a	Maintain stationary position until path is clear or right of way situation is given	Hold vehicle stationary
Front collision* Side collision* <i>* dependent on location or path of collision object</i>	Pulls away from standstill despite the presence of an object in front - that is either crossing (VRU or vehicle) - or is present in the path (VRU, vehicle or infrastructure object)	Acceleration performance, vehicle conspicuity	Right of way considerations - traffic lights - traffic signs - road markings - road layout	Approach speed, direction and priority of other road users	Maintain distance (long/lat), change path around obstacle	Pull away from standstill
Rear collision <i>caused by vehicle behind moving when ego vehicle fails to</i>	Does not pull away despite clear road ahead	Acceleration performance, vehicle conspicuity	Road friction, road incline		Support flow of traffic	

Potential Accident	Hazardous Behaviour (ego vehicle)	Vehicle Factors	Environmental Factors	Behaviour by other object(s) that must be accommodated	Safety Target	Behavioural Competency
Front collision* Side collision* Rear collision* <i>* dependent on road layout and location or path of collision object</i>	Inappropriate lateral steering control (see manoeuvre description for detail)	Vehicle position & heading, vehicle speed, vehicle dimensions, vehicle conspicuity, steering performance	Road layout, road friction	Other vehicles/actors approaching from foreseeable directions (see manoeuvres for details)	Maintain lateral clearance to lane boundary and objects (moving and stationary) in path	Perform lateral steering control
Front collision* Side collision* Rear collision* <i>* dependent on road layout and location or path of collision object</i>	Inappropriate acceleration (see manoeuvre description for detail)	Brake performance/condition, tyre condition, acceleration performance, deceleration performance, vehicle position & heading, vehicle speed, vehicle dimensions, vehicle conspicuity	Road layout, road friction, right of way considerations - traffic lights - traffic signs - road markings - road layout	Other vehicles/actors approaching from foreseeable directions (see manoeuvres for details)	Maintain a safe speed according to prevailing conditions and manoeuvre specific criteria - line of sight - road friction condition - environmental visibility - traffic condition/density	Perform longitudinal acceleration control
	Inappropriate deceleration (see manoeuvre description for detail)					Perform longitudinal deceleration control
	Inappropriate speed (see manoeuvre description for detail)					Adopt appropriate speed



Potential Accident	Hazardous Behaviour (ego vehicle)	Vehicle Factors	Environmental Factors	Behaviour by other object(s) that must be accommodated	Safety Target	Behavioural Competency
Rear collision* Side collision* <i>* dependent on road layout and location or path of collision object</i>	Reverses into an object that is either - stationary, - crossing, or - moving in path	Brake performance/condition, Tyre condition, acceleration performance, deceleration performance, vehicle position & heading, vehicle speed, vehicle dimensions, vehicle conspicuity	Road friction, road incline, road layout	Approach speed, direction and priority of other road users	Maintain clearance (long/lat)	Reverse vehicle
Front collision* Side collision* Rear collision* <i>* dependent on ego vehicle manoeuvre, road layout and location or path of collision object</i>	Perform a manoeuvre that requires indication of intent to other road users, but fails to provide appropriate indication	Vehicle speed, vehicle position & heading	Visibility	n/a	Ensure vehicle is conspicuous through appropriate - front lighting - rear lighting - indication of intentions and - warning of unexpected behaviour	Enhance Vehicle Conspicuity
Front Collision	Travels at too close a distance to lead vehicle	Vehicle speed, vehicle position & heading, vehicle dimensions, vehicle conspicuity	Road friction, road layout	Other vehicles/actors approaching from foreseeable directions or departure of target vehicles, presence of close vehicles/	Maintain a safe distance to moving and stationary objects in lane, entering or exiting lane	Follow other vehicle

Potential Accident	Hazardous Behaviour (ego vehicle)	Vehicle Factors	Environmental Factors	Behaviour by other object(s) that must be accommodated	Safety Target	Behavioural Competency
				objects (e.g. parked cars)		
Front collision Side collision	Lateral Adjustment when not required (when road curvature remains the same or decreases). No lateral adjustment when required (when road curvature increases or to maintain lateral distance to object encroaching ego lane)	Vehicle position & heading, vehicle speed, vehicle dimensions, vehicle conspicuity	Road layout (lane boundary, curvatures), road friction	n/a	Maintain lateral clearance to lane boundary	Maintain lat/long position in lane
	No laterals adjustment when required (when road curvature increases or to maintain lateral distance to object encroaching ego lane)	Vehicle speed, vehicle position & heading, vehicle dimensions, vehicle conspicuity	Road layout, road friction	Lateral encroachments into lane from other objects (stationary and dynamic)	Maintain lateral clearance to objects by adjusting position within lane (while maintaining a clearance to lane boundary), change lane	

Potential Accident	Hazardous Behaviour (ego vehicle)	Vehicle Factors	Environmental Factors	Behaviour by other object(s) that must be accommodated	Safety Target	Behavioural Competency	
Front collision* Side collision* Rear collision* <i>* dependent on road layout and location or path of collision object</i>	Changes lane in the presence of either an oncoming vehicle or a slower or faster moving vehicle in the adjacent target lane. Changes lane without appropriate indication	Vehicle speed, vehicle position & heading, vehicle dimensions, vehicle conspicuity	Road layout, road friction	positions and headings of and approaches from other vehicles/actors from foreseeable directions	Ensure appropriate safe distances to front, side and rear of passed vehicle and other moving and stationary objects	Lane change	Overtake (double lane change)
	changes lane in the presence of either an oncoming vehicle or a slower or faster moving vehicle in the adjacent target lane. Does not complete lane change prior to end of lane or without a target lane present. Changes lane without appropriate indication	Vehicle speed, vehicle position & heading, vehicle dimensions, vehicle conspicuity	Road layout (position of target lane), road friction				Merge/Exit

Potential Accident	Hazardous Behaviour (ego vehicle)	Vehicle Factors	Environmental Factors	Behaviour by other object(s) that must be accommodated	Safety Target	Behavioural Competency
Front collision	Insufficient deceleration to avoid collision or reduce impact severity. Evasive manoeuvre initiated but alternative path occupied	Vehicle speed, vehicle position & heading, vehicle dimensions, vehicle conspicuity	Road layout/clear space, road friction, visibility	Presence of objects in lane ahead or moving towards the lane (e.g., crossing VRUs)	Bring vehicle to a stop before reaching the position of the collision risk (travelling forwards). Reduce impact speed as much as possible. Evasive manoeuvre (depending on availability of alternative paths)	Avoid Obstacle
Side collision	Evasive manoeuvre initiated but alternative path occupied	Vehicle speed, vehicle position & heading, vehicle dimensions, vehicle conspicuity	Road layout/clear space, road friction, visibility	Other road users encroaching on lateral clearance	Lateral evasive manoeuvre, (depending on space available) deceleration or acceleration	
Rear collision	Insufficient deceleration to avoid collision or reduce impact severity	Vehicle speed, vehicle position & heading, vehicle dimensions, vehicle conspicuity	Road layout/clear space, road friction, visibility	Presence of objects behind vehicle or moving towards the vehicle (e.g., crossing VRUs)	Bring vehicle to a stop before reaching the position of the collision risk (travelling in reverse), reduce impact speed as much as possible, evasive manoeuvre (depending on availability of alternative paths)	

Potential Accident	Hazardous Behaviour (ego vehicle)	Vehicle Factors	Environmental Factors	Behaviour by other object(s) that must be accommodated	Safety Target	Behavioural Competency	
	Performs a collision avoidance manoeuvre without an imminent collision risk present	Vehicle speed, vehicle position & heading, vehicle dimensions, vehicle conspicuity	Road layout/clear space, road friction, visibility	Presence and speed of objects behind vehicle or moving towards the vehicle (e.g., crossing VRUs)	Do not perform emergency manoeuvres in situations where they are not required (false positives)		
Front collision Side collision Rear collision	Turns without appropriate indication, turns in the presence of oncoming vehicles that have right of way	Brake performance/condition, tyre condition, acceleration performance, deceleration performance, vehicle speed, vehicle dimensions, vehicle conspicuity	Road layout (number of road entries/exits, junction angle, approach, lane layout, control, signage, markings) road friction visibility	Presence and speed of objects towards intersection (including indication of and potential change of intention)	Maintain longitudinal and later clearances, stop when required to give way, only proceed when the junction is clear and no other vehicle has right of way, make own intentions clear	Navigate Intersection	Navigate Roundabouts
							Navigate T-junctions (left)
							Navigate T-junctions (right)
							Navigate crossroads
							Join/exit traffic
Front collision Side collision	Does not reduce vehicle speed to standstill before the crossing location Moves off before the crossing is clear	Brake performance/condition, tyre condition, acceleration performance, deceleration performance, vehicle speed, vehicle dimensions, vehicle conspicuity	Signage and markings of crossing points, visibility	Presence and speed of objects towards crossing points (including change of intention)	Stop at a safe distance before the crossing for every intended crossing, reduce speed for potential crossing situations	Navigate crossing	Navigate Pedestrian Crossing
							Navigate Rail Crossing

Potential Accident	Hazardous Behaviour (ego vehicle)	Vehicle Factors	Environmental Factors	Behaviour by other object(s) that must be accommodated	Safety Target	Behavioural Competency	
Front collision Side collision	Failure to compensate for specific environmental considerations	Vehicle speed, vehicle dimensions, vehicle conspicuity	Specific environmental aspects concerning road friction, side wind effect road layout, presence of toll booths	n/a	Ensure vehicle adapts its behaviour to be compatible with safe operation within the prevailing environmental conditions	Navigate specific road situation	Navigate bridge
	Failure to ensure vehicle is visible in dark environment		specific environmental aspects concerning road friction, side wind effect road layout, presence of toll booths, presence of toll booths	n/a	Ensure vehicle is conspicuous through appropriate - front lighting - rear lighting - indication of intentions - warning of unexpected behaviour		Navigate tunnel
Front collision Side collision Rear collision	<i>Key consideration in these scenarios is to ensure no other vehicles are impeded and that there is clear space before moving - considerations to be given to VRUs moving about in restricted spaces and ensuring their detection and maintaining clearances that facilitate pedestrian movement</i>					Special manoeuvres	Make u-turn
	<i>Key consideration in these scenarios is to ensure no other vehicles are impeded and that there is clear space before moving - considerations to be given to VRUs moving about in restricted spaces and ensuring their detection and maintaining clearances that facilitate pedestrian movement</i>						Park Vehicle

Potential Accident	Hazardous Behaviour (ego vehicle)	Vehicle Factors	Environmental Factors	Behaviour by other object(s) that must be accommodated	Safety Target	Behavioural Competency
Front collision* Side collision* Rear collision* <i>* dependent on road layout and location or path of collision object</i>	Any of the above, depending on location and layout of road works	Brake performance/condition, tyre condition, acceleration performance, deceleration performance, vehicle speed, vehicle dimensions, vehicle conspicuity	Lack of, temporary markings / signage, contradictory markings/ signage, narrow road layout	Presence of VRUs	Ensure vehicle is able to operate safely within the vicinity of temporary infrastructure (e.g. roadworks)	Navigate Roadworks

Table 6: Mapping between accidents, behavioural competencies and initial safety targets.

Table 7 lists all the supporting safety goals, alongside an explanation and rationale for each.

Table 8 presents a mapping of how the safety goals address and mitigate potential collision events.

SG ID	Safety Goal Description	Explanation and Rationale
4	Follow the rules of the road	<p>This is one of the safety goals that supports all three top level goals.</p> <p>In the UK, the Highway Code sets expectations by describing specific behaviour that is required when driving, while also prohibiting some actions and behaviour. This includes ‘must’ rules, which are linked to legislation, and ‘should’ rules, which are advisory. Ostensibly, this guidance is aimed at informing drivers, but it also steers Courts in making decisions. As the LSAV shares the road transport infrastructure it will also be required to adhere to these rules.</p> <p>Some of these rules will require interpretation for AVs as they are driver specific.</p> <p>This safety goal corresponds to CertiCAV’s Driving Performance Criterion 4.</p>
5	Approach intersections with care	<p>This safety goal supports safety goals (1) and (2), as there is a higher occurrence of collisions due to the sharing of space by many traffic participants. This requires interpretation for specific types of intersections.</p>
6	Drive only into clear space	<p>This particular safety goal extends safety goal (1)’s application. It could apply to both driving forwards, and also reversing</p>
7	Adjust vehicle speed to prevailing conditions	<p>Vehicle speed is a key consideration for road transport safety. Collisions at higher speed are more severe and what is a safe vehicle speed also differs depending on the prevailing conditions. Therefore, in order to support the top safety goals (not causing a collision and avoiding collisions), it has to be assured that the LSAV’s speed is adapted to ensure safe driving, e.g. within friction limits.</p>
8	Prioritise human life while reducing damage	<p>As previously stated, it is foreseen that accidents will still occur and that the LSAV will be involved in collision or near-collision events that will require trade-offs that might involve violating other safety goals. In this case there is a need for a safety goal that prioritises human life over material damage.</p>
9	Drive considerately	<p>This safety goal, supporting safety goals (2) and (3), is aimed at ensuring that the LSAV design is able to make special considerations where necessary to ensure that safety is maintained by being able to make allowances.</p> <p>There are requirements for “considerate” driving in the Highway code that are better aligned to human driving – some interpretation might be required to identify how this can be translated for an ADS.</p> <p>This safety goal corresponds to CertiCAV’s Driving Performance Criterion 8.</p>
10	Provide information to occupants	<p>To support the safety of occupants, it is important that they are given information about the status, progress and actions of the LSAV so that occupants can take appropriate actions themselves. This might include information about setting off, doors closing or announcing of stops.</p>



SG ID	Safety Goal Description	Explanation and Rationale
11	Drive smoothly	<p>This safety goal is supporting the safety goal (3) to protect vehicle occupants by requiring a safe driving style for this particular type of vehicle where not all occupants are expected to be seated and securely fastened at all times.</p> <p>There is a trade-off and possible conflict between this safety goal and safety goal (2).</p>
12	Travel only on appropriate lanes/ road segments	<p>Road Transport overall is based on the principle that vehicles and road users move along road segments and lanes, and rules exist that govern their interaction. Therefore, staying on the appropriate land or road segment is the safer place to be, as leaving the road may result in a number of undesirable outcomes and potential collisions.</p> <p>It is highlighted that this safety goal is one that will always be subject to arbitration in case of imminent collisions.</p>
13	Do not hit a road user travelling ahead from behind	<p>These three safety goals are listed separately from safety goal 4 (to follow the rules of the road), to address some collision scenarios in more detail. For their implementation, they need to be interpreted into parameters that an AV can control. An additional consideration for these safety goals is that the expected behaviour of other road users needs to be taken into account when refining how these goals translate into verifiable performance requirements.</p>
14	Do not obstruct other road users when changing lane	
15	When turning, follow right of way rules	
16	Follow prevailing driving styles	<p>This safety goal also supports safety goal 4 (to follow the rules of the road), but is aimed at the more subtle rules that are less clearly defined, that human drivers instinctively adopt. If there are specific (e.g. local) agreements then these should be accommodated as far as possible.</p> <p>This safety goal corresponds to CertiCAV's Driving Performance Criterion 6.</p>
17	Indicate intentions as per rules	<p>This safety goal covers the all the relevant requirements addressing vehicle conspicuity and signalling. As part of the DDT, the LSAV's ADS must ensure that the vehicle can be seen by other road users and that its intentions when manoeuvring are made clear to other road users.</p>
18	Maintain appropriate safety margins	<p>This safety goal supports all three top-level safety goals, as maintaining safety margins ensures that the LSAV is able to react to unexpected events or stop in case of imminent collision risk. It also supports the comfort and safety of occupants, as it facilitates smoother driving style, which in turns protects occupants from harm in case of harsh vehicle movements.</p> <p>This safety goal corresponds to CertiCAV's Driving Performance Criterion 5.</p>
19	Avoid using behaviour that may not be expected by other road users	<p>This safety goal supports safety goal (2) by requiring the behaviour of the LSAV to be predictable and understandable to other road users. In this way, other road users can adjust their actions accordingly based on their understanding and expectations of how the LSAV will behave. Note that</p>

SG ID	Safety Goal Description	Explanation and Rationale
		the vehicle needs to behave in a way that is compatible with both manually-driven vehicles and also other AVs, potentially developed and operated by other organisations.
20	Avoid obstructing traffic flow	<p>This safety goal also supports the top-level safety goal to avoid collisions by ensuring behaviour that facilitates other road users can also predict the LSAV's behaviour. This includes a vehicle making progress and not stopping and waiting for extensive durations, as this might lead to hazardous behaviour of other vehicles in order to try and make progress.</p> <p>This safety goal corresponds to CertiCAV's Driving Performance Criterion 9.</p>
21	Avoid behaviour not expected by occupants or persons in vicinity of vehicle	<p>In order to ensure the safety and comfort of occupants the LSAV should behave in a way that enables occupants to keep safe. Driving in a predictable way enables occupant to take actions like holding on, for example, before the LSAV brakes or changes directions.</p> <p>This safety goal corresponds to CertiCAV's Driving Performance Criterion 7.</p>

*Table 7: Supporting Safety Goals.*

Hazard Category	Hazardous Event	Collision Type	Example Hazardous Behaviour(s)	Applicable Safety Goal(s)																				
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Harm due to collision	LSAV collides with other road user(s)	LSAV + other road user	LSAV leaves road/lane layout (while driving forwards or reverse). LSAV changes lane into an occupied space. LSAV turns at intersection into an occupied space. LSAV progresses through occupied crossing. LSAV does not achieve collision avoidance. LSAV moves from standstill despite object in front.	X			X	X		X	X				X		X			X				
			LSAV vehicle reverses into occupied space. LSAV does not achieve collision avoidance.	X				X	X	X				X	X						X			

Hazard Category	Hazardous Event	Collision Type	Example Hazardous Behaviour(s)	Applicable Safety Goal(s)																								
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21				
	LSAV collides with stationary object/road infrastructure	LSAV only	Side collision	LSAV leaves road/lane layout (while driving forwards or reverse). LSAV turns at intersection into an occupied space. LSAV passes in lane road user without sufficient clearance. LSAV does not achieve collision avoidance.	X				X		X	X				X		X				X						
			Front collision	LSAV leaves road/lane layout (while driving forwards or reverse). LSAV does not achieve collision avoidance.	X						X	X					X							X				
			Rear collision																									
			Side collision		X					X	X	X					X								X			
			Rollover			X					X					X	X											

Hazard Category	Hazardous Event	Collision Type	Example Hazardous Behaviour(s)	Applicable Safety Goal(s)																					
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
	Other road user collides with LSAV	LSAV + other road user	front collision LSAV makes a turn without following give-way rules. LSAV changes lane without following give-way rules (including overtaking).		X						X	X	X			X		X	X	X	X	X			
			rear collision	LSAV slows down unexpectedly.		X						X	X	X			X		X	X	X	X	X	X	
				LSAV is "stuck/immobile" or does not move on as expected by other road users.																				X	X
	side collision Other road user encroaches into LSAV's lane. LSAV is being cut up. LSAV makes a turn without following give-way rules.		X							X	X	X			X		X	X	X	X	X	X	X	X	
Other road users collide with passengers after disembarking	Other road user (not LSAV) collides with VRU	n/a	Vehicle is brought to a standstill in traffic and passengers leave the vehicle (e.g. due																						

Hazard Category	Hazardous Event	Collision Type	Example Hazardous Behaviour(s)	Applicable Safety Goal(s)																				
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
			to an on-board incident).																					
Harm through unexpected movement of vehicle resulting in a fall	LSAV swerves, jerks or brakes sharply/ unexpectedly causing injury to occupants within LSAV	Non collision -	Sharp accel. Sharp decel. Sharp cornering vehicle tripping/ rolling over			X																	X	X

Table 8: Mapping between Safety Goals and Collision-related Hazards.

As previously stated, the aim of these safety goals is to set the framework for acceptably safe behaviour. There is currently no industry-wide consensus on this, but identical, or at least similar, safety goals are being proposed by other ongoing legislative initiatives. This will be further analysed for the proposal for technical requirements for the approval scheme (see Section 5.4)

It could be argued that the choice of safety goals is to some extent arbitrary in depth and coverage, and the list should not be seen as exhaustive. Furthermore, some safety goals may not be applicable for some systems, particularly those safety goals related to particular traffic scenarios or vehicle functionality (e.g., reversing or turning) – they might be considered non-compulsory in such applications.

Additionally, in particular scenarios, several of these safety goals might be applicable, requiring a choice to be made as to which one to enforce. In other circumstances where there might be conflicts, a decision regarding the priority of safety goals will be required. An example of this would be where the only way to avoid a collision is to apply maximum braking, which is likely to result in injuries to LSAV occupants, particularly standing passengers.

One possible approach would be to make certain safety goals mandatory for all LSAV applications, while others only apply if certain functionality is implemented as part of a particular LSAV application. The manufacturer would therefore need to declare the behavioural competencies of their LSAV, which must be matched to the ODD and intended deployment. The coverage of the safety concept and safety case would then be required to address all applicable safety goals and technical requirements. This approach is further developed in Section 5.4.

There is a balance to be struck when defining these safety goals, to ensure that the benefits of road transport are upheld – after all it could be argued that a vehicle that never moves would be safest. Nevertheless, there is no safety goal currently proposed addressing the need for a LSAV to complete its journey. This would be a safety consideration for certain vehicles (e.g., emergency vehicles like ambulances), but in other cases would be treated primarily as an inconvenience, and hence not included.

A hierarchical view of the safety goals is shown in Figure 6.

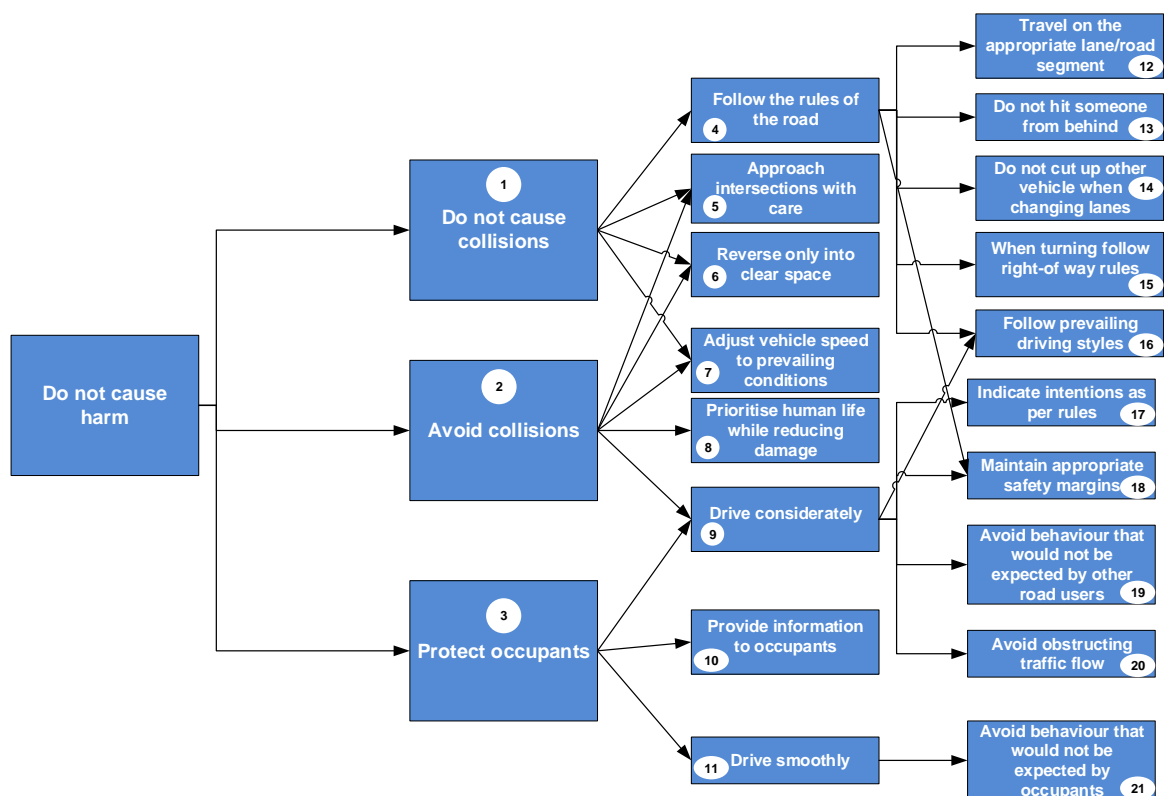


Figure 6: Hierarchical view of the Safety Goals.

Proposed safety goals for non-collision hazards are listed in Table 9. For more information on hazards not related to the ADS performing the DDT, please consult the outputs from Work Package 4 of this project.

Hazard Category		Hazardous Event	Safety Goal(s)
Injuries due to collision (LSAV not involved)	MRC or other stop in location where it is unsafe to exit	Other road user collides with passenger after disembarking	Ensure vehicle is only stopped in safe positions unless unavoidable. Other road users are alerted, and passengers are warned before disembarking.
Injuries from falling load	Non-DDT related (although vehicle motion could have influence)	Loss of load due to incorrect fixing or harsh movement	Could be considered to be partially addressed by Safety Goal (11) Drive smoothly. Otherwise, this is proposed to be addressed as an in-operation requirement.
Injuries from fall from vehicle (passengers)	Non-DDT related	Passenger fall from moving LSAV	Ensure doors are closed before the vehicle starts moving Ensure doors remain closed while the vehicle is moving
Injuries from moving mechanisms	Non-DDT related	Entrapment through moving parts (doors/windows, seats)	Anti-trap mechanisms or slow movements together with appropriate warnings
Thermal event/gas	Non-DDT related	LSAV develops smoke/release of noxious chemicals or a fire	Ensure occupants and operator are alerted and vehicle is brought to a stop
Electric Shock	Non-DDT related	Person comes into contact with live wire or other surface	Ensure persons in the vicinity of the vehicle are protected from exposure to electric shocks

*Table 9: Mapping between Safety Goals and Non-collision-related Hazards.*

The next step is to develop the safety goals into more concrete requirements for an ADS system, or if appropriate, for particular components like its perception, planning and actuation systems. This should include requirements for HMI considerations, specific aspects of the implementation using new technologies (e.g., Machine Learning) and fault and threat conditions.

It is proposed to use aspects from the ALKS (2021) regulation to develop the definitions of the “minimum threshold for acceptably safe behaviour” for particular scenarios that are envisaged to be within the scope of the low-speed driving application covered by the proposed requirements.

In addition to the safety goals, acceptance criteria will need to be established; this topic is discussed in the Section 3.4. It is envisaged that there will be acceptance criteria at various stages or abstraction layers (e.g., whole vehicle, ADS, individual systems, individual test case).



### 3.4 Acceptance Criteria

The overall aim of a regulator when devising a regulatory scheme is to ensure that the scheme sets out requirements that define the minimum acceptable safety to be achieved. This means that the system does not pose an unacceptable level of risk to the public when those requirements have been demonstrated and assessed at the approval stage. These requirements can take the form of specifying required performance or functionality or prescribing processes that aim at ensuring appropriate rigour is applied during development, which in turn translates into a predictable and safe system.

The definition of residual risk, taken from ISO 26262 (2018) is “risk remaining after the deployment of safety measures”, a concept that is illustrated within Figure 7.

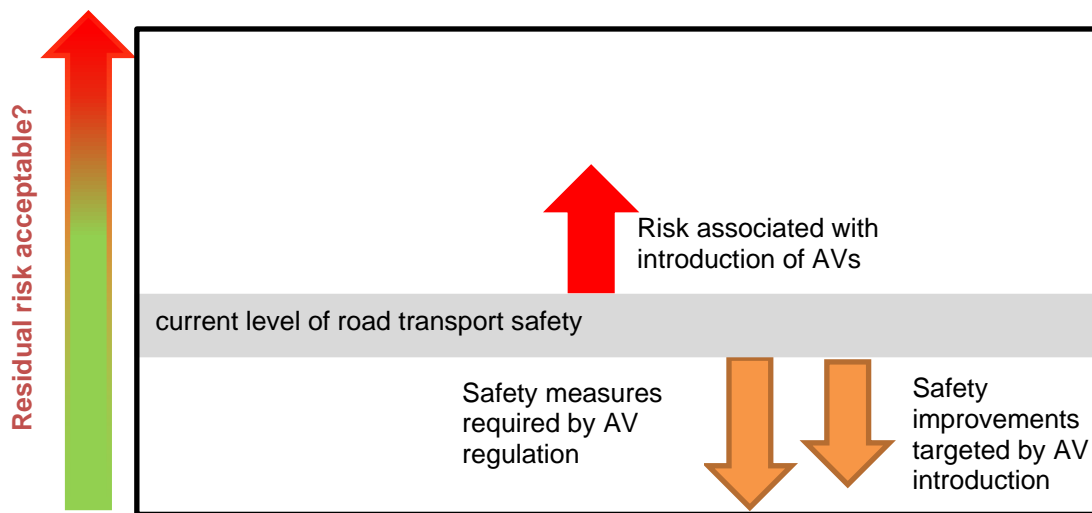


Figure 7: Risk Levels.

The interpretation of acceptable safety for the purpose of the approval scheme is the ability of a low-speed automated vehicle to operate without unreasonable risk to passengers and other road users in the road transport ecosystem. Damage to property or infrastructure and operational effects are also considered if they might result in an indirect potential for harm to people. The decision what “unreasonable” represents requires some criterion to form a judgment. This criterion should facilitate objective decision making which a quantitative safety threshold might provide.

While acceptance criteria are necessary to make decisions about acceptable risk, there are difficulties associated with quantitative safety indicators as typically success is shown as the absence of negative outcomes (e.g., collisions/fatalities/injuries), meaning that the safer an operation or system is, the less data or information about its risk is usually available to analyse. One way to address this problem is to use leading indicators to predict the occurrence of major hazards, but care must be taken with interpreting low incident rates of minor events as it cannot be used as a guarantee that all major hazards are controlled. For example, human drivers often drive at smaller distances to other vehicles than recommended by traffic rules without it resulting in collisions every time. Equally, an ADS might encroach a defined safety margin to a particular object (e.g., a VRU) when needing to balance its distance to multiple moving targets.

Hence it can be seen that metrics need to be appropriate to what is being controlled and that all necessary factors influencing risk need to be taken into account in any risk calculations, including whether risk is considered for an individual, as societal risk or even specific to the location.

It should also be noted that safety, as a property, evolves over time. This is particularly relevant to the topic of automated driving, which, as a technology is being pursued in part to make road transport safer. Lowering fatalities and injuries resulting from collisions has been an aim of many organisations involved in road transport, and vehicle manufacturers, governments and others have invested in technology and measures to make driving safer. This has been measured with a number of metrics over time, such as:

- Simple counts of events of interest (e.g., fatalities, collisions)
- Rates of occurrence of events of interest, using an exposure parameter as a denominator with a count of events as numerator (e.g., collisions per distance travelled or operating hours)

Occurrence rates provide a context which invite their use in comparisons; for example, against human driven miles. But if such a comparison is attempted, care must be taken that the denominator allows this to be meaningful. As noted in the RAND report “Measuring Automated Vehicle Safety” (Fraade-Blanar et al, 2018), “not all miles or operating hours are the same!”

One possible approach for risk acceptance criteria that has been proposed is to prescribe a single absolute risk target value that would be required to be met. Using this approach, acceptable safety would be assessed by measuring undesirable behaviour or specific event occurrences of a particular outcome through setting an upper bound of this acceptable occurrence count or rate.

For a complete automated vehicle, the potential categories might be (fatal or severe) collisions / near collision events etc. per distance travelled or operating hour. Criteria can potentially also be set for system level targets. Examples for those include false positive/ false negative detection rates for objects (e.g., pedestrians).

This has been considered, and provisions included in ongoing draft regulatory proposals, at EU level. Version 4.1 of the “Commission Implementing Regulation for automated driving for urban shuttles” (EU, 2022) proposed the following quantitative target as a requirement:

*[No number] The ADS shall overall be free of unreasonable risks for the vehicle occupants or any other road users and shall ensure a higher level of safety than the level of safety of vehicles driven by persons (indicative target:  $10^{-9}$  fatality per hour of operation).*

In the updated version, this requirement has been modified to set out a target threshold that differs by 2 orders of magnitude, while no longer including that this target value is aimed at achieving an improvement in safety on human driving.

*8.1.1 A safety target for design and development, shall be used. As indicative target,  $10^{-7}$  fatalities per hour should be considered as a minimum for applications covered by this regulation. The manufacturer may use other metrics and method provided it can demonstrate that it leads to an equivalent level of safety.*

The differing numerical values and rephrasing of the requirement show the difficulties involved in setting a quantitative target of what is acceptable.

In order to evaluate the suitability of this type of acceptance criteria for type approval, it is useful to reflect on the definition of risk, with risk being evaluated as the likelihood of an unwanted event (an event that causes harm or loss) combined with its potential consequences. To determine the overall risk associated with a system, operation or asset requires an understanding of the probability of all hazardous events that may occur and their consequence. This is illustrated in Figure 8, which shows examples for the different factors influencing event likelihood and event consequences.

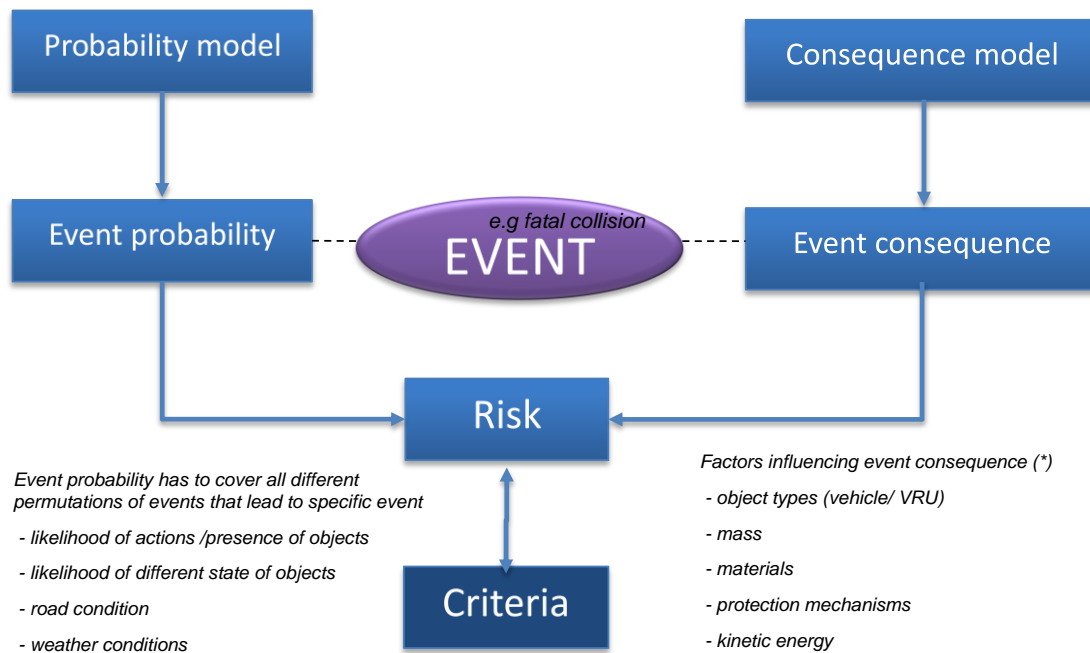


Figure 8: Risk Evaluation.

The hazard list presented in Section 3.2.7 is intended as a starting list for hazardous events when assessing risk, but additional hazardous events and hazards, e.g., those associated with specific technologies, might exist. Attempting to quantify in absolute terms an acceptable risk against which to set targets at the design stage for an LSAV would require controlling for the **causality** and **combinatorics** of the many possibilities that can yield the same event category outcome.

An additional consideration is that there are several ways of categorising unwanted events, e.g., by the involved collision objects (LSAV collision with VRU) or the resulting injury from the collision (fatality - VRU or on-board). The outcome of the same unsafe LSAV behaviour may lead to events that fall into different categories depending on the surrounding circumstances, e.g., not slowing down sufficiently might result in a fatal injury if the resulting collision is with a VRU, or in material damage only if the collision object is a roadside barrier. Equally, different unsafe behaviours may result in the same outcomes. A fatal head-on collision might be the result of overtaking into oncoming traffic or progressing into an incorrect lane during a turn manoeuvre. And for each of these possible collision events there are near-infinite permutations of causes:

Not only do “traditional” technical failures need to be considered (traditionally covered by Performance requirements and reliability targets for mechanical system and Functional Safety for EE systems), but also:

- The cases where external factors compound to a situation or scenario that is not covered in the design or V&V phase (SOTIF)
- Malicious external threats (Cybersecurity), and
- Errors by other road users

No solution or precedence is currently available to show how an absolute target for a risk event could be:

- Allocated across systems and subsystems during design to determine both performance targets and target failure rates for each system (particularly non-deterministic systems)
- Substantiated with V&V evidence with a sufficient confidence interval

As stated previously, it is possible to set requirements and specify criteria that can be directly observed at a more detailed level, e.g., that VRUs shall be passed at a specified minimum lateral distance, but it will not be possible to translate a violation of the minimum passing distance into a contribution to an

overall risk value for VRU collisions, as other factors will be contributing to the outcome at each occurrence of such a scenario.

Because an overall target threshold would also need to ensure that the target that is set results in an improvement on current road traffic safety, there remains the challenge of how to make a statistically valid comparison between AVs and human-driven vehicles, as comparable data that is representative of the TOD would be required. Measuring the achieved safety in operation, on the other hand, with metrics such as KSI (killed or seriously injured) rates will be possible over time acknowledges (Fraade-Blanar et al, 2018), and that is where a future safety target could either be derived from or put in a place as a target with the potential for a target improvement in safety over time.

In other industries, the link between risk events and their outcomes is more directly coupled (for example, in the aviation sector, where quantitative failure targets do exist). Here, aviation safety standards set limits for probabilities per flight hour of failures resulting in different types of outcomes (catastrophic/ hazardous/ major/ minor). Based on a risk graph that determines acceptable risk as a function of the classification of the outcome of a particular failure, and the probability of that failure, limits are set for each failure mode (dependent on aircraft class). These permissible failure limits are derived from statistics based on decades of data and historical contributions per system to accidents. These limits are also expressed per “average probability per flight hour” to ensure comparability, using averaged flight durations, profiles and assumed total number of service life flights. Additional specific design principles for particular systems or subsystems (e.g., no single point failure resulting in catastrophic failures) are also required by these safety standards.

The application of a target value in this industry is focussed on ensuring that the contribution of system failures to hazardous events is below a set threshold, with the nominal performance of the aircraft systems considered to be safe and environmental influencing factors controlled and understood. The automotive industry requires a way to measure safety that covers both

- the performance aspects (including considerations of how contributions from actors in the different complex “open context” operating environments influence safety) and
- the acceptable failure targets in combination

in order to enable qualitative assessments to be made where it can be determined if a minimum safety “bar” has been achieved.

This principle has been described in Favarò, F. M. (2021) with the following formula:

$$\text{Frequency of occurrence of (certain level of) harm} \leq \text{Threshold Value}$$

which highlights 2 potential interpretations, and hence uses, for Positive Risk Balance (PRB).

Firstly, PRB may be interpreted as a “safety assurance metric for a completed system” insofar as the achieved threshold value for the completed ADS system and the resultant harm from its deployment would have to be shown to achieve a lower threshold value than is currently experienced.

The second interpretation lends itself more to guiding decisions during the development of an ADS system as to what the threshold value should be for individual hazards or hazardous events.

Both interpretations can be used independently, meaning that an overall risk framework setting a positive risk balance as an overarching safety goal could be supplemented by the use of PRB as an individual criterion supporting the setting of an acceptability target.

Based on the arguments put forward, this report proposes that at the current time, such an absolute risk target cannot be shown to be met at design approval stage in a way that it can predict the performance in operation. Instead, this report supports the view expressed as part of the Law Commissions consultation by the Faculty of Advocates that “the definition of the safety standard needs to be more nuanced than can be achieved in a single sentence” (Law Commissions, 2022).

An alternative approach proposed in literature, for example in ISO/PAS 21448 (2019), Shalev-Schwartz, S., Shammah, S., & Shashua, A. (2018), and SaFAD (2019), is based on assessing the performance of an automated vehicle by evaluating the performance of its ADS in individual scenarios, each with their own separate criteria based on factors that are relevant for the particular scenario (see Section 5.9 for more information on ‘scenario-based testing’). This risk evaluation approach includes defining

and assessing safety by means of a comparative human performance model, and evaluating the capability of the ADS that controls the vehicle to handle scenarios safely, with criteria for:

- expected behaviour/outcomes in individual scenarios;
- expected verification and validation evidence;

with the aim of improving road safety. Comparison to human drivers is considered necessary, at least in the short term, since the abilities and flaws of (certified) humans reflect the level of risk currently tolerated by the general public. The Law Commissions' report considered views on whether:

- (1) as safe as a competent and careful human driver;
- (2) as safe as a human driver who does not cause at-fault accidents; or,
- (3) safer than the average human driver;

were appropriate comparisons, with none of the options receiving a *majority* response.

The key detail that remains with either of these options is how to describe and quantify the performance of any human driver as huge ranges of capabilities exist, based on either age, training status, health status, distractions, environmental conditions, and many other factors.

Acceptance of safety for an ADS should combine the review of quantitative and qualitative evidence to make a judgement about whether a particular AD system presented for type approval is safe enough for introduction, while acknowledging that residual risk will remain even when all safety requirements and safety measures contained in the regulation have been implemented.

This residual risk arises from unknown hazards, known limitations in mitigation for known hazards, and from uncertainty in assurance of mitigations needs to be monitored on an ongoing basis during operation with a process in place for resolution of issues that are found.

The approach put forward is based on assessing the performance of an ADS by requiring manufacturers to show that the LSAV can achieve behaviour that, in the context of its intended operating environment, is considered acceptably safe. The acceptably safe behaviour is expressed in safety goals and associated criteria. High-level safety goals have been formulated as objectives in 3.2.7 that are refined into more specific requirements next.

The manufacturer will be required to describe and show evidence of how their LSAV design meets the safety goals, e.g., by describing and arguing perception capabilities, environmental conditions that are within scope, and safety envelopes implemented based on models of other road users' behaviour. Additionally, this evidence will provide the input to the analysis and interpretation of in-use monitoring data to ensure action can be taken when necessary.

However, at the Type Approval Stage, it will not be possible to quantify the overall residual risk that remains in the design in such a way that an upper threshold of occurrences can be predicted, due to:

- functional insufficiencies that may be triggered by unknown hazardous scenarios;
- residual faults in the design;
- the complexity of the systems, their operating environments, and the modalities of incidents that they may be involved in.

## 4 Definition of the System & Deployment

### 4.1 Operational Design Domain and Target Operating Domain

#### 4.1.1 Background

##### 4.1.1.1 Definition of the Problem Addressed

A key stage early in the development of a safety case is the definition of the scope, including what functionalities the system is designed to perform and in what surrounding environments it will perform them. This is essential to ensure all stakeholders, including industry regulators, share a common understanding of what is in scope and what is not. Safety analysis techniques, including the assessment of risks, depend upon a clear understanding of this. Furthermore, it will be a key influence on the range of scenarios that are assessed within the test programme.

This section therefore addresses how to define the functionality that the system provides, the domain that the system is designed for, and the domain that the system will be deployed within. However, it does not address the definition of the technical solutions utilised within the system, which should also be provided in order to support the analysis of functional safety, safety of the intended function and cybersecurity, as defined within standards relating to these fields.

##### 4.1.1.2 Current State of the Art

#### Operational Design Domain and Target Deployment Domain

The acronym 'ODD' refers to the Operational Design Domain, i.e. the surrounding conditions in which the system is designed to operate. A taxonomy for how an ODD might be defined is set out within BSI PAS 1883 (2020). However, it is equally important within the safety of AVs that the scope of the intended environment in which the system will be deployed is also considered; the domain that the system is *designed* for may not match the reality of the environment that the system is *deployed* within. Figure 9 summarises how the ODD and the deployment environment relate to each other, using the name 'Target Operating Domain' (TOD) for the latter.

By definition, both the ODD and the TOD have to sit within the wider category of every permutation that is possible in the world (grey circle). In the ideal situation, the TOD should overlap with the ODD such that the system is only ever required to perform tasks it was designed for (green area). Any permutations that are within the TOD but not within the ODD (blue area) carry residual risk and should therefore be eliminated by adjusting the ODD or TOD scope when they are identified. It must be acknowledged, however, that there will remain a residual risk of permutations that are outside the ODD remaining undiscovered in the actual deployment location; due diligence in characterising the real location within the TOD definition will mitigate this risk, but not eliminate it.

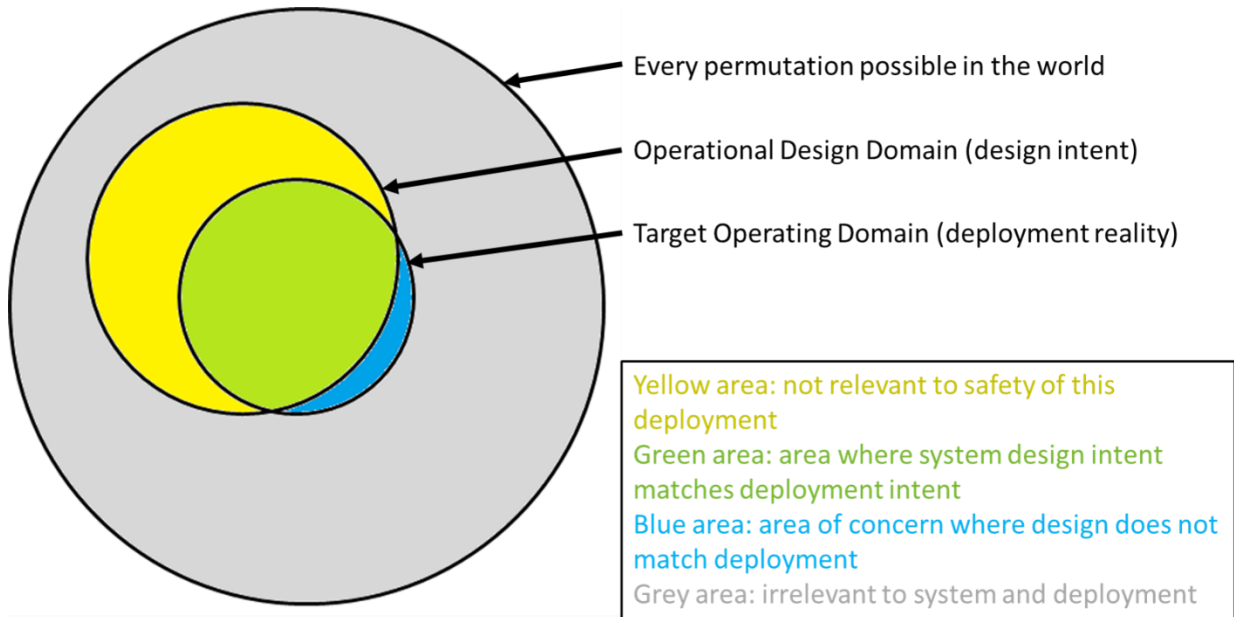


Figure 9: Venn diagram showing the relationship between the operational design domain and the target operating domain for the deployment.

As an example, an ODD may include roundabouts with a radius of lane centre between 4m and 25m, and therefore the safety assurance process, including safety analysis and testing, would need to demonstrate coverage of this range. However, a TOD for a particular deployment may only include roundabouts between 4m and 15m radius, meaning that only this range would be pertinent to that deployment (green area). Roundabouts greater than 15m and not exceeding 25m would not be pertinent to that deployment, but may be pertinent to other deployments that the vehicle type could be used for (yellow area), and any roundabouts above 25m radius would be out of scope for both (grey area). If, however, analysis of the deployment route(s) or area(s) identified a roundabout with a 3.8m radius, this information would need to be captured in the TOD, resulting in a discrepancy (blue area) that would need to be addressed.

For each individual deployment, the yellow area is therefore of low relevance, as it relates to permutations found in the real world that are not found within the particular deployment. Depending upon the amount of variety found within the possible scenarios that could occur within the ODD and the TOD, it could be the case that the yellow area is many orders of magnitude larger than the green area, meaning that analysing and testing the entire ODD would entail vastly more expenditure of time and cost relative to analysing and testing the TOD.

At the other end of the spectrum, if the system was developed solely for one particular deployment upon a specific route or in a specific geofenced area, the yellow area would cease to exist as a result of the TOD and ODD being identical, and the 'scenario space' that needs to be assessed and tested would be significantly reduced; only the smallest circle in the figure would need to be covered. Similarly, if the system was designed to cover an ODD that is broader and more abstract than one specific route or area, such that the yellow area still exists, but the regulatory testing focussed upon assuring safety only within the TOD (rather than covering the broader ODD), only the small circle would need to be assessed by the regulator.

In addition to introducing the term TOD, to contrast with ODD, this report will also refer to 'specific' and 'generic' definitions of the ODD and TOD. A specific definition of an ODD or TOD is one that is linked to one or more actual or intended deployment routes or area(s) such that any safety analysis and testing that is performed to this specific definition is particular to the characteristics of that geographical location, i.e. to one particular place within the world. On the other hand, a generic definition is one that only describes the attributes of the domain in an abstract manner, and does not link it to any particular geographical location, being potentially applicable to many deployment locations.

This report also uses the term 'COD' (Current Operating Domain) to describe the surroundings being experienced by the LSAV at any given instant in time during deployment within the real world. Whilst all CODs encountered by the system should ideally lie within the scope of both the ODD and TOD, there will always be a residual risk of encountering a COD that lies outside the ODD and/ or TOD; for example, due to occurrences that were previously unforeseen or deemed impossible (so called 'black swan' events).

## Relationship to the Test Programme

The requirements for the test programme are covered in Section 5.9, and therefore will not be examined in detail here. Nonetheless, one of the key challenges in gaining a sufficient level of safety assurance to support an approval will be gaining an adequate level of coverage of the range of scenarios that the system could reasonably be expected to encounter in service, which will require significant time and resources. As the TOD effectively defines the range of possible permutations that the test programme has to take samples from to provide such coverage of the challenges the vehicle will face in service (the 'scenario space'), the way the TOD is defined will have a key influence upon the practicability of the test programme, and hence upon the assurance and approval process as a whole. This section will therefore consider the key aspects of the test programme that relate to how the TOD and the system functionalities are defined.

As a result of a report by RAND (2016) on the required test mileages to gain statistical confidence of automated vehicle (AV) safety, there is a widespread consensus within industry that accumulating mileage within uncontrolled conditions on public roads is not, by itself, a viable solution to gain sufficient safety evidence. This has prompted a move towards 'scenario-based testing, where systems are interrogated within a designed experiment, each test case being selected such that as a collective whole, they provide suitable coverage of the challenges that an AV could face when deployed in the real world (SaFAD, 2019; ISO/TS 4804, 2020), including the less common 'edge cases'.

This requires a test programme that takes as an input the various parameters that have been identified as being applicable to a scenario (e.g. lane width, speed of cyclist) and the range of values they could assume, and selects multiple test scenarios from this multi-dimensional scenario space in order to sample system performance. The problem comes in how the number of test scenarios escalates exponentially as the number of parameters increases; if viewed as a full factorial experiment, with continuous variables quantised into a number of discrete strata, the number of test cases would be equal to the number of strata for each parameter multiplied together.

As a simplified example, if a scenario could be described with 20 parameters, and each of these parameters is quantised into 10 strata, the number of test cases would be  $10^{20}$ . This would be neither plausible nor reasonable in practice. Methods such as orthogonal arrays or Latin hypercubes with multidimensional uniformity can be used to make the sampling of the problem space sparser, and intelligent sampling methods such as equivalence partitioning, boundary testing, confounding of variables and search space optimisation can be used to focus test effort where it provides most valuable data. Such approaches would allow more efficient sampling such that the reduction in the likelihood of uncovering any given system defect is compromised as little as possible as the number of test cases reduces, but care should be taken, and it should not be assumed that down-sampling by a factor of billions or trillions to arrive at a practicable test programme could be done without any negative impact upon coverage quality (HumanDrive, 2020). This is especially true in the case of automated vehicles, where the complexity of the systems, the behaviours they must perform and the environments that they must operate within results in systems having low interrogability, making it challenging to reliably define equivalence classes and decision thresholds.

The above is not intended to be a specification for how scenario-based testing will be conducted, as this is covered elsewhere within this report and within work package 3, but is provided to highlight the following key issues that have a direct bearing upon the definition of the ODD and TOD:

- The parameters that can be varied and the ranges they can be varied within needs to be thoroughly understood in order for a scenario-based testing programme to provide acceptable coverage.
- Increasing the range that a parameter could adopt (e.g. increasing the range of lane widths that the system could encounter) will either mean more samples have to be taken (i.e. more test cases), or a lower density of sampling will have to be accepted (i.e. increased opportunity for system flaws to remain undiscovered).



- Increasing the number of parameters that can be varied within the test programme (e.g. increasing the number of roadside object types such as post boxes or traffic signs, that would each in turn have parameters to define positions, heights etc.) will increase the size of the test programme *exponentially*.
- It is therefore important that the definition used to specify the bounds that scenario-based testing must provide coverage of is sufficiently detailed and narrow to constrain the scope, such that testing the system remains practicable.

## Existing Regulations and Standards

### Advanced Driver Assistance Systems

For production vehicles, active safety features and advanced driver assistance systems (ADAS) are not linked to a specific location and are therefore able to be operated in a wide variety of locations that fit the high-level specification for the conditions in which the system is designed to operate (broadly equivalent to an ODD, although necessarily not defined as such). Given that such features are designed to supplement the performance of the human driver, who is ultimately responsible for ensuring safe performance of the driving task, there is no attempt to explore all the possible scenario permutations, including rare edge cases, within safety analysis and testing.

As such, there remain many permutations where such active safety systems will provide false negative responses (i.e. fail to intervene correctly when needed) or false positive responses (i.e. make interventions that are not appropriate) due to events that can be reasonably expected within the operating environment (equivalent to the TOD). However, this should not be regarded as unacceptable residual risk, due to the aforesaid responsibility that the driver ultimately holds.

For example, EuroNCAP has performed test procedures upon a range of active safety and ADAS features such as autonomous emergency braking and lane keep assist (EuroNCAP, 2021). The test procedures are based on extensive research into real-world accident modalities and are designed to test vehicles in ways that represent a significant number of accidents in which road users are killed or seriously injured, using simple road geometries that don't attempt to replicate specific locations. As such, the high level and abstract descriptions of the intended operating conditions that are provided to customers are entirely in keeping with the abstract and generic nature of the test scenarios, and entirely in keeping with the intended role of such systems.

The same principle does not hold true for automated driving systems of level 3 or above within the SAE definitions (SAE J3016, 2021), as without a driver who is required to be attentive to the surroundings and ready to intervene as required, the residual risk presented by scenario permutations that are not covered by a limited suite of tests becomes a major concern. Whilst there will always be some rare edge cases that remain unknown, nonetheless it is incumbent upon system manufacturers to minimise this residual risk through exhaustive analysis and testing of the range of scenario permutations.

This is consistent with ISO/PAS 21448 (2019), the standard for SOTIF (Safety of the Intended Functionality), where efforts are made to convert 'area 3' (unknown unsafe scenarios) to 'area 2' (known unsafe scenarios) through a process of discovery. Further engineering development can then be undertaken to move them from 'area 2' to 'area 1' (known safe scenarios), the scope of the system can be reduced to eliminate the problem, or a justification can be put forth that their residual risk within 'area 2' is acceptable due to the low exposure to the triggering conditions. This underlines the level of diligence and detail required when specifying a target operating domain, to ensure that the range of possible permutations is adequately identified such that safety assessments and testing that occur downstream provide appropriate coverage.

UNECE regulation 79 for steering equipment (UNECE, 2021) uses the term "boundary of functional operation" and states that this "defines the boundaries of the external physical limits within which the system is able to maintain control." Whilst the reference to a system 'boundary' rather than an 'ODD' is more in line with traditional systems engineering practice and jargon, nonetheless it is used such that it is synonymous with ODD, and is described on a 'generic' (using the jargon of this report) basis. This is appropriate as regulation 79 does not cover highly automated systems at present, and the ADAS and active safety features that are included should be expected to operate on any road that fits a broad ODD, or 'boundary of functional operation', description, with the driver remaining ultimately responsible for ensuring safe operation.

### Automated Driving

UNECE regulation 157 for ALKS (Automated Lane Keeping Systems) applies a methodology not dissimilar to EuroNCAP (ALKS, 2021), where generic tests are specified in order to target what may be expected to be the most prevalent and/ or high-risk scenarios. Whilst these scenarios do not attempt to capture the entire range of reasonably foreseeable permutations that could be encountered, the regulation does include provisions requiring manufacturers to assume responsibility for making sure the system is safe, including assurance of functional safety and SOTIF, and type approval authorities can request further tests to satisfy themselves of adequate safety. However, no framework is provided with regards to how this coverage should be achieved.

Such high-level requirements could make it challenging for manufacturers to ensure compliance, and for regulators to reach objective and consistent decisions. The regulation does take steps to mitigate this by specifying some of the aspects that should be parameterised (e.g. table 2 in Annex 4), although this is far from exhaustive. It should also be noted that ALKS represents a very different use case to the LSAVs examined within this report, with the presence of a user-in-charge and the restriction to particular traffic conditions upon divided highways arguably reducing the challenge of achieving suitable scenario coverage. The regulation does include real-world testing, but by definition it would not be possible to cover the actual deployment route(s) or area(s) within this testing as there is no defined deployment location; the regulation is purely 'generic', as opposed to 'specific'. Furthermore, as described previously, mileage accumulation is not a practicable means to provide sufficient and statistically valid assurance of safety.

ISO/PAS 22737 (2021) is a recently released safety standard for Low-Speed Automated Driving (LSAD) systems, in which a limited suite of generic scenarios are tested, with no process to adapt the test cases to ensure coverage of the challenges within the actual deployment route/ area. As an example, the 'driveable area' test uses an extremely abstract, geometrically drawn definition for an area where the driveable zone narrows, with the narrowing point being perfectly square on a perfectly straight road with dimensions that are all either directly fixed, or fixed as a function of the vehicle dimensions. This means that successful negotiation of the test scenario will provide extremely limited confidence that the system is able to operate safely within complex and chaotic real-world environments where the drivable area would typically narrow according to a more organic, less geometric layout.

The standard applies a similarly high level and generic approach to other tests, such as those for pedestrians and cyclists, where testing is limited to a small number of pre-set generic scenarios that do not parameterise the variables such that the full range of permutations of movements by the other actors is sampled, and that do not account for how the artificial set-up translates to the actual geometries that will be encountered when deployed in the real world.

As such, this approach does not adequately identify, document, analyse and test the full range of permutations that a system could encounter in the real world, compromising the level of safety assurance that is able to be provided. As the vehicle, rather than the driver, is responsible for ensuring safe operation, this lack of coverage of edge cases, or even of high-probability, day-to-day occurrences, is a significant concern. This deficit could have been addressed by developing a methodology that defines the TOD in a manner that captures the specifics of the intended deployment, and then samples realistic scenarios to gain coverage of this specific TOD.

### **ODD and Scenarios**

Significant discussion of methods to define the characteristics of the deployment domain for a system have been undertaken within the USA. In particular, a report by the Automated Vehicles Safety Consortium (AVSC), a programme created by the Society of Automotive Engineers (SAE), observes that an ODD could either be defined in a 'bottom up' manner where a specific location is identified first and then the abstract attributes of that location are identified and defined, or a 'top down' manner where the abstract attributes are described first, and then a location that fits them is subsequently identified (AVSC, 2020).

Furthermore, it is observed that current trials and pilot deployments use the 'bottom up' approach, i.e. are defined for a specific location: "A bottom-up approach leverages specific, mapped routes and makes the challenge of identifying objects and scenarios more tractable. This best practice recommends the bottom-up approach. It enables a better understanding of (local) environmental conditions, roadway geometries, physical infrastructure, zones, and the behaviours of other road users. This recommendation and approach may be modified as technology advances and more knowledge and experiences are gained". The document goes on to describe a recommended workflow for defining the ODD in a bottom-up manner, and provide a worked example.

The findings of the AVSC report are aligned with what we propose in this report; although we propose a different terminology, with the term 'TOD' introduced to cover the specific deployment location, nonetheless the intrinsic link between the system, the assurance programme and the specific deployment location is maintained.

The National Institute for Standards and Technology (NIST) have built upon the AVSC findings in a report (NIST, 2021) that proposes the term 'operational envelope specification', or OES. The OES can take three forms:

- OES<sub>Nom</sub> defines the 'nominal' conditions, i.e. the conditions that are intended and expected. This could be interpreted as being equivalent to the proposed term 'TOD' within this report, i.e. be a description of what is expected within the actual deployment, although there are some contradictory statements, and in places the OES<sub>Nom</sub> appears to be a more detailed description of the ODD, not of the TOD, or a continually evolving repository of all the parameter values observed in the OES<sub>Act</sub> (see below) instances that have been accumulated over time. The report doesn't explicitly consider the distinction between design and deployment domains.
- OES<sub>Act</sub> defines the 'actual' driving conditions. In general, this seems to refer to the surroundings of a specific vehicle in a specific location at a particular instant in time, corresponding to 'COD' in this report, although again there is a lack of consistency in how the term appears to be used throughout the document
- OES<sub>Ref</sub> is the 'reference' OES; this is a template of operating condition names and parameter definitions, which provides the format used to create an OES<sub>Nom</sub> or OES<sub>Act</sub>. The document does not attempt to list what names and definitions would be within the OES<sub>Ref</sub>; the description is purely abstract, leaving some uncertainty as to how it would work in practice.

Because of the lack of consistency and clarity in the use of the terms, and because the report represents early findings rather than a published standard, it would not be appropriate to align with the terminology they have presented; otherwise, there is a risk that the meanings of the OES definitions could deviate significantly before they reach the point of being more clearly defined, meaning that their reuse here would distort the intended meanings of the requirements and guidance. Therefore, the terms TOD and COD are retained, although it is recommended that the work on OES is monitored, and opportunities sought to align upon terminology if and when it becomes clear that the terms are synonymous. Similarly, there has been some use of the term 'TOD' within the AV standards community, but no agreed definition, and therefore this is an area of development that should be monitored to ensure that, if and when a formal definition for TOD appears within other documents, it is suitably aligned to the usage here.

In terms of providing a taxonomy for an ODD definition, there are two main references available; the PEGASUS Project, and BSI PAS 1883. The PEGASUS Project defines a hierarchy of six 'layers' that are required to describe the ODD elements that could exist within a scenario (PEGASUS, 2019). These are illustrated in Figure 10. On the other hand, BSI PAS 1883 (2020) defines three categories at the top level of the taxonomy, as shown in Figure 11.



Figure 10: Scenario layers as defined in the PEGASUS method. Source: PEGASUS (2019)

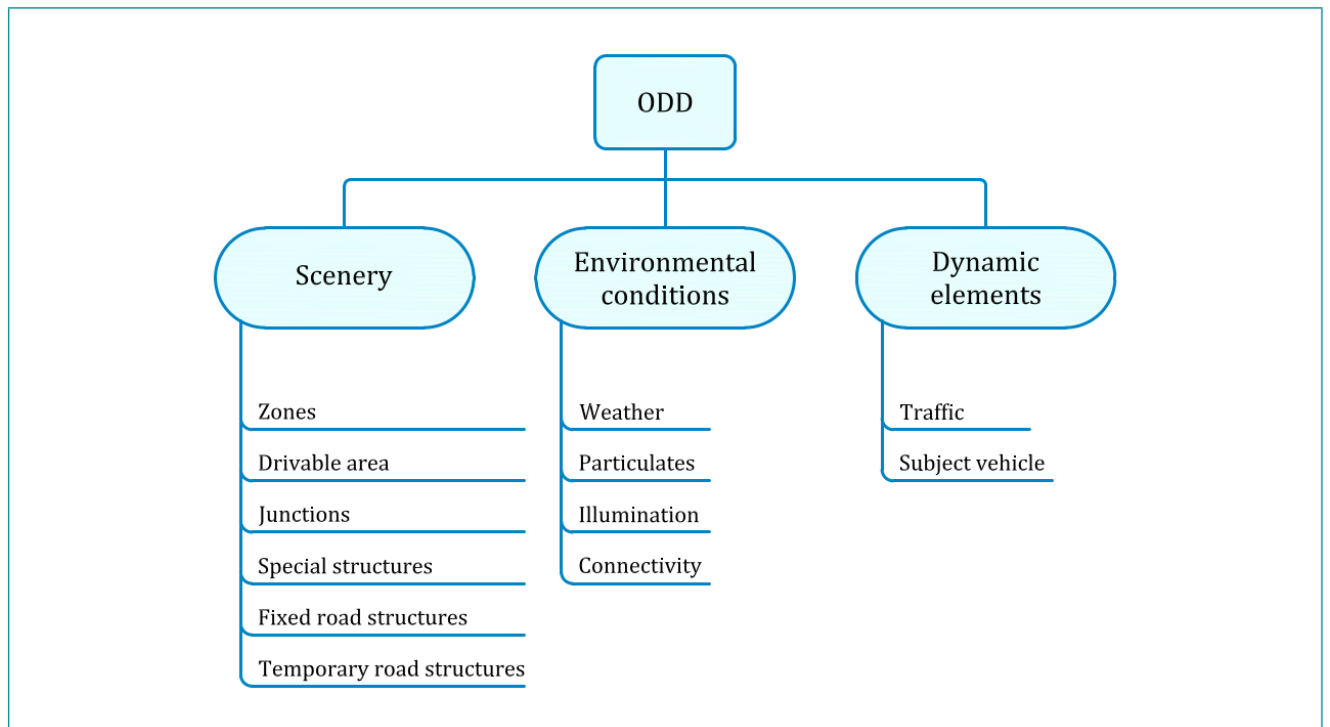


Figure 11: Categories within the ODD taxonomy. Source: BSI PAS 1883 (2020)

There are a few key points to note about the similarities and differences between the two approaches. Firstly, it can be seen that layers 1, 2 and 3 ('road level', 'traffic infrastructure' and 'temporary manipulation of layer 1 and layer 2' respectively) from PEGASUS are amalgamated into a single

category, 'scenery', in BSI PAS 1883. It could be argued that it is reasonable to put 'road level' and 'traffic infrastructure' in the same category as they will both remain relatively static throughout the deployment lifecycle (although the latter tends to be less static), meaning that the safety analysis, test programme and in-service monitoring will treat them in a similar way.

However, the 'temporary road structures' that are also placed in the scenery category of PAS 1883, and the equivalent 'temporary manipulation' within PEGASUS, is fundamentally different to manage as it remains dynamic through the deployment Lifecycle, and can be parameterised in a dynamic manner within the test programme (e.g. by placing traffic cones in various positions during physical or simulation testing). As will be examined further in Section 4.1.2.1 (suggested TOD structure), this requires it to be categorised differently when defining the target operating domain.

Similarly, 'connectivity' is placed within the 'environmental conditions' category in PAS 1883, whereas it sits within a separate 'digital information' category, Layer 6, in PEGASUS. It is questionable whether connectivity counts as an environmental condition in a common understanding of the terms, but more importantly, there is a fundamental difference in that the digital information is something that can be designed and controlled, whereas there is much less opportunity to control the other environmental parameters. This again suggests that it should form a separate category, as per the PEGASUS method.

BSI PAS 1883 also includes the 'subject vehicle' amongst the dynamic elements; this is a questionable inclusion, since the ODD is generally interpreted to be a definition of the operating conditions that can exist at the system boundary, not the system itself. Instead, it would have been more valuable to further enumerate the different road user types, such as cars, pedestrians, animals or cyclists, the category of 'traffic' providing a limited descriptor for an umbrella category of these.

BSI PAS 1883 provides some useful subcategories that can serve as a checklist for some of the permutations that should be considered; in particular, the 'environmental conditions' category includes many permutations that may not have been considered within a desk-based brainstorming activity. Unfortunately, however, the other categories, in particular the key one of 'dynamic elements', contain far less information, and therefore provide limited value in this respect.

A more fundamental issue relates to the need for the specific deployment route(s) or area(s) to be identified, recognising the need for a 'bottom up' approach, as per the above reference to the AVSC report and in line with the principles outlined in this report. While it could be argued that a definition of the specific route(s) or area(s) may be added under the 'zones' subcategory within the 'scenery' category of PAS 1883, there is no explicit mention of the need for such a key part of the definition to be included, and no description of how it should be captured. Instead, the focus is upon abstract, generic descriptions of types of features that could be encountered; whilst such descriptions, derived in a bottom-up manner from the specific route(s) or area(s), are essential to support many elements of the safety case (e.g. the functional safety analysis), they are insufficient for supporting a scenario-based testing programme for a highly-automated vehicle, or the implementation of operational safety measures, as detailed within this report. It should be noted that BSI PAS 1883 pre-dates the concept of design and deployment domains being distinct consideration (as per the AVSC and NIST reports previously referenced), perhaps explaining the lack of consideration of the specific deployment.

At the time of writing, an ISO standard for ODD definition (ISO 34502) is being drafted, which has been reported within the working groups for this project to be closely based upon PAS 1883. This is another area that should be monitored as future standards emerge, and depending upon the direction that the ISO standard takes, it may prove possible to align GB regulatory requirements with it, particularly if the aforementioned limitations of BSI PAS 1883 are addressed.

The link between the ODD and the test scenarios undertaken is summarised by the safety case guidance for users of CAM Testbed UK, produced by Zenzic (2021); this relationship is shown in Figure 12, where it can be seen how scenarios are derived from a combination of the operating conditions that the vehicle could reasonably be expected to find itself within and the behaviours that the system will be required to perform within those conditions. This shows why consideration of test programmes is of such primary importance when identifying a suitable means to define the target operating domain.

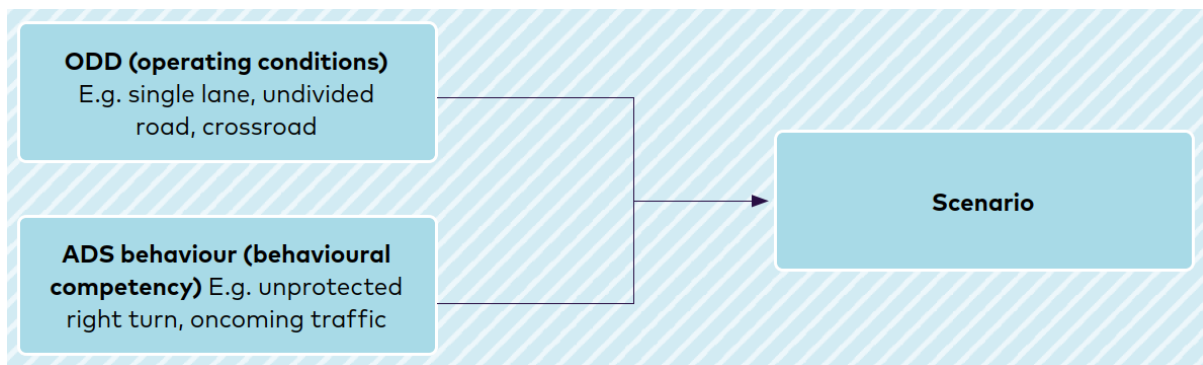


Figure 12: Relationship between operating conditions, behaviours, and scenarios. Source: Zenzic (2021)

The Zenzic guidance also notes how the ODD definition can become very complex in practice, due to interdependence between parameters: “for example, a trialling organisation may define the top speed of the vehicle as 70 mph during daytime on dry day, but may reduce the top speed to 40 mph in the presence of rain. Such interdependence can be valuable to allow the system to be exposed to the broadest range of challenges possible in the given conditions, or looked at conversely, to ensure that the most challenging situations are only presented to the vehicle where there are other factors to mitigate risk such as an absence of nearby traffic.”

Significantly, it also notes: “furthermore, the trial ODD will be intrinsically linked to the route selection; choosing an alternative route to avoid one particular ODD element may mean adding or removing other elements that are associated with the routes under consideration”. The report refers only to the ODD, and doesn’t consider the TOD as a separate aspect, as again, it pre-dates this concept.

#### UL4600

Another significant document relating to AV safety is UL4600 (2020), which describes what aspects of safety should be considered within a safety case, covering a wide range of considerations from assurance of safety management systems to fault detection and response strategies. The standard has limited references to the ODD, and no clauses that directly indicate what attributes should be defined or whether it should be linked to the specific deployment route(s) or area(s), but it does specify the need for the test programme to sample from the range of reasonably foreseeable permutations which could occur within the ODD, thereby highlighting the need for an ODD (or TOD) that is complete, accurate and practicable to test.

UL4600 describes the concept of a ‘minimum equipment list’ (MEL) within section 10.2, which is derived from the aerospace industry. This is the list of items that make up the system, that are required to be in a fault-free condition in order for the system to operate within a particular mode. While normal operation may require all, or almost all, of the items within the system to be functioning, failure of certain items might allow the system to continue operation within a degraded mode; for example, a sensor failure might result in limitations to the allowable vehicle speed, or to the environmental conditions that the system can operate within. When the lower threshold of a MEL is crossed, the system would have to enter a lower MEL that tolerates less items being available. Behaviour should also be defined for when the lowest MEL available cannot be met, e.g. performing an MRM. This concept is particularly valuable as it is important to view the design and deployment domains in the context of the behaviours the system is permitted to perform and the presence of faults rendering items unavailable; such faults may require limitations to the behaviours or the deployment domain in order to maintain acceptably safe operation.

An interesting concept introduced in UL4600 is ‘partial conformance’ (section 17.3.4 of UL4600). This allows an independent assessor to review and accept the portions of the safety case that don’t relate to the testing of the vehicle or the collection of early in-service data, thereby allowing a level of assurance to be provided before testing commences based upon non-testing aspects such as analysis of functional safety or safety management systems. This concept is developed further in Section 4.1.2.1 of this report, under the guise of ‘provisional assessment’.

#### Waymo Pilot Deployments

Waymo (2020c) describe how their systems are tested and assessed extensively in the applicable location before driverless deployments can occur, and note that: “Waymo’s operational design domain is defined by elements such as geographies, roadway types, speed range, weather, and time of day.

An operational design domain can be very limited: for instance, a single fixed route on low-speed public streets or private grounds in temperate weather conditions during daylight hours. However, Waymo aims to have a broad operational design domain to cover everyday driving. We're developing self-driving technology that can navigate complex city streets in a variety of weather conditions and times of day within broad geographic areas".

It is further noted that "Waymo's system is also designed so each vehicle does not operate outside of its approved operational design domain. For example, passengers cannot select a destination outside of our approved geography, and our software will not create a route that travels outside of a geo-fenced area, which has been mapped in detail (see "How We Build a Map for a Self-Driving Vehicle"). Similarly, our Waymo Driver is designed to automatically detect sudden changes (such as a snowstorm) that would affect safe driving within its operational design domain and come to a safe stop (i.e. achieve a "minimal risk condition") until conditions improve".

Whilst only a high-level description, this appears to show a desire to expand the operation area balanced against a pragmatic need to control the practicability and safety of the operations, and shows that the geographical location is limited by geo-fencing.

### Stakeholder Feedback

One stakeholder within the first round of consultation, representing a group of vulnerable road users, expressed the need for the full range of parameters within the operating environment to be captured; for example, a cyclist may typically travel at 10 or 15 mph, but some riders may be travelling at 30 mph, and the system needs to be able to react safely to all. They also remarked how this could change as a function of the operating environment – for example, it may not be possible for cyclists to reach higher speeds in busy urban areas – and that understanding the range of parameters will be more practical for a more limited scale deployment. Similarly, multiple interviewees commenting on the challenge of capturing the full range of environmental conditions or road users that a system could be exposed to. Many stakeholders expressed the importance of understanding the operating environment as a precursor to verifying the level of coverage provided by the test programme.

One stakeholder observed that existing AV uses are geofenced, and whilst there is an ambition to achieve level 5 systems that can go anywhere, it is not a reality yet.

Another stakeholder expressed the opinion that identifying the particular location of a deployment will be important as part of the hazard analysis and risk assessment process, with concern that such assessments would be inadequate unless they are linked to the location. This would allow consideration of particular types of hazard, such as if the route passes close to and eye hospital or a school.

The second round of stakeholder feedback was conducted later within the project, once mature proposals were in place to solicit comments upon. A survey was used, with questions targeted at areas expected to be contentious. This showed that there was a significant minority of respondents who objected to the need to define a target deployment domain, and felt that the deployment could be covered with an ODD; on balance, however, it was deemed that the design and deployment phases could be significantly different, and that there is a growing trend of referring to deployment domains within the CAV community (as identified by some respondents), and therefore that there is a justification for including TOD as well as ODD.

Similarly, there was wide disagreement on the subject of the TOD being specific to the deployment route(s) or geofenced area(s). Some of this was on the basis that an ODD could be made specific to a location, with SAE J3016 including a reference to "...geographical and time of day restrictions", which could allow inclusion of specific locations. We have opted to allow the ODD to be made specific to a location if the manufacturer so chooses, but have opted to have a separate ODD and TOD, as justified within this section, and to only make it mandatory for the TOD to be specific to the deployment location(s).

There were also concerns logged about the effect that requiring testing upon the specific location would have upon the scalability of LSAVs, with some respondents having a desire to be able to follow an 'approve once, deploy anywhere' model for economic reasons. Whilst we recognise that this should remain a long-term ambition, we, and many of the respondents, do not believe this is realistic given the current and foreseeable state of the art. Therefore, whilst the requirements should seek scalability where this is practical, this should not come at the expense of safety assurance.

It should be noted that, regardless of the requirement to perform scenario-based testing upon the specific route(s) or geographical area(s) of the deployment, operational safety measures for the deployment assurance will require specific locations to be defined to support assurance steps such as the operational risk assessment and the review of the suitability of the route, for which there was widespread support. It is also worth noting that there was strong support for requiring scenario-based testing upon specific locations (and their representative equivalents such as digital twins), with only one respondent objecting. This is in agreement with the evidence and proposals set out within this section, but by definition would require the specific locations to be defined in order to support the testing.

Therefore, it is recommended that the TOD should be specific to the deployment location, as should a significant proportion of the scenario-based testing programme, but that opportunities should be sought to allow scalability where this is possible without compromising the level of safety assurance – for example, by allowing a lower volume of testing on the specific route(s) or area(s) to be justified by the manufacturer if they have already acquired evidence from prior deployments that the system is able to operate safely in similar environments and is able to extrapolate to new ones successfully. Nonetheless, some testing upon the specific route(s) or area(s) should be required for all deployments within the foreseeable future, in order to provide adequate confidence that the system will operate safely within the particular characteristics and demands of the deployment.

### Approval Phases

CertiCAV (2021) recommended a two-stage process where the system receives an approval covering its entire operational design domain in phase 1 (which would be to a ‘generic ODD’ under the nomenclature used within this report), and each specific deployment of that system is then subjected to an operational approval process in phase 2. This is illustrated in Figure 13. The result of this would be that the scenario-based testing programme would need to parameterise all attributes of the physical infrastructure such that all the possible permutations within the ODD are covered, followed by phase 2 merely covering operational safety and not providing scenario-based testing coverage of the system safety upon the actual locations that the system would operate upon.

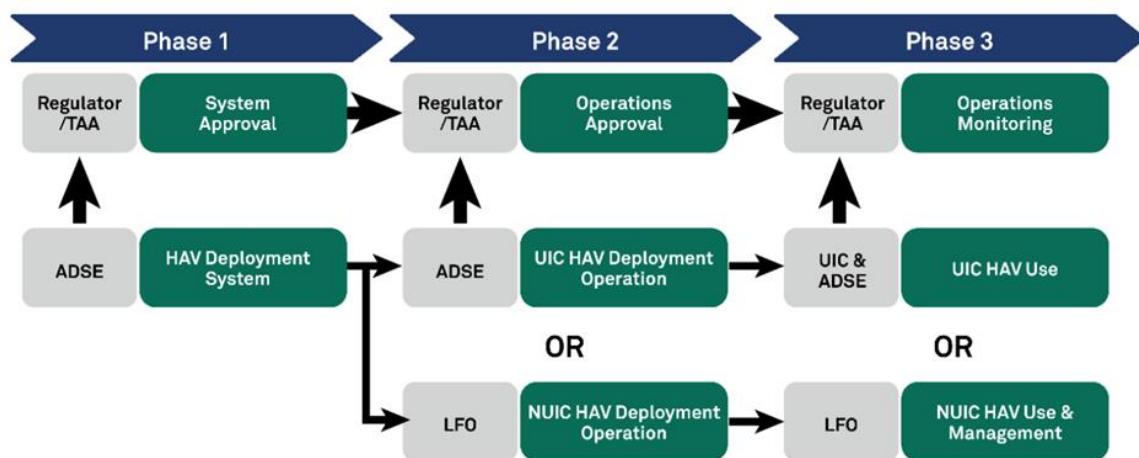


Figure 13: Illustration of how a multi-phase approvals process might work. Source: CertiCAV (2021)

For reasons outlined subsequently within this report, it would not be feasible within the foreseeable future to parameterise all aspects of the road and road infrastructure within phase one. Therefore, this approach would only be viable if the domain definition used as the basis for the test programme in phase 1 met the requirements this report proposes for a TOD, including the need for it to be ‘specific’ to the deployment route(s) or area(s).

Whilst scenario-based testing needs to be linked to a ‘specific’ TOD, it is likely to be feasible to approve many other elements of the safety case on a ‘generic’ basis, where these elements do not depend on specific road geometries (e.g. functional safety analysis, subsystem verification tests), provided that the scenario-based testing programme takes place within a subsequent ‘specific’ phase. This suggests the need for a flexible approach that can accommodate some of the safety evidence being created and approved for a TOD that defines a specific location, and other evidence being approved for a generic



ODD that could be applied to multiple TODs; indeed, some aspects of the safety case, such as electromagnetic compatibility, could even be agnostic to the ODD. This would allow as much of the scalability proposed within the CertiCAV report to be retained, whilst also making safety assurance more practicable.

Regulations under development in Germany, as presented to UN ECE WP29 GRVA (UN ECE, 2021), propose a three-phase process, as follows:

1. Approval from national type approval authority, including audit, simulation and real-world test driving
2. Regional approval to verify that the vehicle is able to adapt fully to automated driving on the specified ODD ("limited geographically to a defined environment, e.g. A to B or even A<sup>2n</sup>")
3. Registration and issuing of license plate, based on approvals in previous two phases

The intention is that the verification within phase 2 would consist of a finite number of days operating the system on the intended route(s), i.e. mileage accumulations rather than scenario-based testing. It is unclear how the regional authority would work with the national authority, or what testing would be in scope for each phase. However, it suggests that some level of approval to a generic ODD is followed by some level of approval to a specific TOD (in the terminology of this report), with phase 3 being merely administrative.

The limitation of this approach is that phase 1, much like the phase 1 defined in CertiCAV, would require parameterisation of every possible road and road infrastructure parameter within a 'generic' ODD; given the number of extra parameters that would have to be introduced to cover lane width, bend radius, camber etc., a test programme to cover such a generic domain would require significantly more test cases and/ or would result in significantly sparser coverage. The testing that is for a specific TOD, in phase 2, is a welcome acknowledgement of the need to assess the system on the actual deployment TOD, but the safety assurance able to be provided by a limited duration mileage accumulation programme will in practice be inadequate to gain suitable confidence that the performance is acceptably safe (RAND, 2016).

It should also be noted that the Law Commissions (2021) provide proposals for how the phases within a GB approval scheme for AVs should be structured, and this should also be a key consideration within the development of a GB Safety and Security Scheme for LSAVs.

### 4.1.1.3 Conclusions Drawn

#### Limitations of Generic Definitions

For a manufacturer's scenario-based testing programme to provide sufficient assurance that 'unknown unsafe' scenarios (in ISO/PAS 21448, or SOTIF, parlance) have been searched for, it must provide coverage that is representative of every possible road layout that could be encountered. For a 'specific' location, this would entail testing the actual road layouts that exist in the route(s) or geofenced area(s). However, if the testing is to be done on a generic basis, this would require significant research in order to identify all reasonably foreseeable parameters and ranges that can be found within Great Britain, followed by a test programme that provides coverage of them. This would involve parameterising attributes such as (to provide non-exhaustive examples):

- Lane width
- Bend radius
- Gradient
- Angle of junction
- Number of roads at junction
- Radius of roundabout
- Number of lanes
- Dimensions and locations of obstacles obscuring the field of view
- Camber
- Crown height
- Lane marking type
- Lane marking degradation

- Nature of the road edge (e.g. pavement, grass verge, barrier)
- Road surface material
- Surface damage
- Presence of surfaces that may cause sensor problems (e.g. steel drain covers with radar)
- Prevailing speed limit
- Presence of priority markings/ signs such as 'stop' or 'give way'
- Presence of traffic lights, zebra crossings etc.

Even with a subset of examples such as that above, it can readily be appreciated how quickly the complexity of the test programme will escalate. This parameterisation would have to be performed in addition to parameterisation of the other, non-static elements within the scene, such as the weather or other road users, creating a truly vast scenario space.

Furthermore, note that each parameter can assume many values within a single scene; for example, parameters such as lane width, gradient or bend radius could vary continuously throughout the scene, resulting in infinite possible permutations for each single parameter itself, even before considering the possible combinations with other parameters.

Providing coverage of the range of permutations that can be found in the fixed infrastructure, as per the parameter examples above, would also be challenging on a practical level; whilst it may be possible to find real locations that feature a value that closely matches what was selected within the sampling methodology used to create the test programme (e.g. a lane width of 2.4m, a bend radius of 36m or a junction angle of 85°), it may prove significantly more challenging to find real locations that match all of these parameters, and will become exponentially harder when the full suite of fixed infrastructure parameters is identified and locations have to be found for a vast number of test scenarios. This can be partially overcome by modelling generic road layouts to fit these attribute values in simulation or upon a proving ground, but real-world testing would still be required to provide validation that such modelling is representative.

### Advantages of Specific TODs

The scope defined for this report limits the operating environment to “a fixed route of fixed geographical area”; as has been examined, linking a test programme to a fixed location in such a manner presents benefits in terms of the size of the ‘problem space’ that the test programme needs to provide coverage of.

In addition to lowering the number of test cases required to gain an equivalent quality of coverage, this will also reduce the problem of identifying all the physical permutations that can be found in the real world, as only a limited, finite area needs to be investigated. It will also make it more practicable to develop processes for parameterising the various scenario attributes to construct a test programme, as there would be no need to parameterise any of the fixed physical attributes of the scene, only the non-fixed attributes. The use of ‘specific’ TODs could also be argued to simplify the task of identifying what dynamic elements need to be included as parameters, and in what ranges, as it becomes more practicable to survey road use at a finite number of locations than to attempt to form a generalised model of what events and scenario permutations occur in the wider world.

The test programme would therefore take each location within the deployment route or area (e.g. a stretch of straight road, a T junction, or a roundabout) and ‘fuzz’ the system with the full range of scenario permutations that could be experienced upon it (e.g. the presence of different types of road users and the different trajectories they could follow). Each location within the route(s) or area(s) would require this fuzzing, meaning that the number of test cases for each location is multiplied by the number of locations to arrive at the overall test volume (broadly speaking – in practice, different locations may need a different volume of tests).

This is vastly preferable to attempting to provide coverage on a ‘generic’ basis where, as described previously, the number of test cases goes up exponentially as the number of parameters that can be varied increases. By removing the need to parameterise aspects such as lane width, camber, junction angle etc., the size of the problem space is significantly reduced, as is the reliance on the system to be able to interpolate between sample points with respect to the fixed infrastructure.

Thought about the other way, if a company wanted to deploy a system on a specific location in Milton Keynes and safety of this was to be demonstrated solely by testing upon roads in Greenwich and Coventry (or their digital twins), without the actual deployment location in Milton Keynes being included,

there would be a risk that differences between the deployment and test locations could result in hazards being undetected.

For example, slight differences in the layout of line markings could cause significant differences in system behaviour, invalidating test evidence from locations that were presumed to be similar. Furthermore, the differences that trigger such responses may appear insignificant to human eyes, or may go unnoticed altogether, such as an object that reflects radar in a manner that confuses sensors or an object that momentarily blocks the line of sight at a junction. This would make it very difficult to be confident that none of the locations in the Milton Keynes deployment contain any such differences to the test locations in Greenwich and Coventry; assurance of the absence of such triggers could only be gained by testing upon the actual route in Milton Keynes, and this evidence would therefore need to be on a 'specific' basis.

As a point of reference, it may be helpful to consider the pilot driverless deployment operated in Phoenix, Arizona by Waymo. This has included the accumulation of extensive mileage within the specific location of the driverless service, and indeed for every other specific location where they have conducted testing that is yet to result in a fully driverless public service, with simulation and physical testing in a controlled environment also used extensively to replicate scenarios relating to the specific deployment location (Waymo, 2020a; Waymo, 2020b). Even after the vast, and arguably unprecedented, volume of safety testing that they have undertaken, it is not possible to transplant the vehicle to a new location and commence driverless operations without having to undertake significant testing that exposes the system to a new range of scenarios that are representative of the new location.

On this basis, it is concluded that even the most apparently simple deployments need to be analysed, tested and approved for the specific deployment location, and the specific features contained therein. Whilst the volume of new analysis and testing can be adjusted to be proportionate to both the complexity of the TOD and the level of confidence build up from past deployments (a 'proven in use' argument in functional safety terminology), it would not be appropriate to allow any deployments to commence without appropriate consideration of the TOD, and without that TOD defining a specific geographical location.

This does not mean that locations outside the deployment route(s) or area(s) cannot be tested upon or used within the safety evidence. It may be the case that the manufacturer of the system wishes to use it across multiple deployments and therefore does development work and testing that covers a range of locations (and potentially entirely fictitious locations created in simulation), and as noted above, it would be possible to argue in the safety case that successful operation in previous deployments that feature similar scenarios means that less testing on a new deployment location is required to gain acceptable confidence in the safety of the system. However, even in the case where safety evidence has been accumulated for other locations, it would still be necessary to 'top up' the test evidence with data for the new location, and to reconsider other aspects of the safety case (e.g. SOTIF) to either argue that they remain valid for the new location, or to update them as necessary.

This represents a significant difference in comparison to existing type approval and driver licensing; existing type approvals are only specific to the vehicle type and existing driving licences are only specific to the person and the broad category of vehicle, but neither specify limitations to the specific route(s) that the approval applies to. However, such an approach is not unprecedented in other industries; in the rail industry, for example, drivers (who perform the equivalent function to an ADS) have to be trained and assessed as competent for both the specific locomotive or multiple unit they will be driving and for the specific route that they are driving it upon (RSSB, 2022). Drivers of road vehicles are approved on a 'generic' basis, but this relies upon humans possessing general intelligence such that they can interpolate and extrapolate from past experiences to make sense of new and unfamiliar ones; artificial general intelligence has not been achieved, and therefore machines cannot be relied upon to adapt in the same way.

It should be noted that concerns relating to the 'brittleness' of system behaviour when faced with subtle changes between scenarios also apply where the characteristics of a specific location change over time, such as through wear and tear, seasonal effects or roadworks. Management of this is considered within Section 6.2, and changes identified within the TOD should be subject to an impact assessment, triggering an update to the safety case where appropriate.

## Why Tightening a Generic ODD is Insufficient

A completely generic approval of test evidence within a safety case would be a go-anywhere solution that allows manufacturers to deploy the system to multiple different locations, with no further testing being needed to revalidate the vehicle for each deployment. Whilst superficially attractive due to the scalability, this report has already detailed why it would not be practicable to test and assess a system upon a generic basis.

There have been some suggestions within the working groups for this project that an alternative solution would be to use a 'generic' ODD, but to define the attributes of this ODD so restrictively (i.e. decomposing the attributes into extremely fine detail and limiting each to as narrow a range as possible) such that the size of the 'scenario space' that has to be considered within the safety testing and analysis is restricted. However, the following limitations of such an approach should be considered:

- If the ODD is described tightly enough that only one viable deployment location can be found within the world that matches it, then a 'generic' approach would offer no scaleability advantage over a 'specific' approach. Indeed, unless a specific location is considered when defining such a tight ODD, there would be a risk that, after all the investment in the test programme was completed, no viable deployment could be found to make use of the approval. It is therefore advised that at least one specific location that is commercially viable for deployment should always be identified prior to the type approval process commencing, and that steps are taken early in the process to confirm this viability with any authorities who would be stakeholders in approving this deployment.
- If the ODD is not defined so tightly that it is only limited to a single deployment location, then road layouts and roadside infrastructure would need to be parameterised to ensure every possible deployment location is covered. As has been examined previously, this presents difficulties due to the scale of the 'problem space', the practicalities of identifying suitable test locations that provide the right combination of parameters, and the residual risk that subtle differences between ostensibly similar scenes may trigger hazardous behaviour.

## The Need for Data Acquisition

Much of the above analysis has focussed upon the current state of the art regarding the specification of the roads and associated infrastructure. However, it is equally important to define the other, less stable, attributes of the TOD: the temporary infrastructure, dynamic elements, environment and digital information. The latter can largely be defined based upon the design of the system and any offboard infrastructure that forms part of the wider system; the 'highly automated supersystem (HASS) under the definition proposed by the CertiCAV (2021) project.

However, temporary infrastructure, dynamic elements and the environment are challenging to define as they depend upon factors outside the control of the manufacturer or operator, such as the presence and behaviour of other road users or the weather at any given point in time. It is therefore necessary for the safety case to include consideration of what permutations could occur within these categories (e.g. presence of an e-scooter, or volcanic ash in the atmosphere) and what range of values could parameters associated with these permutations assume (e.g. possible speeds and directions of e-scooter, and possible sizes and concentrations of particulate matter in the air). Only by having a comprehensive understanding of what could be experienced in the real world is it possible to determine what is and isn't within the TOD of the system.

For example, it could legitimately be decided that operating in volcanic ash levels above a certain threshold is not permitted, and therefore it will be defined as being outside the TOD. However, if the possibility of significant levels of volcanic ash was never considered in the first place, the system design would not be analysed and tested to assess suitability for operation when there is a significant concentration of volcanic ash in the air, and there would be no clear definition of how the system should respond in such circumstances.

This highlights how important it is to collect significant volumes of data on what permutations are reasonably foreseeable within the TOD; the safety case should therefore be expected to show due diligence in undertaking a study of what permutations could be experienced, such that these permutations can be further considered in the safety case.

Many of these permutations will follow some form of distribution (e.g. a Gaussian distribution, or 'bell curve'), with some values for a parameter occurring far more frequently than others. This will be

important in the creation of a test programme to ensure that the distribution of scenarios considers, even if it doesn't attempt to replicate, real-world distributions (see section 5.9), and will also be important in determining what residual risk can be accepted in terms of triggering events for SOTIF hazards.

In the longer term, it is proposed that a study is undertaken to capture a large volume of data on what permutations and parameter ranges can be observed in the real world, in order to support the creation and assessment of the safety case. However, in the absence of any such database, the onus should be upon the manufacturer to undertake a suitable study to acquire the required information. This could include, for example, surveys of road use within the specific deployment route(s) or area(s) to better understand what road user types and movements may be expected, or a study of historical weather data to understand what weather events are reasonably foreseeable. Naturally, such a task becomes significantly easier when considering a specific location rather than having to consider all possibilities within Great Britain, although care should be taken when undertaking surveys to ensure sufficient data is collected over an appropriate time period – for example, data collection during the summer would fail to capture weather permutations or pedestrian clothing styles observed in winter.

### Reasonableness of ODD and TOD Definitions

In terms of defining what is in or out of scope, objects and attributes of the scenes surrounding the system can broadly be placed into three categories:

- Indisputably within the ODD/TOD: situations that the vehicle is expected to be able to handle well without interruption to 'normal' operation.
- Indisputably outside the ODD/TOD: situations for which the vehicle is not intended to continue operation, requiring an MRM, but where the MRM is 'non-emergency'.
  - For example, poor weather may justify the vehicle stopping, but does not require an emergency response, as the rain is not an obstacle that could be struck.
  - In some cases, it will be possible to transition to a broader ODD/TOD that provides equivalent safety by limiting aspects of the vehicle safety (e.g. one that accommodates heavy rain by reducing the maximum permitted vehicle speed); again, the same principle applies that the parameter would be indisputably outside the 1<sup>st</sup> ODD/TOD, and the transition to the 2<sup>nd</sup> ODD/TOD would be a non-emergency one.
- Grey area for ODD/TOD: situations in which the system is not intended to continue normal operation, but may have to perform an emergency avoidance manoeuvre for the hazard that this situation presents.

Although the first two categories are clear cut, there is no guidance available in any of the literature on whether the third category should be classed as in or out of the ODD and TOD. When the possibility of suddenly encountering something (e.g. horse rider, e-scooter) cannot be eliminated, and there could be a need to perform an emergency reaction to that very feature immediately. There are two possibilities for how this could be accounted for within the safety assurance process:

- (1) Include it in the ODD and TOD such that the avoidance is part of the 'normal' (i.e. within intended conditions) DDT (even if it may result in more conservative behaviour than 'normal' driving such that it looks not dissimilar to an MRM).
- (2) Exclude it from the ODD and TOD such that avoidance is via an MRM (or a transition to a broader ODD/TOD associated with degraded functionality), as a result of exiting the allowable operating conditions (even if it may require evasive manoeuvring that looks not dissimilar to performing the 'normal' DDT).

Non-urgent MRMs as a result of situations that fall under the 'indisputably outside the ODD/TOD' category (e.g. heavy rain, snow) can be treated as an equivalence class for test purposes, such that the scenario-based testing programme would need to test each MRM in a range of scenario permutations, but would not need to parameterise each triggering condition that leads to the MRM along with all the other test parameters (light level, speed of vehicle ahead etc.). The triggering conditions for the MRM (i.e. the vehicle's ability to detect when it has exited the TOD) would need to be analysed and tested, but this could be undertaken via traditional systems engineering verification methods, without requiring scenario-based testing. For example, the ability of the vehicle to detect temperatures below

the allowable limit would not vary as a function of whether a pedestrian crossing the road is moving at three or four mph.

However, occurrences that could require an immediate and bespoke response would need full parameterisation within the test programme. For example, an e-scooter would need to be able to be detected and classified at a range of speeds, positions and angles, and the response of the vehicle would also vary according to the behaviour of the e-scooter, including the need to react to emergency scenarios where heavy braking or swerving is required to avoid or mitigate a collision. This would be true whether the permutation falls under the 'indisputably within the ODD' category (i.e. the manufacturer wishes the system to be able to operate in the vicinity of e-scooters as part of its normal operation) or under the 'grey area' (i.e. where the manufacturer is not intending for the system to operate in the vicinity of e-scooters as part of normal operation, but nonetheless their presence cannot be ruled out and therefore the system may have to react urgently).

Because of the similarity between safety assurance needs for occurrences that are 'indisputably within the ODD/TOD' and occurrences that fall within the 'grey area', this report recommends that the requirements mandate option 1 above for permutations that fall within the 'grey area', i.e. that any objects or attributes that can't be eliminated from the vicinity of the vehicle, and that the vehicle could potentially have to react to urgently, should always be classed as part of both the ODD and the TOD. This will ensure that the safety assurance process is in line with that for the intended 'normal' driving conditions such that the full range of parameter combinations is considered. It will also avoid the risk of potentially hazardous situations being casually overlooked by stakeholders on the basis that "we don't need to worry about those, they're outside the ODD".

## 4.1.2 Recommendations

### 4.1.2.1 Proposed Requirements

#### Specifying the Design Scope

The manufacturer applying for system approval shall provide an ODD specification that defines the attributes of the operating environment that the system is designed to function in. This shall be to a sufficient level of detail to support the analysis of all downstream safety analysis and testing aspects, such as functional safety and safety of the intended functionality (SOTIF), and shall define what features and parameter ranges are in scope with regards to:

- The road level, including characteristics such as surface material, friction levels and acceptable geometry ranges (e.g. maximum gradient, minimum width)
- The traffic infrastructure, including road signs and lane markings that the system is designed to detect and speed limit zones the system is allowed to enter
- Temporary infrastructure such as traffic cones, temporary lane markings and moveable barriers
- Dynamic elements such as cars, cyclists, pedestrians or animals that the system is designed to react appropriately to
- Environmental conditions that the system is designed to operate within, including weather, lighting and particulates

If the system relies upon digital information such as wireless communications, wired communications or digital maps in order to operate safely, these shall also be included within the ODD definition. See Section 5.6 for more information on external inputs (wired and wireless communications).

It is permissible for the manufacturer to provide multiple ODD definitions in order to accommodate degraded performance; for example, operation may be restricted to a narrower range of road level or traffic infrastructure permutations if rain, snow or fog exceed a defined threshold.

The manufacturer shall define a set of behavioural competencies that the system is designed to perform. This shall include:

- Functionality that is essential for safe operation within the ODD, e.g. follow path defined by lane markings (if the ODD features lane markings), adjust speed in response to vehicle ahead (if the ODD includes other vehicles).
- Manoeuvres that are not essential for safety but the LSAV is nonetheless designed to perform, e.g. lane change to overtake slower vehicle, reverse bay parking.
- The underlying OEDR (object and event detection and response) competencies that are required to complete the above behaviours. For example, 'detect parking bay markings' or 'detect static objects to the rear of the vehicle' would, inter alia, be required to perform reverse bay parking.

Requirements for defining behavioural competencies are examined in more detail within Section 4.3.

It is permissible to provide multiple behavioural competency definitions such that degraded functionality may be provided in response to certain ODD permutations (e.g. reverse bay parking only being available if rain does not exceed a threshold or if no pedestrians are present) or to allow for faults within subsystems.

The manufacturer shall define the internal system requirements for operation, including a 'minimum equipment list' (MEL) to define what elements internal to the vehicle or system (e.g. sensors, processors, actuators) must be functioning with no detected faults present in order for a journey to commence or continue. Where a manufacturer has provided multiple ODDs or multiple behavioural competency definitions, it is permissible to provide multiple MELs such that degraded functionality may be provided in response to subsystems or component being unavailable (e.g. operation being limited to a narrower range of weather conditions in response to a sensor being unavailable).

A MEL should include not just functional failures where a subsystem or component becomes incapacitated in a discontinuous, binary sense, but also failures where performance levels on a continuous spectrum fall outside a permissible tolerance band, e.g. where the soiling of a sensor or wear of an actuator results in the performance lying outside the acceptable range. Capturing such acceptable ranges within the MEL will be important to facilitate consideration of these factors within the analysis and testing of the system, e.g. by allowing testing to consider the worst-case permutation, or by parameterising within the range as part of the test programme.

The ODD(s), Behavioural Competency Definition(s) and MEL(s) shall be combined to produce a definition of the overall 'System Design Capability', which defines the overall scope of operations that have been designed for. This shall include at least one ODD definition, at least one Behavioural Competency Definition, and at least one MEL definition. Where more than one of any of these are provided, mapping between the definitions shall be provided such that every possible combination is defined as either being permitted or not permitted. An example of this is shown in Figure 14, where a full factorial list of possible permutations for a fictitious system has been provided, each has been defined as to whether it is a permutation that the system is permitted to enter or not, and a justification has been given.

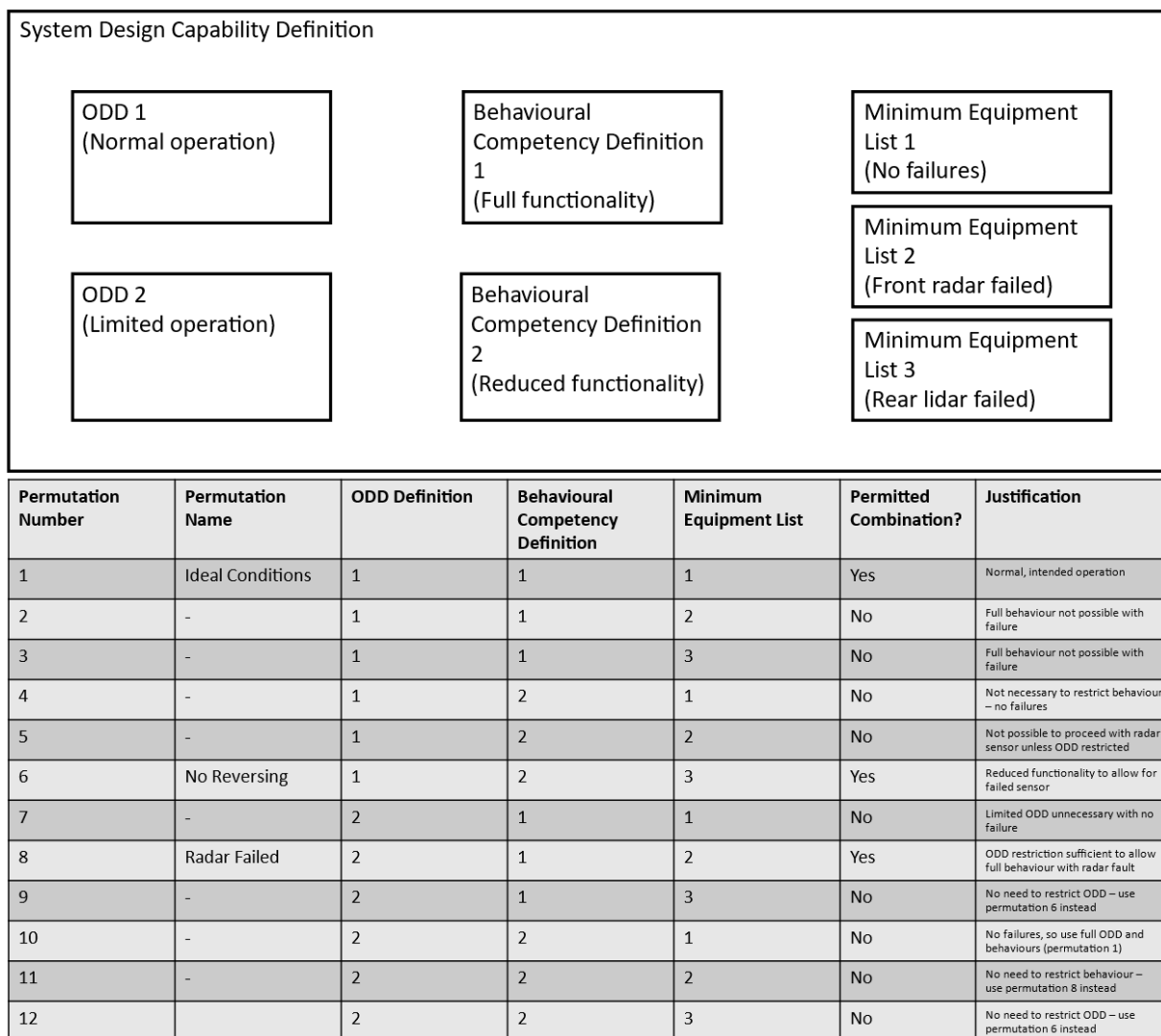


Figure 14: Definition of the possible ODDs, Behaviour Definitions and MELs that are available for a fictitious system, and a table defining which combinations are permitted as part of the system design

The System Capability Definition submitted shall be sufficient to support a ‘provisional assessment’ by the regulator; a provisional assessment, defined in more detail later in this section, is an optional interim review of the aspects of the safety case that are not specific to a particular deployment location, allowing the assessment of such evidence to be common to multiple deployments. The information provided within the System Capability Definition must therefore provide at least as fine a level of detail as the system and its surroundings are considered at within any evidence submitted for provisional assessment (e.g. the functional safety, SOTIF and cybersecurity analyses); otherwise, it would not be possible to validate that the safety case is compatible with the characteristics of a deployment. The advantages of a provisional assessment are that it allows a level of assurance to be gained before testing starts on public roads and that it allows safety evidence that has been deemed compliant at this point to be reused across multiple deployments, thereby enhancing scaleability and efficiency where possible.

### Specifying the Deployment Scope

The Operator shall submit a definition for the Target Operating Domain (TOD) in order to support analysis of the operational safety of the deployment; this shall form part of the deployment safety case report. However, it is also required that a TOD is submitted in order to support the testing of the vehicle upon the actual route(s) or area(s) of the deployment; as will be subsequently examined, it remains an open point as to whether this testing of the vehicle should form part of the evidence within the Vehicle Safety Case Report or the Deployment Safety Case Report; if the former, it shall also be required that the manufacturer submit a TOD during the vehicle type approval phase in order to support this testing.



The TOD shall contain all the top-level categories that were required for the ODD. However, it may optionally choose to further refine the level of detail such that the TOD may form a subset of the ODD. Over and above this, the TOD definition shall also include a definition of the specific route(s) or geofenced area(s) that will be used within the deployment, requiring a clear definition of the specific location(s) involved (e.g. defined via latitude and longitude coordinates) such that the location within the world is unambiguous to all stakeholders. An example of a suitable structure for a TOD definition is shown in Figure 15; however, it should be noted that sections 2a and 2b of the TOD may optionally be amalgamated, that section 6 may be omitted if the system does not rely upon digital information in order to perform the DDT safely, and that other structures may optionally be used provided they capture equivalent information.

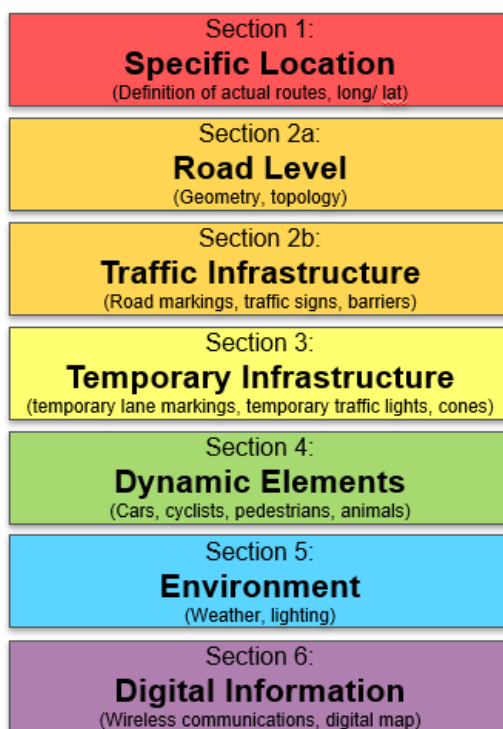


Figure 15: Example of how a TOD definition could be structured in order to provide the required information

Where the specific location refers to a predefined area rather than a predefined route or set of routes, it is permissible for that area to exclude certain routes or geographical areas within the wider area, such that there are 'islands' within which the AV is not permitted to operate. Indeed, it is expected that many deployments will feature such limitations, so that operation upon unsuitable roads can be excluded whilst still permitting a viable service elsewhere within the locality. Where such exclusions exist, they shall be clearly documented within the TOD.

Although provisional assessment of some elements of the safety case, prior to completion of the test programme, may be undertaken by the regulator based upon an ODD within a System *Design* Capability Definition that is 'generic' (i.e. describes attributes of a domain in an abstract way, but does not define the specific location), a complete approval for a system shall only be granted once the TOD has been added to create a 'System *Deployment* Capability Definition' such that the definition is now 'specific' (i.e. identified as being applicable only to the specific deployment location or routes). Optionally, the System Deployment Capability Definition could include a reduced subset of behavioural competencies compared to those given within the System Design Capability Definition; for example, 'turn left at T junction' could be removed if there are no junctions in the targeted deployment route where this would occur, thereby simplifying the test programme. The MEL specification, however, shall not change from that used within the System Design Capability Definition, as any such change would invalidate any safety assessments performed by the regulator on the basis of the System Design Capability Definition.

It is mandatory for the TOD to be submitted to the regulator prior to a test programme being assessed, as the identification of the specific deployment route(s) or area(s) is essential to allow a test programme

to be practicable whilst providing appropriate safety assurance. To undertake a test programme that attempts to cover any location that fits the generic ODD would require an implausibly large number of test cases to allow appropriate coverage of the possible scenario permutations, or looked at alternatively, would provide insufficient coverage of the possible road layout permutations that a system could encounter when deployed. Furthermore, parameterising fixed infrastructure to create a test programme that provides appropriate sampling of the range of permutations would pose significant practical difficulties.

As such, given the current novelty and immaturity of the technology, whilst test data from other locations can be used to support a safety argument, it is a requirement that the majority of the whole-vehicle test scenarios used as evidence of safe operation shall be for the specific route(s) or area(s) of the intended deployment, and should provide coverage of the entirety of the route(s) or area(s); it is not permissible to deploy an AV in a location where it has not been tested, or not been tested sufficiently to provide reasonable likelihood of detecting hazardous behaviour that may result from particular characteristics of the specific route(s) or area(s). Note that as well as testing upon the actual route, this 'specific' test evidence could also make use of representative equivalents of the actual deployment, such as a digital twin in simulation or a mock-up upon a proving ground, provided that such testing is validated by assessing correlation against real-world tests.

A definition of the System Deployment Capability Definition for the specific route(s) or area(s) is also essential to support the assessment of operational safety within the Deployment Safety Case Report, which should consider the particular hazards and operational practicalities (i.e. traffic flows and how they may be affected) that are specific to the location.

It should be noted that it is permissible for System Design Capability Definition and the System Deployment Capability Definition to be identical. This would be the case if the system is designed, analysed and tested with one single deployment location in mind. In order for the System Design Capability Definition and the System Deployment Capability Definition to be identical, the ODD would have to be identical to the TOD, meaning that the ODD must meet the TOD requirements, including the identification of the 'specific' route(s) or area(s). The two permissible methods, where the ODD is 'specific' or 'generic', are shown in Figure 16. Although starting with a generic provisional assessment may be attractive because of the ability to reuse it across multiple deployments, nonetheless it should be considered whether this is the optimal approach given that reference to the specific deployment route(s) or area(s) may make the task of defining a complete set of ODD attributes, and of analysing the system design, significantly more practicable.

Any provisional system assessment shall not be classed as an approval, and no approval certification shall be issued; it shall merely consist of reports such as test reports and audit reports for some aspects of the safety case prior to the System Deployment Capability Definition being available, and therefore prior to the point where it is possible to type approve the system. Such test reports may then be used across multiple type approvals, in line with existing practice for approving traditional vehicles.

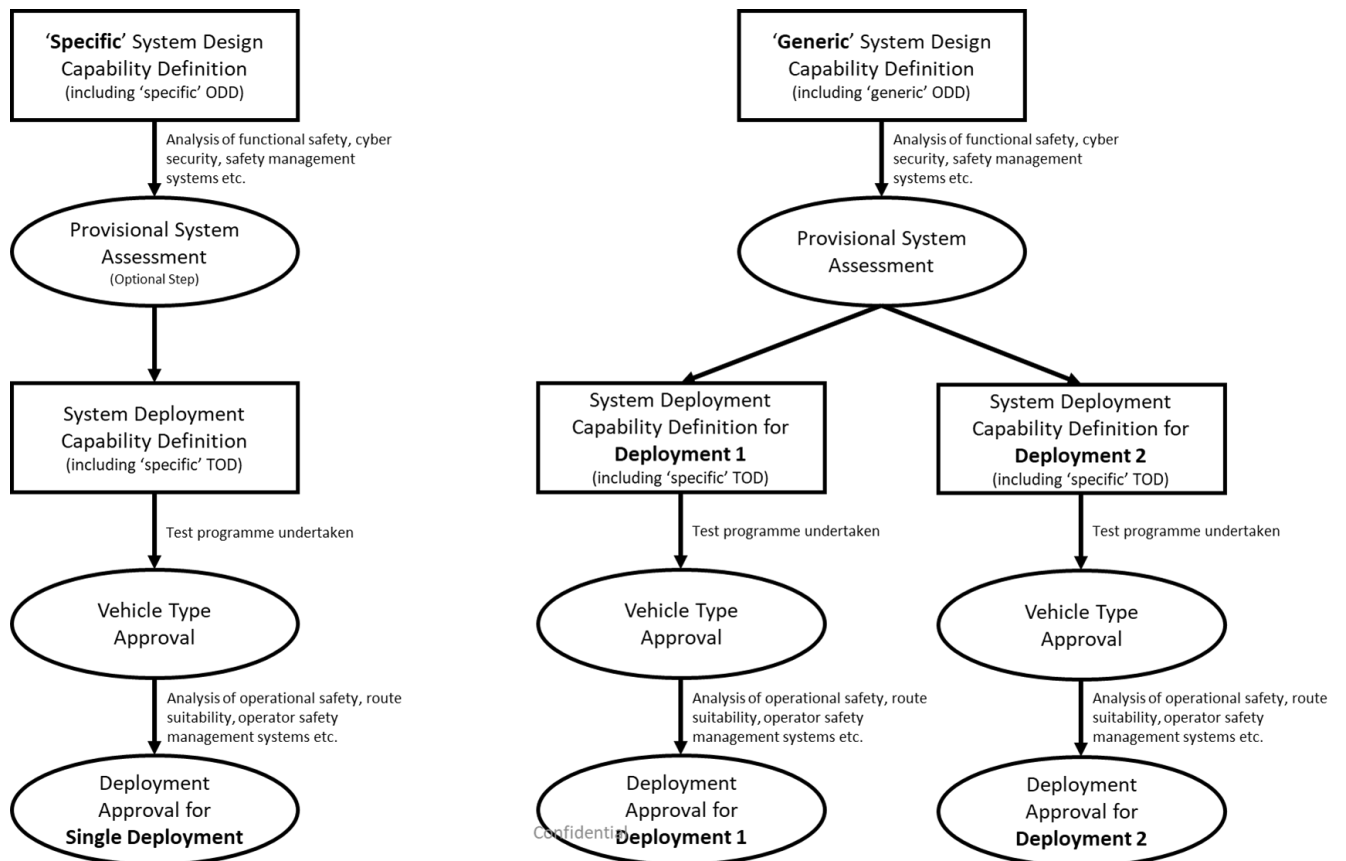


Figure 16: Illustration of how a 'specific' design definition is suitable when the system is developed and tested for a single deployment, whereas a 'generic' design definition supports a provisional approval that can be re-used for multiple deployments

### Applicability of ODD and TOD to System and Deployment Safety Case Reports

Within meetings to support the production of this report, there have been differing opinions as to whether the scenario-based testing of the vehicle upon the actual deployment location(s) should form part of the vehicle or the deployment approval phases, and therefore be evidenced within the vehicle or the deployment safety case report, with no clear consensus emerging for either approach. As described previously, the scenario-based testing upon the deployment route(s) or area(s) would need to provide coverage of a scope that is defined by a TOD rather than an ODD, such that it is specific to the challenges presented by those route(s) or area(s). This therefore means there are two possible approaches for how the ODD and TOD could be allocated to phases in the regulatory assurance process:

1. If scenario-based testing were to take place within the vehicle approval, the TOD would be needed within the system safety case to support this. As such, the system safety case would include both the ODD and the TOD, as per the approach provisionally proposed above, with the caveat that the ODD could be made identical to the TOD if the manufacturer does not wish to exploit the scalability of performing some aspects of the safety assurance against a generic ODD (noting that it is mandatory for the TOD to be specific, whereas it is permissible for the ODD to be made either specific or generic). This has the potential disadvantage of complexity due to the split between ODD and TOD within the system approval phase, as was shown in Figure 16.
2. If scenario-based testing upon the deployment route(s) or area(s) were to take place within the deployment approval, there would be no need to define the TOD within the system safety case. This would create the potentially simpler solution of only requiring the ODD within the system safety case and only requiring the TOD within the deployment safety case. Potential disadvantages of this approach are that it would need to be ensured that adequate technical

expertise is available for the regulator to audit the test programme within the deployment phase, which will be otherwise less technical, and there could be question marks as to whether a system should be described as 'approved' and the end of the system approval step if that system has not been subjected to the scenario-based testing programme.

Ultimately, this decision will depend significantly upon the regulators ability to provide such technical oversight within the deployment phase, which will in turn depend upon decisions about which regulatory bodies should be responsible for each phase. As such, whilst this report has provisionally put forth option 1 as a means to ensure robust test assurance within the system approval, the decision should be left open at this point such that an informed choice can be made when the regulatory structure has been further defined; this report concludes that either option 1 or option 2 would be appropriate.

If option 1 is ultimately selected, it should be considered whether the regulatory process could be further simplified by removing the ODD altogether and requiring a TOD for both phases, given that a separate ODD would not be necessary for all approvals. The scalability of being able to reuse some elements of the safety case across multiple deployments would not necessarily be lost, as it would be possible for the regulatory body to reuse test reports or other such assessment documentation across multiple approvals provided that compatibility is shown.

Lack of an established consensus on this topic within meetings to support the development of the report precludes a prescriptive approach, so this is left open for future consideration. However, from a technical assurance and a risk management perspective, it is of limited consequence which approach is selected provided that thorough scenario-based testing upon the actual deployment location(s) forms part of the Safety and Security scheme *at some stage*. As such, whilst the decision on the administrative approach is deferred, the achievement of a level of consensus on the need for 'specific' testing supported by a 'specific' TOD should be seen as a key point of progress.

### ODD and TOD Reasonableness

Whilst some aspects of what is included or excluded from the ODD and TOD are at the manufacturer discretion, depending upon the functionality and performance that the system is desired to be capable of, there are some attributes of the ODD and TOD that have direct safety implications. The ODD and TOD shall therefore be audited by the regulator to ensure that there are no omissions that could compromise safety. An object or attribute shall be recorded as being within scope in both the ODD and TOD, and may not be regarded as out of scope, if both of the following are satisfied:

- It is reasonably foreseeable that the vehicle will encounter the object or attribute (i.e. it cannot be eliminated from the vicinity of operations, and the exposure to it cannot be made so rare that the residual risk can be accepted)
- The vehicle could be required to take immediate emergency avoiding action (e.g. braking or swerving) to avoid or mitigate a collision as a direct consequence of that object or attribute being present.

This is to ensure that the vehicle is designed and tested such that it is able to respond safely to all permutations it may be expected to experience in service, thereby ensuring that all events that need complex and or emergency manoeuvres are assessed with at least the same thoroughness as the main behaviours within the dynamic driving task. If the permutation does not require a bespoke emergency response, and can be addressed with a defined MRM or a transition to a different ODD/TOD, then it can be defined as outside the ODD.

For example:

- For a deployment on a university campus, the presence of e-scooters typically cannot be ruled out, and it could be necessary to take emergency action to avoid them. Therefore, they must be in the TOD.
- For a deployment that is exposed to the open air, heavy rain cannot be prevented. However, there would be no need to take emergency action to avoid the rain, or any object present as a direct, immediate consequence of the rain, meaning that an MRM could be performed in a controlled manner. Therefore, rain that exceeds a quantitative threshold could be excluded from the TOD.

- For a deployment in a 30mph zone, other vehicles operating at 70mph could result in emergency avoiding action being necessary to avoid a collision. However, if it can be argued that the exposure to such a permutation is sufficiently rare, it could be excluded from the TOD. Exclusion of such outlier behaviour can help make development and testing of the system more practicable, but would require evidential justification.

Whilst the above principle shall be the primary means of determining ODD and TOD reasonableness, including for rare permutations, it is anticipated that there will be broad trends in what is permissible for each of the subcategories within the definitions:

- 'Road level' or 'traffic infrastructure' elements that are part of the deployment or could foreseeably become part of the deployment (e.g. damaged sign through vandalism) would have to be included within the ODD and TOD.
- Temporary infrastructure permutations could be excluded from the ODD and TOD provided that either the exposure to them is suitably rare or that the vehicle is able to detect them with enough prior notice that emergency avoidance is not required (allowing a controlled MRM before the vehicle reaches the infrastructure).
- Dynamic elements that the vehicle can reasonably be expected to encounter would have to be included within the ODD and TOD - even if the vehicle responds in a conservative manner by slowing or stopping such that it looks not dissimilar to an MRM, the object detection and responses would have to be fully assessed, including being parameterised within the test programme, as opposed to an MRM which could be treated as an equivalence class.
- Environmental elements may be excluded from the ODD and TOD (an exception would be where a sudden response may be needed, such as if there is potential for an LSAV to be exposed to waves breaking over a sea wall).
- Digital information may be excluded from the ODD and TOD if the system doesn't depend upon them for safe operation.

### 4.1.3 Supporting Guidance

#### 4.1.3.1 Supporting Information on Requirements

The underlying premise of the proposed requirements relating to the definition of the ODD, TOD, operational behaviours and minimum equipment list(s) is that whilst there may in some cases be a desire on the part of the manufacturer to develop a system that can be used within many deployments, including some not foreseen at the time of type approval, nonetheless it must be recognised that the testing of the complete vehicle within representative scenarios must include significant testing that provides coverage of all locations within the specific deployment route(s) or geofenced area(s).

This is to ensure that the test programme is able to provide appropriate coverage of the range of permutations possible, which becomes problematic if the characteristics of the road and surrounding infrastructure have to be parametrised along with the dynamic and environmental aspects due to the size and complexity of the resulting test programme. Furthermore, over-reliance upon test evidence from other locations risks hazardous behaviour resulting from subtle differences between ostensibly similar locations remaining undiscovered until the system is deployed.

Therefore, whilst some elements of the approval can optionally be undertaken on a 'generic' basis, the regulatory text ensures that the scenario-based testing and mileage accumulation testing elements of the regulatory process must be done on a 'specific' basis, i.e. must be primarily evidenced by test data for the deployment route(s) or area(s). This may include testing on the actual deployment location, testing on a mock-up of the location within a private facility, or testing upon a digital twin; for the latter two, it is essential that robust evidence is provided to validate the realism of the scenarios through comparing correlation to the real-world location. The testing should provide coverage of every possible location within the deployment route(s) or area(s), approached from every direction possible, with the non-fixed elements such as environmental conditions and dynamic actors being parameterised to gain acceptable coverage for each location. Robust identification of the location(s) within the TOD is therefore an essential precursor to the safety assurance.

Furthermore, once the system has been approved, it will be subjected to a deployment approval, which must also be conducted on a 'specific' basis to allow consideration of the practicalities of operating on that route or area; for example, there may be particular concerns around disrupting traffic flows at a particular junction that is a pinch-point for traffic flows during rush hour. In practice, flexibility should be allowed for the deployment to be approved concurrently, or potentially even before, the system. This is to allow manufacturers and operators to reduce the risk that significant resources could be invested in undertaking the scenario-based testing of the system, only for it to be rejected on the basis of traffic management or concerns about competition with other local transport systems, for example.

### 4.1.3.2 How ODDs & TODs could be defined in practice

It is envisaged that the different sections within the ODD and TOD would use different methodologies to identify what permutations could be encountered and determine what is in or out of scope. In the case of the 'road level' and 'traffic infrastructure' sections (2a and 2b in the proposed ontology, although as noted previously, it would be acceptable to combine them into a single section), it may typically be the case that the generic information in these sections would be directly derived from examining what physical permutations are present in the specific location (see Figure 17). However, if the ODD was only described generically (i.e. section 1, defining the specific location, was blank), it would be necessary when defining the TOD to research possible deployment route(s) or area(s) to find one where all the locations conform to the generic description that had been within the ODD; otherwise, aspects of the safety case that were produced to match the ODD would be invalidated. Naturally, there would be significant concern at the prospect of no viable deployment route or area being identified that exactly matches the generic ODD, which is why it is envisaged, and recommended, that manufacturers and operators should have the specific route(s) or area(s) in mind from commencement of the regulatory process.

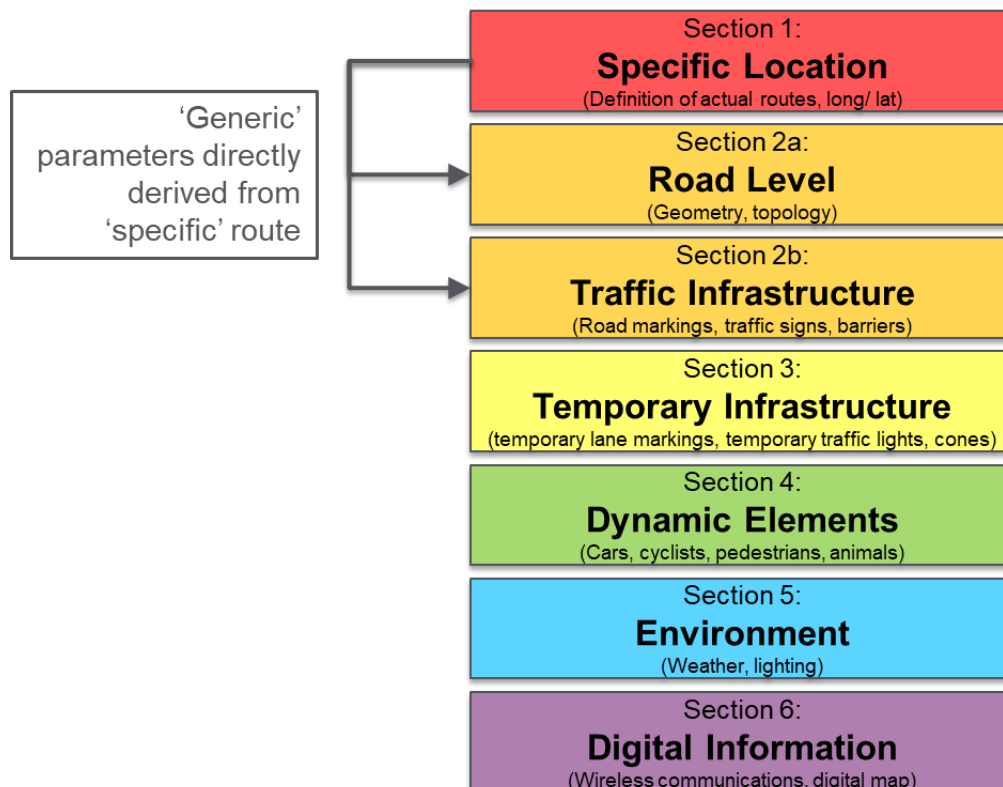


Figure 17: 'Road Level' and 'Traffic Infrastructure' sections can be directly derived from examination of what features exist within a specific location identified for deployments

The elements within the deployment that can change on a continuous, dynamic basis, on the other hand, cannot be identified in such a concrete manner from examination of an actual route(s) or area(s),

and will therefore need studies to determine what permutations can reasonably be expected to occur, estimating future probabilities based upon past statistics. This could, for example, include surveying a specific route over time (e.g. to understand what other vehicles are likely to be present and assess probability distributions), studying wider UK patterns (e.g. data on weather within the region) and consideration of the limitations that would be imposed within the deployment (e.g. limitations on the times of day for operation may affect ranges of lighting or temperature expected). This is illustrated in Figure 18; again, note that the arrows would be reversed if the objective is to find a specific TOD that matches a pre-existing generic ODD, and that this would carry the aforesaid risks relating to being confident that a suitable deployment domain is available.

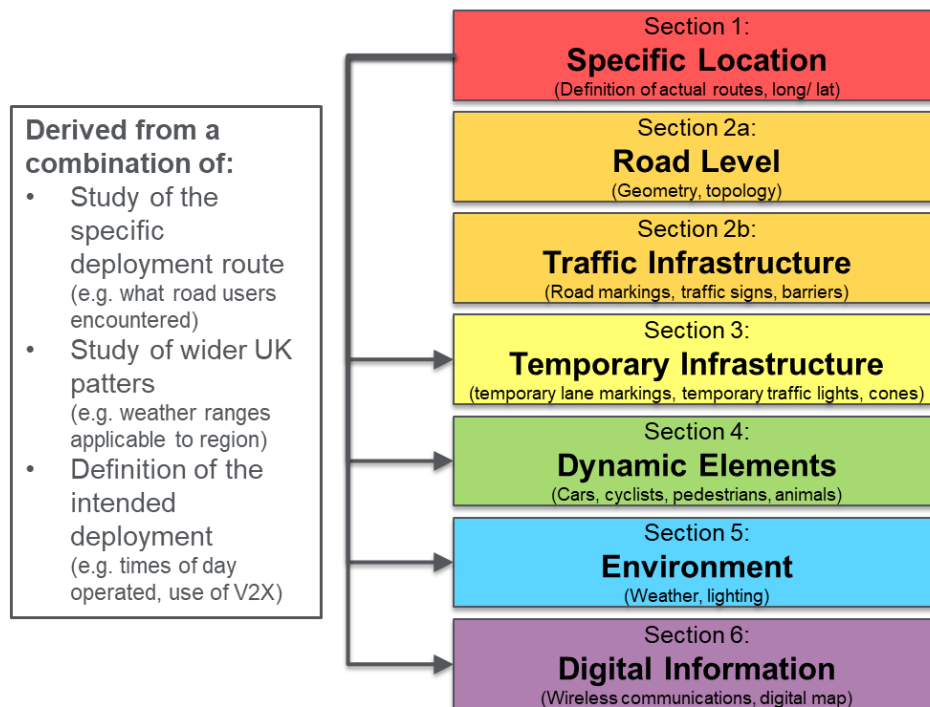


Figure 18: Defining the 'Temporary Infrastructure', 'Dynamic Elements', 'Environment' and 'Digital Information' that could be encountered in a 'Specific Location', to facilitate determination of what is in scope for the ODD/TOD

Where there is a wish to apply elements of a system safety case approval to new deployment routes or areas, the attributes that are described generically must correspond – the requirements for confirming such a match should be aligned with the requirements for confirming compatibility between an ODD and TOD, which is examined within Section 4.2 of this report. This will ensure that any aspects of the safety case that are based upon a generic ODD description (e.g. functional safety) are able to be reused without safety being compromised. However, by definition, it would not be possible to carry across the specific location section to the new deployment location, meaning that any portions of the safety case that are specific to the original deployment location (e.g. scenario-based testing or operational risk assessment) would have to be reassessed.

As was the case for the first deployment, the matching of the specific location to the generic attributes (sections 2-6) in deployment 2 could occur in either direction, i.e. the generic attributes could be checked for compatibility with an already-known location (and indeed, the original definition may have borne in mind this second deployment from the start), or a new deployment location could be sought that matches the pre-existing generic attributes. The relationship between the ODD and/or TOD sections for two deployments are illustrated in Figure 19.

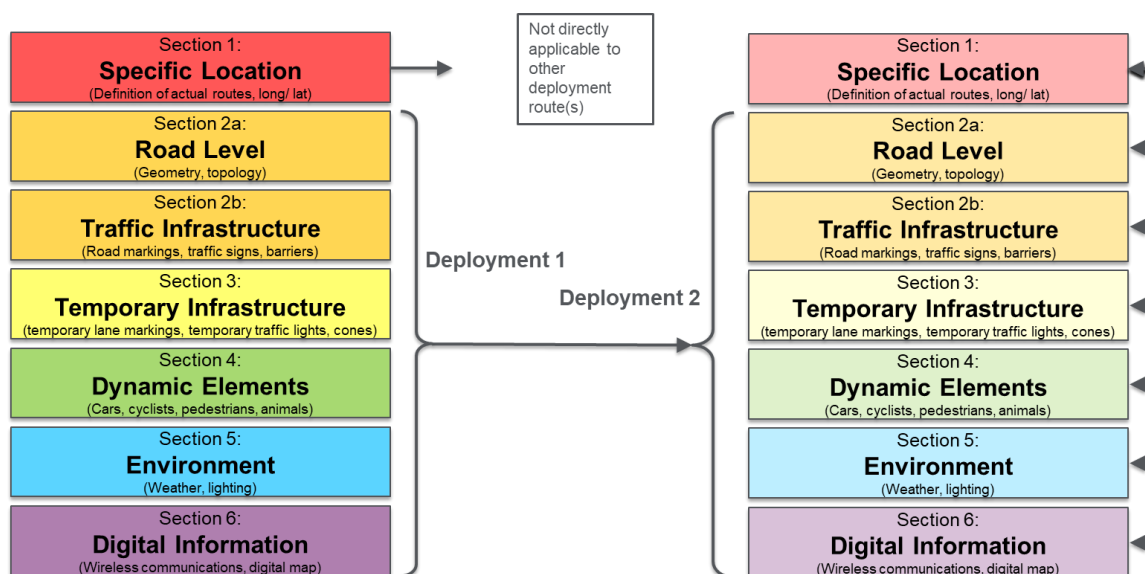


Figure 19: Illustration of how the generic parts of a safety case could be carried across provided that the generic attributes of the TOD match (sections 2a to 6), but the 'specific location' (section 1) would by definition be different for each, so safety evidence tied to this could not be reused

## 4.1.4 Future Considerations

### 4.1.4.1 Areas for Further Work

There is currently limited harmonisation within the field of ODD specification. Furthermore, whilst the industry is starting to recognise the need for a means to define the deployment domain, this has not been captured beyond the very early work published by AVSC and NIST. It is therefore recommended that regulators should monitor work being undertaken in this field and seek opportunities for harmonisation. In particular, the OES concept presented by NIST may, once it is more clearly defined, prove to be compatible with the TOD definition in this report. Ongoing standardisation relating to the ODD should also be monitored. There is also significant work taking place within the industry in relation to scenario definition languages, and this should again be monitored due to the inherent link between scenario and ODD/TOD definitions, which ultimately seek to describe the same attributes.

This report has deliberately avoided prescribing a single structure for ODD definition, because there is no universal agreement within industry on what requirements a format needs to satisfy in order to provide utility, or indeed whether there is evidence that standardisation on an ODD format will provide value at this point in time. As such, whilst it is advised that the rapidly proliferating standards landscape is monitored, arriving at a single approach that is ready to be universally mandated should not be seen as a priority at this point.

The concept of a MEL has been used in other industries, and is used within UL4600. Much like for the ODD and TOD, there doesn't appear to be an immediate need to standardise upon a single format; as long as the manufacturer arrives at a format that allows them to capture the relevant information, it will fulfil the purposes of all stakeholders. Therefore, there is no recommendation for further work to standardise MELs – such standardisation done too early would not be productive and could actually inhibit development – but nonetheless, it is advised that any working groups or standards body developing proposals in that area should be monitored.

There is even less standardisation when it comes to behavioural competencies, and whilst many stakeholders within the project have acknowledged the need for these to be defined such that the scope of a system and deployment and be understood and a representative test programme undertaken, there is limited guidance in this field. As per the above considerations, it is not advised that an attempt at standardisation would provide value at this point in time, but nonetheless, further research to better understand how operational behaviours and behavioural competencies could be identified, defined,



analysed and tested would help inform both manufacturers and regulators, thereby supporting more informed safety cases and regulatory decisions.

The concept of a 'system design capability definition' and 'system deployment capability definition' is a new one within this report; while previous reports such as CertiCAV (2021) have referred to an operational design condition (ODC) that captures both the ODD and any internal failures (as per the MEL concept used here), such an approach omits consideration of the behaviours that the system is able to provide. In practice, it is expected that the domain the system operates in, the subsystems that are available in a fault-free condition, and the manoeuvres that the system is permitted to undertake will be intrinsically linked. For example, a failure could result in a reduction in the behaviours permissible and/ or a reduction in the range of conditions permitted in the TOD, or deterioration of weather outside the bounds of the normal operation TOD may result in transitioning to a TOD that corresponds with a different, more limited set of behaviours.

As such, it is recommended that the concept of an overall definition that covers the external surroundings, the internal subsystem availability status and the behaviours available should be explored further, and consensus for such an approach sought in international discussions. This will result in a far more holistic understanding of the system than mere categorisation of the ODD, which only forms part of the full picture.

#### 4.1.4.2 Effects of Technology Evolution

At some point in the future, processes may be developed whereby systems can be tested and approved on a 'generic' basis such that the test evidence provides coverage of a potentially limitless number of physical locations. Furthermore, AV technology may develop such that systems can be trusted to interpolate and extrapolate sufficiently to be able to be deployed in new locations where they haven't previously been tested. This is not the case now or in the foreseeable future, and therefore it is important that requirements mandate 'specific' testing; provisions to allow for 'generic' testing to accommodate the hoped-for developments in the long-term future would introduce risk of systems using such provisions in a way that wasn't intended in the short term, compromising safety.

However, if and when a situation arose where safety assurance testing of the whole vehicle in realistic scenarios can be undertaken on a generic basis, the requirements would have to change to accommodate this. In particular, it would mean that, whilst the deployment approval given to the operator would remain on a specific (TOD) basis, the entirety of the system approval could be done generically, and therefore the system could be approved solely against the ODD. Any claims of systems being ready for such 'go anywhere' approvals should be treated with caution, however, and robust evidence should be required to support such claims.

#### 4.1.4.3 Future Expansion

It should be noted that the scope of this report is for vehicles that fit the SAE level 4 automated driving description, and that operate in a fully-driverless manner. Therefore, the requirements and guidance have been shaped to suit vehicles that operate in complex areas without a human present to act as a fallback, meaning very high robustness is required, even in the face of a vast range of challenging scenarios that could materialise in service.

As such, the demands placed upon the vehicle, and upon the regulatory system, are significantly different to those for a Level 3 ALKS system, which is restricted to road types that are relatively simple and consistent in nature and have a user-in-charge. If the contents of this report were to be adapted to suit systems intended for divided highways within specific traffic conditions and with a user-in-charge present, it may be reasonable to relax the absolute requirement that testing should be performed upon the actual deployment route(s) or area(s); indeed, retaining such a requirement would be clearly infeasible for an ALKS system that is not restricted to any specific route(s) or area(s).

## 4.2 Validating the Compatibility of the System and its Deployment

### 4.2.1 Background

#### 4.2.1.1 Definition of Problem Addressed

A key component of the process will be to ensure that the design of the system is compatible with the route(s) or area(s) upon which it will be deployed. This is necessary to ensure that any safety analysis or testing of the design, and any approval of the resulting safety evidence by a regulator, is truly applicable to the challenges that the system will face when deployed; without this assurance, there would be no traceability between the design assurance and the deployment, and therefore no means to know whether the system is operating within conditions for which the safety case remains valid. The definition of the design and deployment domains was examined within section 4.1, and this section builds upon these definitions by examining the level of decomposition to which the regulator should require the definitions of the ODD and TOD to be aligned.

#### 4.2.1.2 Current State of the Art

The various standards relating to the definition of the ODD have been reviewed in detail within Section 4.1.1.2, and therefore will not be repeated here. There is limited literature regarding the definition of the deployment (as opposed to design) domain, hence the introduction of the 'Target Operating Domain' (TOD) within this report. As a result, there is no literature available that directly sets out how design and deployment domains should be compared.

One standard that is relevant is the ISO standard for safety of the intended functionality (ISO/PAS 21448, 2019). This examines how to assess the suitability of the system design for providing the required behaviours within the required operating environment. As such, whilst it doesn't identify the deployment domain as a separate concept to the design domain, nonetheless it does capture many of the underlying principles required to ensure compatibility. In particular, the standard details how the process needs to identify 'unknown unsafe scenarios' ('Area 3') such that they become 'known unsafe scenarios' ('Area 2'). Following this, they can then be addressed via modifications to the system such that they become 'known safe scenarios' ('Area 1'), or can be moved out of scope for the system via tailoring of the system functionality or ODD.

As such, SOTIF provides a key methodology to compare the ODD with the needs of the vehicle in operation, and it is therefore essential that the TOD, and the validation of the compatibility between the ODD and TOD, is aligned with the needs of SOTIF assurance.

The Zencic Safety Case Framework (Zencic, 2021) notes how Area 3 (unknown unsafe scenarios) could be further broken down into two subcategories, which present fundamentally different challenges:

- a) Untested on the vehicle but captured within database;
- b) Untested on vehicle and not captured in database.

It further states that "The first subcategory contains scenarios that have not yet been discovered to be hazardous for the ADS, but will be discovered to be in due time as the test programme proceeds. However, the second subcategory is more difficult to address, as scenario-based testing will only expose the ADS to scenario permutations that are already known to be possible within that ODD, leaving the possibility of flaws in the ADS remaining uncovered due to gaps in the database. This highlights the importance of a comprehensive database when applying scenario-based testing."

The guidance refers to scenario databases, on the principal that the coverage provided by testing can only be as comprehensive as the set of scenarios identified within the database of scenarios. However, as the scenarios need to provide coverage of the range of possible permutations within the TOD, it therefore follows that failure to adequately identify the possible permutations within the TOD will ultimately result in the failure of the test programme to uncover any unknown hazardous behaviours that relate to unidentified permutations.

ISO/PAS 21448 sets out a process for SOTIF assurance that allows iteration such that the system, functionality and ODD are progressively updated, and the hazards reassessed until acceptable safety is arrived at. In many cases, this will involve progressive decomposition of attributes within the ODD definition; for example, if the ODD includes operation upon 'roundabouts', but it is identified that the system doesn't have the necessary sensor suite to support lane changes whilst negotiating a roundabout, the ODD could be decomposed such that single-lane and multi-lane roundabouts form a separate category, with one being in scope and one out of scope.

The SOTIF process within ISO/PAS 21448 includes both a desk-based analysis of the suitability of the design (e.g. are the sensor fields of view sufficient to cover the required detection capabilities) and the testing described above. Whilst the desk-based aspect could potentially be done on a 'generic' basis, a manufacturer may find it more convenient to do on a 'specific' basis, as consideration of the permutations possible within the deployment route(s) or area(s) will support a clear understanding of what is required for the system to operate safely. However, as described in detail within Section 4.1, the scenario-based test programme, an essential component of SOTIF assurance, would need to be performed for the specific deployment route(s) or area(s). This results in a situation where it would be possible for some of the SOTIF case to be done on a generic and some on a specific basis. Where this is the case, it would be essential that the ODD and TOD definitions match to the level of decomposition used within the SOTIF analysis to ensure neither part of the analysis is invalidated by the system being designed for or operated within a domain which doesn't match a portion of the SOTIF case.

The automotive functional safety standard, ISO 26262 (2018) also requires a level of consideration of the surrounding environment in order to assess the risks posed by failures; for example, the severity and controllability scores for a hazard may be different for a system that operates on 20mph urban roads, in comparison to a system that operates upon 70mph motorways. However, this analysis is able to be done at a high level of abstraction, and there is typically no need to decompose attributes of the ODD to a low level (indeed, the standard cautions against excessive decomposition, as this can result unrepresentatively low exposure scores for each hazard). Therefore, the functional safety analysis is suitable to do on a generic basis such that a single functional safety case for a vehicle type would be able to cover multiple deployments.

BSI PAS 1883 (2020) sets out an ontology for how an ODD could be defined at a high level, but notes that ODDs can be defined to different levels of abstraction/ detail depending on the stakeholders it is aimed at. It provides no indication of what level of abstraction would be appropriate for any particular stakeholder, or how to undertake an assessment of the ODD, however. Similarly, the PEGASUS (2019) methodology only examines the high-level breakdown of categories and doesn't specify an appropriate level of decomposition.

### 4.2.1.3 Conclusions Drawn

Where an ODD has been used that isn't identical to the TOD, it is essential to ensure that any safety evidence, and its review by the regulator, is compatible with the TOD such that it is not invalidated. Therefore, it needs to be ensured that:

- Every object, attribute or event that is listed as in or out of scope in the ODD is also given the same categorisation within the TOD.
- The ranges permitted within the ODD are equal to or greater than those within the TOD (i.e. they must either match, or the TOD must be a subset of a wider range available within the ODD).

It therefore follows that the level of detail to which the definitions match must be at least equal to the level of detail within the ODD definition; the TOD can optionally decompose attributes by further subdividing them such that only portions are in scope, but may not stop at a lower level of decomposition than the ODD. The manufacturer may wish to select the level of decomposition in the ODD to be the minimum required (i.e. at a high level of abstraction) in order to complete all the analyses it is used as the basis for (e.g. functional safety, pedestrian protection), as any decomposition beyond this would restrict the ability to re-use safety evidence upon new deployments whilst providing no extra safety assurance.

However, in addition to the ODD matching the TOD, it is also vital that the TOD is an accurate reflection of the real world. Regulators must therefore be satisfied that due diligence has been shown in identifying the range of permutations that could be encountered, including consideration of typical conditions and

how they may vary over time (e.g. leaf fall in autumn), reasonably-foreseeable edge cases, and a process to monitor the attributes of the deployment domain whilst the vehicle is in service to identify any changes or omissions.

As such, there would be two stages of validation required, as illustrated in Figure 20. If the ODD is identical to the TOD, then the validation between the two can be minimised, requiring only an overcheck to ensure there are no administrative errors in copying one to the other.

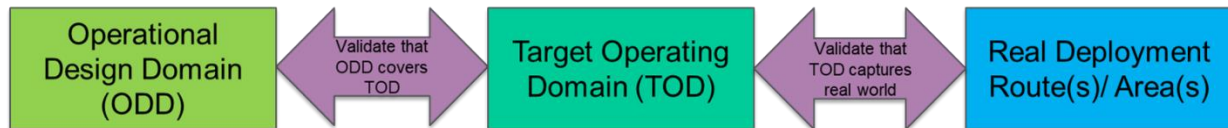


Figure 20: Validation between the ODD, the TOD, and the real deployment route(s) or area(s)

## 4.2.2 Recommendations

### 4.2.2.1 Proposed Requirements

#### Compatibility of the TOD with the Real World

The TOD shall be justified with evidence relating to the conditions present, or likely to be present, within the deployment route(s) or area(s). This shall include:

- Identification of all permanent features and their attributes within the specific deployment route(s) or area(s), including the road layouts and geometries and also the traffic infrastructure such as road markings or signs
- Identification of what non-permanent features are reasonably foreseeable within the specific deployment route(s) or area(s) (e.g. traffic cones, cyclists, rainfall) and what attributes they could reasonably be expected to have (e.g. range of possible positions, speeds, rainfall rates etc.),

This shall be decomposed to a sufficient level of detail to support all elements of both the system safety case and the deployment safety case. The regulator shall audit the completeness and accuracy of the information presented within the TOD by identifying a sample of features and attributes within the route(s) or area(s) and confirming that the TOD captures these features (regardless of whether it defines them as in or out of scope), and shall also audit the processes used by the manufacturer and/ or operator to develop the TOD.

However, it shall be the responsibility of the manufacturer and/ or operator to ensure that the TOD decomposes the features and attributes to an appropriate level of detail, given that excessive or insufficient detail could make it impractical to conduct or audit other elements of the safety case later on, such as the scenario-based testing programme. If the level of detail within the TOD proves insufficient for any downstream safety tasks or evidence, the regulator shall require that the TOD be updated to remedy the issue before any safety evidence that makes use of the TOD definition can be accepted.

The level of decomposition within the TOD – for example, whether mobility scooters are broken down into different sub-categories before identifying which are in or out of scope – shall be at least as detailed as any categorisations used within the system architecture. For example, if an occupancy grid used for a perception subsystem to pass information to a path planning subsystem has separate subcategories for ‘walker’ and ‘runner’ under a broader ‘pedestrian’ category, then the TOD shall capture this level of detail at an absolute minimum.

Beyond this level of detail, regulators may work with the manufacturer or operator to help ensure appropriate decomposition within the TOD, but there are no objective measures by which to determine appropriateness; the TOD definition could in theory be decomposed until the ‘pedestrian’ category is sufficiently broken down that each person within the world is individually identified, but in practice this is clearly a long way from feasible, and some appropriate point to cease the decomposition has to be determined.

As the definition of the TOD (and indeed the definitions for the ODD, MEL and behavioural competencies) do not provide any assurance of safety in their own right, and are merely intermediate steps to support the creation and review of the safety analysis, testing and mitigations to cover the scope of the initial definition, the key factor in determining the level of decomposition that is appropriate is the level of decomposition required for such safety evidence; for example, if the functional safety or SOTIF cases determine that some levels of rain are permissible for operation but others are not, then the TOD must also decompose to this level.

As such, any rejection resulting from inadequate decomposition shall be for the downstream evidence that is unable to be supported by the TOD rather than rejection of the TOD itself. In practice, the definitions for the TOD and ODD should be seen as live documents that are able to be updated accordingly as downstream analysis takes place. Therefore, rather than being rejected outright, the any flaws identified in the level of decomposition would be required to be addressed via updates to the definition.

### Compatibility of the ODD and the OD

A manufacturer may, optionally, choose to make the ODD identical to the TOD; this would be appropriate in situations where the system is engineered, and type approval sought, for a single specific deployment. Where this is the case, the regulator shall perform a check to confirm that the content of the ODD and the TOD are indeed identical with regards to what is included within and what is excluded from the two domains.

Where the ODD and the TOD are not identical, the regulator shall perform a check to confirm that the ODD includes in scope every feature that is included within the TOD scope, and that the ranges of all parameters identified in the TOD match or exceed the ranges of the corresponding parameters of the ODD. As such, the TOD may be a subset of the ODD, but the ODD may not be a subset of the TOD. This is to ensure that all permutations that are foreseeable within the TOD are compatible with the ODD such that any safety evidence that relies upon the ODD (for example, the functional safety case) will not be invalidated within the targeted deployment.

Similarly, whilst the TOD may decompose features to a finer level of detail than the ODD – for example, so that the TOD only includes certain radii of mini roundabouts where the ODD caters to all radii, the ODD may not decompose to a finer level of detail than the TOD, as this would make it impossible to confirm whether some permutations of a feature that could be experienced in the TOD would be incompatible with the ODD. For example, if the ODD decomposed roundabouts according to the number of lanes, with some permutations identified as in scope but some out, and the TOD merely identified an umbrella category of roundabouts as in scope, there would be no traceability as to whether the roundabouts encountered in the deployment would be ones that are in or out of scope of the ODD.

### 4.2.2.2 Supporting Information

The proposed requirements set out a key distinction; it is possible for the regulator to audit the information that has been captured and the processes used to capture it such that a TOD can be rejected where due diligence has not been shown, but it is not possible for the regulator to determine what is an appropriate level of decomposition – for example, whether ‘pedestrian’ needs to be broken down into different levels of subcategory according to height, clothing, speed of movement etc.

The report therefore sets out an absolute minimum level of decomposition for the TOD – that it must decompose at least as far as the system itself does – and the TOD may be directly rejected by the regulator if this is not met. However, it is anticipated that many of the work products within the safety case that rely upon the TOD (e.g. SOTIF analysis, scenario-based testing) may require a level of decomposition beyond this, and the required level of decomposition can only be determined when assessing these downstream work products themselves; there is no objective means to determine whether the TOD, in isolation, has decomposed the information to a sufficient depth.

As such, there are no absolute requirements for the level of decomposition within the TOD beyond the minimum of the level to which the system itself decomposes scene attributes, the TOD merely being an intermediate work product to support safety assurance within other elements of the safety case. However, it is hoped that an ongoing dialogue between the regulator and the creator of the TOD will help identify problems early. In particular, it is likely that the SOTIF analysis will require progressive

decomposition of the TOD as the iterative cycle identifies further changes to the scope and/ or system design in order to ensure safety.

Furthermore, scenario-based testing will require consideration of the possible features and attributes in significant detail; for example, testing would need to cover the range of permutations possible, so even if the perception subsystem within the ADS places all pedestrians into the same category, it is essential to test a wider range of permutations that sit within the pedestrian category, e.g. to confirm that persons whose walking motion or face is obscured by clothing can be detected. As the test programme will be directly derived from the TOD, with the aim being to provide coverage of the range of permutations that could be encountered within the parameter bounds specified in the TOD, it therefore follows that a TOD that provides appropriate decomposition to support scenario-based testing is an essential precursor to support safety assurance.

As was shown in Figure 20, it will also be necessary to ensure that the ODD is compatible with the TOD. The clause in the proposed requirements is intended to ensure that the ODD at a minimum covers the full extent of what is in scope for the TOD; the ODD may cover more, but is not permitted to cover less. This allows safety assessments performed against the ODD to be able to cover multiple deployments should a broad, generic ODD be selected, whilst ensuring that the TOD being assessed within the particular regulatory approval will not invalidate the ODD-related assumptions within the applicable safety evidence. Given that the TOD will also have been validated against the real world, these two steps in combination provide reasonable assurance that the CODs encountered by the vehicle in service will be an acceptable match to the ODD, although it must be recognised that there will always be residual risk of unforeseen, and possibly unforeseeable, events resulting in situations where the COD is outside the scope prescribed by the TOD and ODD.

## 4.2.3 Future Consideration

### 4.2.3.1 Areas for Future Work

The information currently available in the public domain relating to ODDs is based upon a combination of abstract theorising and of knowledge gained from trials (as opposed to full commercial deployments). Furthermore, the industry is only starting to address the need for the target deployment domain (TOD) to be considered as a separate concept to the design domain (ODD). As such, the specification of the ODD and TOD should be seen as an immature area of industry where practical experience and empirical evidence is lacking.

In order to address this, it is necessary for ODDs and TODs to be analysed in detail for actual AVs and deployments; repeating the thought-experiments or trial-based ODDs that make up the current state of the art will do little to uncover the 'unknown unknowns' and allow approaches to defining and validating ODDs and TODs to be tested in a realistic manner. As such, the emphasis for further work should be upon:

- Undertaking more advanced trials that rely on system safety, rather than relying upon the fallback of a safety driver to mitigate for an immature system as per typical trials to date (Zenzic, 2021). Only this will truly pressure test the detail and accuracy of an ODD and TOD such that processes and formats can be further developed.
- Ensuring that the ODD, the TOD, the methods used for their creation, and the validation techniques applied, are shared within the public domain as part of the requirements for advanced trials that are in receipt of government funding. An analysis of the effectiveness of the methods used, and of areas for improvement, should also be required.
- Updating the requirements within the GB Safety and Security LSAV scheme to reflect real-world, proven practice from advanced trials (i.e. trials that don't feature a safety driver) and from early commercial deployments, making use of data as it becomes available.

## 4.3 Behavioural Competencies

### 4.3.1 Background

#### 4.3.1.1 Definition of Problem Addressed

The need to define the ‘Behavioural Competency Definition’ of a system was examined in Section 4.1, which looked at the behaviours that the system will perform within its intended deployment route(s) or geofenced area(s). This section further examines the required behaviour of the system by looking at how ‘behavioural competencies’ should be captured, bearing in mind that different competencies will be at a different level of abstraction. This supports a scenario-based testing programme because traceability and coverage can then be analysed between the test scenarios undertaken and the behavioural competencies that need to be tested, with these behavioural competencies in turn being traceable to the system requirements.

#### 4.3.1.2 Current State of the Art

The aim of behavioural competencies is to describe a high-level set of fundamental driving behaviours that all road users must be capable of performing to operate safely in a defined environment. These should be independent of how the vehicle is operated or equipped, and therefore apply to all levels of automation as well as manual driving. Behavioural competencies can be used to support a consistent and well-defined basis for validation.

The test programme as a whole will need to sample from the range of things the system can do (i.e. behavioural competencies) within the range of operating environments it could be required to do them in (i.e. the TOD). As such, both should be seen as inputs that define the ‘problem space’ that the test cases need to provide coverage of, as illustrated in Figure 21.

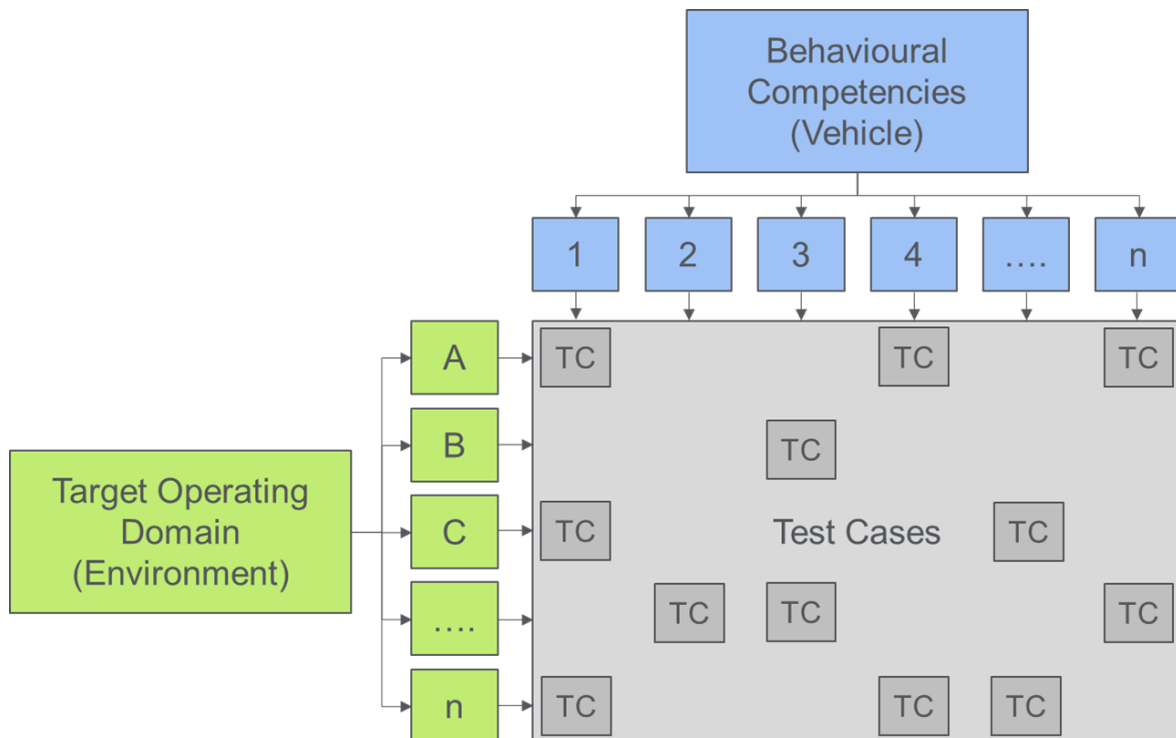


Figure 21: Illustration of how the test cases selected for the system need to sample not just the range of target operating domain permutations possible, but also the range of behaviours that the system could be required to perform within these environments

In addition to scenario-based testing needing to cover the range of possible behaviour and operating environment permutations (HumanDrive, 2020), Safety of the Intended Functionality (SOTIF) requires each functionality that the system must perform to be analysed to identify potentially hazardous scenarios and their triggering conditions (ISO/PAS 21448, 2019). Within this context, the behavioural competencies could be viewed as broadly analogous to the subset of the required functionality that relates directly to the DDT that the ADS must perform, albeit potentially at a different level of abstraction; development of a SOTIF case typically involves decomposition of functionality and scenarios until an appropriate level is reached.

Given that different levels of abstraction can be used to define the behaviours, an ontology that allows a hierarchical structure may be deemed appropriate. This is something that has not been examined for behavioural competencies, but there are many synergies with the task of ODD and TOD definition, which also involves different levels of abstraction, and therefore standards and guidance relating to ODD definition such as BSI PAS 1883 (2020) could be viewed as a plausible starting point for the ontology structure.

Within the USA, NHTSA (the National Highway Traffic Safety Administration) published a set of behavioural competencies (NHTSA, 2018), which are shown in Table 10. These in turn incorporate prior work by California PATH (2016). Waymo (2020c) have made use of some of the NHTSA competencies (Table 11), but have extended the list to incorporate novel ones of their own (Table 12). Some competencies may not be directly relevant to the UK, such as special rules around overtaking school buses.



Broad Grouping of Categories	Categories of Behavioural Competencies	Specific Behavioural Competencies
<b>Tactical Manoeuvres</b>	Parking (Note: ODD may include parking garages, surface lots, parallel parking)	<ul style="list-style-type: none"> <li>• Navigate a parking lot, locate spaces, make appropriate forward and reverse parking manoeuvres</li> </ul>
	Lane Maintenance & Car Following (Note: ODD may include high and low speed roads)	<ul style="list-style-type: none"> <li>• Car following, including stop and go, lead vehicle changing lanes, and responding to emergency braking</li> <li>• Speed maintenance, including detecting changes in speed limits and speed advisories</li> <li>• Lane centering</li> <li>• Detect and respond to encroaching vehicles</li> <li>• Enhancing conspicuity (e.g., headlights)</li> <li>• Detect and respond to vehicles turning at non-signalized junctions</li> </ul>
	Lane Change (Note: ODD may include high and low speed roads)	<ul style="list-style-type: none"> <li>• Lane switching, including overtaking or to achieve a minimal risk condition</li> <li>• Merge for high and low speed</li> <li>• Detect and respond to encroaching vehicles</li> <li>• Enhancing conspicuity (e.g., blinkers)</li> <li>• Detect and respond to vehicles turning at non-signalized junctions</li> <li>• Detect and respond to no passing zones</li> </ul>

	<p>Navigate Intersection (Note: ODD may include signalized and non-signalized junctions)</p>	<ul style="list-style-type: none"> <li>• Navigate on/off ramps</li> <li>• Navigate roundabouts</li> <li>• Navigate signalized intersection</li> <li>• Detect and respond to traffic control devices</li> <li>• Navigate crosswalk</li> <li>• U-Turn</li> <li>• Car following through intersections, including stop and go, lead vehicle changing lanes, and responding to emergency braking</li> <li>• Navigate rail crossings</li> <li>• Detect and respond to vehicle running red light or stop sign</li> <li>• Vehicles turning - same direction</li> <li>• LTAP/OD (left turn across path/ opposite direction) at signalized junction and non-signalized junction [NOTE: for UK roads, this would be reversed to consider a right turn across the path of oncoming traffic]</li> <li>• Navigate right turn at signalized and non-signalized junctions</li> </ul>
	<p>Navigate Temporary or Atypical Condition</p>	<ul style="list-style-type: none"> <li>• Detect and respond to work zone or temporary traffic patterns, including construction workers directing traffic</li> <li>• Detect and respond to relevant safety officials that are over-riding traffic control devices</li> <li>• Detect and respond to citizens directing traffic after an incident</li> <li>• N-point turn</li> </ul>
<p><b>OEDR (Object and Event Detection and Response) Capabilities</b></p>	<p>OEDR: Vehicles</p>	<ul style="list-style-type: none"> <li>• Detect and respond to encroaching, oncoming vehicles</li> <li>• Vehicle following</li> <li>• Detect and respond to relevant stopped vehicle, including in lane or on the side of the road</li> <li>• Detect and respond to lane changes, including unexpected cut-ins</li> <li>• Detect and respond to cut-outs, including unexpected reveals</li> <li>• Detect and respond to school buses</li> <li>• Detect and respond to emergency vehicles, including at intersections</li> <li>• Detect and respond to vehicle roadway entry</li> <li>• Detect and respond to relevant adjacent vehicles</li> <li>• Detect and respond to relevant vehicles when in forward and reverse</li> </ul>

	OEDR: Traffic Control Devices and Infrastructure	<ul style="list-style-type: none"> <li>• Follow driving laws</li> <li>• Detect and respond to speed limit changes or advisories</li> <li>• Detect and respond to relevant access restrictions, including one-way streets, no-turn locations, bicycle lanes, transit lanes, and pedestrian ways</li> <li>• Detect and respond to relevant traffic control devices, including signalized intersections, stop signs, yield signs, crosswalks, and lane markings (potentially including faded markings)</li> <li>• Detect and respond to infrastructure elements, including curves, roadway edges, and guard rails</li> </ul>
	OEDR: Vulnerable Road Users, Animals, Objects,	<ul style="list-style-type: none"> <li>• Detect and respond to relevant static obstacles in lane</li> <li>• Detect and respond to pedestrians, pedal cyclists, animals in lane or on side of road</li> </ul>
<b>Failure Modes</b>	ODD Boundary	<ul style="list-style-type: none"> <li>• Detect and respond to ODD boundary transition, including unanticipated weather or lighting conditions outside of vehicle's capability</li> </ul>
	Degraded Performance/ Health Monitoring, Including Achieving Minimal Risk Condition	<ul style="list-style-type: none"> <li>• Detect degraded performance and respond with appropriate fail-safe/fail-operational mechanisms, including detect and respond to conditions involving vehicle, system, or component-level failures or faults (e.g., power failure, sensing failure, sensing obstruction, computing failure, fault handling or response)</li> <li>• Detect and respond to vehicle control loss (e.g., reduced road friction)</li> <li>• Detect and respond to vehicle road departure</li> <li>• Detect and respond to vehicle being involved in incident with another vehicle, pedestrian, or animal</li> <li>• Non-collision safety situations, including vehicle doors ajar, fuel level, engine overheating</li> </ul>
	Failure Mitigation Strategy	<ul style="list-style-type: none"> <li>• Detect and respond to catastrophic event, for example flooding or debilitating cyber attack</li> </ul>

Table 10: NHTSA List of Behavioural Competencies. Source: NHTSA (2018)

Number	Description of Behavioural Competency
1	Detect and Respond to Speed Limit Changes and Speed Advisories
2	Perform High-Speed Merge (e.g., Freeway)
3	Perform Low-Speed Merge
4	Move Out of the Travel Lane and Park (e.g., to the Shoulder for Minimal Risk)
5	Detect and Respond to Encroaching Oncoming Vehicles
6	Detect Passing and No Passing Zones and Perform Passing Manoeuvres

7	Perform Car Following (Including Stop and Go)
8	Detect and Respond to Stopped Vehicles
9	Detect and Respond to Lane Changes
10	Detect and Respond to Static Obstacles in the Path of the Vehicle
11	Detect Traffic Signals and Stop/Yield Signs
12	Respond to Traffic Signals and Stop/Yield Signs
13	Navigate Intersections and Perform Turns
14	Navigate Roundabouts
15	Navigate a Parking Lot and Locate Spaces
16	Detect and Respond to Access Restrictions (One-Way, No Turn, Ramps, etc.)
17	Detect and Respond to Work Zones and People Directing Traffic in Unplanned or Planned Events
18	Make Appropriate Right-of-Way Decisions
19	Follow Local and State Driving Laws
20	Follow Police/First Responder Controlling Traffic (Overriding or Acting as Traffic Control Device)
21	Follow Construction Zone Workers Controlling Traffic Patterns (Slow/Stop Sign Holders)
22	Respond to Citizens Directing Traffic After a Crash
23	Detect and Respond to Temporary Traffic Control Devices
24	Detect and Respond to Emergency Vehicles
25	Yield for Law Enforcement, EMT, Fire, and Other Emergency Vehicles at Intersections, Junctions, and Other Traffic Controlled Situations
26	Yield to Pedestrians and Bicyclists at Intersections and Crosswalks
27	Provide Safe Distance From Vehicles, Pedestrians, Bicyclists on Side of the Road
28	Provide Safe Distance From Vehicles, Pedestrians, Bicyclists on Side of the Road

Table 11: Set of Behavioural Competencies used by Waymo that are derived from NHTSA ones.  
Source: Waymo (2020c)

Number	Description of Behavioural Competency
29	Moving to a Minimal Risk Condition When Exiting the Travel Lane is Not Possible
30	Perform Lane Changes
31	Detect and Respond to Lead Vehicle
32	Detect and Respond to a Merging Vehicle
33	Detect and Respond to Pedestrians in Road (Not Walking Through Intersection or Crosswalk)
34	Provide Safe Distance from Bicyclists Traveling on Road (With or Without Bike Lane)
35	Detect and Respond to Animals
36	Detect and Respond to Motorcyclists
37	Detect and Respond to School Buses
38	Navigate Around Unexpected Road Closures (e.g., Lane, Intersection, etc.)
39	Navigate Railroad Crossings

40	Make Appropriate Reversing Manoeuvres
41	Detect and Respond to Vehicle Control Loss (e.g., reduced road friction)
42	Detect and Respond to Conditions Involving Vehicle, System, or Component-Level Failures or Faults (e.g., power failure, sensing failure, sensing obstruction, computing failure, fault handling or response)
43	Detect and Respond to Unanticipated Weather or Lighting Conditions Outside of Vehicle's Capability (e.g., rainstorm)
44	Detect and Respond to Unanticipated Lighting Conditions (e.g. power outages)
45	Detect and Respond to Non-Collision Safety Situations (e.g. vehicle doors ajar)
46	Detect and Respond to Faded or Missing Roadway Markings or Signage
47	Detect and Respond to Vehicles Parking in the Roadway

*Table 12: Set of Behavioural Competencies used by Waymo that are additional to the NHTSA ones.  
Source: Waymo (2020c)*

The main limitations observed with the NHTSA behavioural competencies are that they are only given very broad definitions, and are therefore open to interpretation, but most importantly, there is a lot of potential overlap across the tactical and OEDR sections. 31 out of the 54 NHTSA competencies are defined as “Detect and Respond to...” different road features and users, but in many cases at least the detection task is common to other competencies (i.e. those that could be viewed as being at a different level of abstraction). This creates a lot of opportunities for inconsistent interpretation and duplication of effort through the subsequent requirements analysis. It should also be noted that there are significant differences between the two lists of behavioural competencies, suggesting a difficulty in highlighting a definitive list that is applicable to all systems and deployments.

It should be noted that both the NHTSA and Waymo competency lists do not just include the manoeuvres that the system would be required to undertake during ideal driving conditions, but also consider detection and appropriate response to adverse events such as subsystem failures and adverse weather conditions.

Note that behavioural competencies were also touched upon within Section 3.3 of this report (Safety Goals and Risk Framework), in particular Table 5 and Table 6. Table 5 used a similar list of NHTSA-derived behavioural competencies, at a relatively high level, as an input to support the analysis of what hazards could result if such behaviours are not performed adequately; these hazards are captured in Table 6, with the behaviours described being hazardous behaviours, which the system safety analysis and testing should provide assurance against, in contrast to the desired behaviours. The behavioural competencies in Section 3.3 have been kept at a high level of abstraction as the aim was to arrive at a list that would be appropriate to a wide range of vehicles, in order to support the setting of performance requirements for each specific vehicle type, but it is envisaged that the behavioural competency definition for a specific vehicle type would have to be more detailed and less abstract in order to provide a comprehensive understanding of what the vehicle is required to do, thereby supporting elements of the vehicle safety case such as the SOTIF and functional safety analyses.

### 4.3.1.3 Conclusions Drawn

Whilst there is a clear need to define the behaviours that the system is required to provide in service, there is no established consensus upon either the format in which they should be defined or upon an absolute list that is ubiquitous to all deployments. As such, the regulatory requirements should propose a flexible approach that allows the list used by regulators to expand over time in order to support analysis, but that also includes a process to ensure that behaviours that are not currently on the list can be captured. In the interim, the NHTSA list of competencies should be used as a baseline.

Regardless of the length of the list, it should only be seen as a prompt, and it should not be expected that all vehicles and deployments have to include all identified behavioural competencies; for example, if the vehicle is not intended to perform lane changes at any point, then behavioural competencies relating to lane changes would not be applicable. Therefore, whilst the regulator should seek to develop

a broad list covering all known competencies in order to support analysis, the actual list of behavioural competencies for a vehicle type should be bespoke to that vehicle type.

## 4.3.2 Recommendations

### 4.3.2.1 Proposed Requirements

The manufacturer shall document the Behavioural Competencies that the system is able to perform; these shall be captured in a 'Behavioural Competency Definition'. The behavioural competencies shall be considered within the system safety case, including within the consideration of what hazards may arise from the system's intended behaviours (or functionality, within the SOTIF parlance). The Behavioural Competency Definition shall also be used within the development of a test programme, to provide acceptable coverage of the range of behaviours in the range of possible TOD permutations. The regulator shall audit of the coverage provided by the test programme, to confirm that the full range of behaviours was observed, each within a representative range of surrounding TOD permutations.

The Behavioural Competency Definition should be made available to the Operator such that the behaviours can be considered within the deployment safety case (e.g. within the operational risk assessment, as described in Section 6.1).

The regulator shall audit the completeness of the Behavioural Competency Definition, and satisfy themselves that the following categories of behavioural competencies have been adequately covered:

1. Behaviours that every vehicle should be expected to provide. For example, it would be reasonable to require detection of and response to pedestrians for all systems operating upon public roads – even if pedestrians are prohibited, it is nonetheless a foreseeable misuse. Similarly, all systems must be able to react appropriately to faults detected within subsystems or components (e.g. by switching to a degraded mode or performing an MRM). Within this report, the safety goals defined in Section 3.3 provide a required outcome for the system behaviour, regardless of operating environment, and the behavioural competencies help ensure they are met through defining required behaviour that is particular to that vehicle type and operating environment.
2. Behaviours that regulators should expect the system and deployment safety cases to consider. For example, 'navigate roundabouts' or 'perform lane change' are both reasonable behaviours that the regulator shall expect to see consideration of within the safety case, but the safety case may conclude that they do not apply (e.g. if the route that the system is intended for does not contain any roundabouts or multi-lane sections). Where it is claimed that a behavioural competency is not applicable, justification shall be given.
3. Behaviours that are not foreseen by regulators or captured in a standard list, but nonetheless are necessary for safe operation within the deployment route(s) or area(s). These are difficult to capture, and could be viewed as 'unknown unknowns', but nonetheless could have significant safety implications. By definition, the regulator cannot audit against a list of competencies that are hitherto unknown, and therefore the audit shall instead focus on ensuring that the manufacturer and operator have applied due diligence within their processes such that it can be trusted that the Behavioural Competency Definition provided for the vehicle type is complete. This process shall include identification of new behavioural competencies required for operation within the TOD via analysis, via discovery during testing, and via in-service monitoring, and shall result in the safety case being updated where new behavioural competencies are identified. Where the manufacturer and Operator are separate entities, the process shall include a viable means for new Behavioural Competencies identified by the manufacturer or the operator to be passed on to the other body such that both safety cases are updated.

It is recommended that the regulator compiles and maintains a list of competencies that can be used as a checklist for auditing safety cases; this should also be made available to manufacturers and operators to act as a prompt and aid the completeness of the Behavioural Competency Definitions that they provide. All behaviours falling into category 3 in the above list shall, upon being identified, be logged within the system safety case and deployment safety case reports that are provided to the

regulator, in such a manner that they are readily apparent to the regulator. The regulator shall then use these to further enhance the list of competencies.

The safety management systems identified within both the system and deployment safety case reports shall include a process such that new additions to the list of competencies held by the regulator shall be periodically checked, any new behaviours reviewed, and the Behavioural Competency Definition for the system and deployment updated where applicable. The regulator could support this by providing a means to readily identify new additions to the list.

### 4.3.2.2 Supporting Information

It is anticipated that the list of behavioural competencies maintained by the regulator would initially be based upon those produced by NHTSA and Waymo, combined with the safety goals identified within this report (Section 3.3), and potentially further competencies identified via a review of the road rules identified by Work Package 2 of this project. The list would then progressively expand over time as more experience of AVs is gained by industry, thereby further enhancing safety by reducing the number of necessary behaviours that are missed within the safety analysis and testing.

There is significant overlap between behavioural competencies and the 'functionality' considered within SOTIF, and it is therefore likely that the behavioural competencies will be significantly refined, possibly involving decomposition to a lower level of abstraction, as the SOTIF analysis progresses. Behavioural competencies should be defined prior to the SOTIF process commencing, to ensure that there is a clear definition to support the analysis, but may be modified by the process of 'tailoring' the functionality within SOTIF such that unsafe functions are removed; as such, an initial list of competencies could be seen as an input to SOTIF, with a refined list being an output.

In order to support different levels of abstraction within behavioural competencies, it is recommended that a hierarchical structure is used such that multiple lower-level competencies can be nested under each higher-level one. This will reduce the level of duplication, thereby making traceability easier to manage. A structure such as that used for ODD definition within BSI PAS 1883 (2020) could be applied.

It is expected that a decomposed list of behavioural competencies will be of particular use when assessing the coverage that a test programme has provided. The definition of the TOD will allow the environmental aspects of scenarios to be assessed for coverage, but it is also essential that the test scenarios cover the range of behaviours expected of the vehicle. Decomposition will help make this coverage analysis more complete; for example, if the high-level behaviour is merely 'negotiate roundabout', coverage could be achieved through repetition of relatively similar scenarios, whereas if the competency is further decomposed to consider taking multiple different exits to the roundabout, requiring different turn signals to be given and potentially requiring different lanes to be used, it would be possible to observe if the test programme has been biased towards one permutation, and corrective action could be taken.

This is particularly important within the testing of AVs, as opposed to active safety systems for driver assistance, as the definition of a particular test case doesn't necessarily determine what behaviour the vehicle will provide, this being under the control of the vehicle rather than those performing the tests. For example, where one AV reacts to a cyclist by decelerating to follow them, another AV may react to the same scenario by maintaining speed but adjusting the trajectory to overtake them, both being potentially safe and reasonable behaviours in such a scenario. As such, retrospective analysis of the behaviours observed within the test programme will be vital to ensure appropriate coverage was achieved, and to identify any areas where further testing, with adapted test parameters, may be needed to stimulate a particular behaviour that is under-represented in the data.

The close link between behavioural competencies and the TOD should be noted; for example, if the route(s) available to the system only allow left turns to be taken at roundabouts, then scenarios involving other exits would not be relevant to the safety assurance. Therefore, the creation of the behavioural competencies definition should not just consider what it is desired for the vehicle to be able to do, but also what the nature of the possible deployment route(s) require the vehicle to be able to do if it is to perform safely.

## 4.3.3 Future Considerations

### 4.3.3.1 Areas for Future Work

This is an aspect of AV safety assurance that will need significant further work to build a more complete picture of what behavioural competencies a vehicle may be expected to perform. This should be led by the regulator, who should set up a process to allow new behavioural competencies to be captured and shared such that learning from past approvals and from in-service monitoring can be used to improve safety over time. This will require it to be permitted for any behavioural competencies submitted within the safety case to be captured and shared by the regulator.

Experimentation via practical application is also needed in order to develop more mature methods to define the behavioural competencies; whilst this has been done to a certain extent by Waymo, there is limited information available in the public domain, and it may be expected therefore that formats and processes will evolve as further practical experience is gained.

As such, the regulator should monitor approaches being used and look to identify any emerging state of the art that would work for a range of deployments and systems and would be scaleable such that it could support deployments with a far broader OD. It is proposed that practical experience of analysing a real deployment, including the full complexity of the real world that it must operate in, would be far more beneficial than an analysis of a hypothetical deployment or of a small-scale R&D trial that relies upon the safety driver to provide behaviours that are beyond the vehicle's capabilities; only a real deployment can fully pressure test the processes used and allow a scaleable methodology to emerge. However, as a short-term solution to gain a better understanding of how to define behaviours, it is recommended that government-funded R&D trials should be asked to define and share the behavioural competencies, as whilst this won't be sufficient to completely solve the problem for full commercial deployments, nonetheless it will give a better benchmark than that provided by the extremely limited data currently available.

### 4.3.3.2 Future Expansion

It is reasonable to expect that early implementations of AVs will be relatively limited in the route(s) available and the behaviours that they are expected to undertake, when compared to human-driven vehicles or to the longer-term vision for AVs. This makes it much simpler to capture the required behaviours. However, as the complexity of deployments increases, it may be expected that the behavioural competencies definition will also get more complex, this being further compounded as increasing experience of AVs results in an ever-expanding list of competencies maintained by the regulator.

Similarly, expanding the scope of the GB Safety and Security Scheme (e.g. to accommodate higher speeds) will result in an expansion to the behavioural competencies definition, thereby requiring solutions that are scalable.

As experience of more complex deployments is gained, it may become possible to provide more prescriptive requirements, but the current high-level requirements will provide sufficient flexibility such that they are expected to remain applicable to systems and deployments that feature broader scopes.



## 5 Assurance of System Safety

Safety is an emergent property of a system and is dependent on how a system behaves when it performs its required functionality within the ultimate operating environment. System Safety covers both product safety (the safety of the product, the LSAV in this case, in use without failure) and functional safety (safety of the product in case of failure). To achieve system safety a number of engineering principles, techniques and processes need to be applied during all phases of a systems lifecycle.

In this section, these different topics are reviewed and recommendations for the assurance scheme proposed, starting with existing safety guidance in the automotive industry as defined by international standards, but also covering areas that fall outside the scope of existing standards. Proposed requirements and supporting guidance are provided for all aspects of the system safety assurance.

### 5.1 Functional Safety

#### 5.1.1 Adherence to ISO 26262

Functional Safety (FS) is a well-established discipline in the automotive industry; an automotive series of standards for FS has been in existence for over a decade, and pioneering working in the UK goes back to the early 1990s. This series of standards is widely accepted in the automotive industry, being derived from the Functional Safety generic standard IEC 61508, and gives guidance on assessing risk associated with malfunctioning behaviour of E/E (electrical and/ or electronic) systems in vehicles. It contains a risk classification scheme that can be used to assess hazards and, based on the required risk reduction identified in the hazard analysis and risk assessment, gives guidance for system, software, and hardware development activities.

Having first been developed almost 20 years ago, some concepts around the traditional “driver-in-the loop” model and its consideration of controllability need to be tailored to make them more applicable to automated driving.

ISO 26262 (2018) provides requirements on how to achieve functional safety of a vehicle system (“item” in ISO 26262 terminology) through the implementation of a safety lifecycle during product development that provides an approach to risk management. It provides a particular risk model that has been adapted around a driver control model. Although not setting any quantitative targets for safety, there is an implied “accepted” level of risk that application of ISO 26262 gives. However, it should be noted that this risk is concerned with malfunctioning behaviour only and does not cover risk due to the general use of the product, the vehicle, within the road transport environment. Instead, the scope of ISO 26262 is limited to providing guidance on mitigating risk arising from malfunctioning behaviour of automotive E/E systems, caused by software or hardware faults.

Functional Safety will be an important part of assuring the safety of automated vehicles, but on its own, is not sufficient, as it only covers some causes of hazards associated with automated driving. Nevertheless, as an established reference of “best practise” for the development of EE systems in the automotive industry, it is proposed to include a requirement that malfunctioning behaviour of all EE systems on an LSAV needs to be shown to be adequately addressed as part of type approval, but without making compliance with ISO 26262 mandatory. The manufacturer will be required to produce evidence that malfunctioning behaviour has been addressed sufficiently and the existence of a Functional Safety management system will need to be demonstrated, with adherence to it during the development of the LSAV confirmed through audit evidence. It is also envisaged that the evidence from functional safety activities will form part of the safety case that the manufacturer submits. The evidence provided can be produced by following a process based on ISO 26262, or an equivalent approach if it can be shown to achieve the same objective.

In its 2018 update to Edition 2, the scope of the standard was extended from its initial applicability to passenger cars to also include trucks and buses. The main addition this provided to the standard was an approach to how different vehicle variances could be treated in the functional safety analysis; this approach might suit manufacturers who are pursuing both goods and passenger vehicles.

When reviewing the evidence, it is proposed that the regulator should focus their review on assessing:

- How a manufacturer has approached Functional Safety at vehicle-level to ensure that the interactions between all items contributing to the DDT are addressed, and ;
- That they has ensured that consistency and traceability across all the items is achieved, with a particular focus on the items that provide the ADS functionality.

Of additional importance is to ensure that the automated driving use case has been considered when performing the risk assessment and when defining appropriate safety concepts, and that safety concepts for cross-vehicle features or systems such as networks, power supply, diagnostics and software updates are appropriate. This should be reviewed in the context of any MRM capability that the manufacturer has declared.

It is not proposed that the requirements should seek to duplicate or mandate the application of any existing functional safety standard; instead, the process should allow flexibility for the manufacturer to justify a reasonable methodology by which they have achieved functional safety.

### 5.1.2 The Challenge of Qualifying AV Simulation Tools

If a software tool is used to support or automate all or part of a process, then there is a potential for the tool to introduce systematic faults into the design, or to fail to detect the presence of systematic faults within the design. Additionally, if a software tool is used, then its appropriateness for compliance with the process requirements of ISO 26262 are to be established.

In addition to addressing hazards due to malfunctioning behaviour of E/E systems, ISO 26262 also provides guidance on ensuring the tool is adequate for supporting the intended tasks within the development process and that malfunctions of software tools do not increase the potential for systematic faults.

This is achieved by ensuring that every software tool used in the development is subjected to a process that aims to identify its:

1. Intended use, including its inputs and outputs, procedures, environmental and functional constraints
2. The possibility and scope of a tool malfunction impacting safety
3. Any existing measures against software tool induced errors in the design (e.g. process steps or redundancy in tasks or even software tools)

Based on the potential impact of a tool malfunction and the existing measures against it, there are a number of methods required in ISO 26262 to ensure that software tools in the development process are appropriate to the level of risk they themselves may pose.

Possible measures may include (in order of decreasing rigour)

- Developing the software tool in accordance with a safety standard
- Validation of the software tool
- Evaluation of the tool development process
- Developing confidence in the tool from use

As well as ensuring that malfunctions of a tool do not introduce faults into an 'item', ISO/FDIS 21448 extends on the requirements on qualification of software tools in ISO 26262 by requiring that, as well as confidence in the use of software tool with respect to malfunctions introducing errors into the design, there is confidence in the capability of a tool in the context of its contribution to the safety of the intended functionality. Examples given within the FDIS include ensuring confidence of the simulations tools that represent real world parameters appropriately and also the accuracy of real-world data measurements).

The draft technical specification for automated cars produced by the European Commission (2022) contains a Part 4 titled "Principles for Credibility Assessment For Using Virtual Toolchain In ADS Validation", itself derived from UNECE's Annex III within VMAD (2022), which aims to ensure that manufacturers provide evidence that a virtual toolchain used for ADS validation is appropriate. It builds upon some of the methods proposed in ISO 26262 by requiring an assessment of the criticality analysis of the tool in accordance with ISO 26262, together with a definition of the scope/use of the tool (see

Figure 22). To ensure that the risk associated with the use of the tool chain ('tool residual risk') is appropriate, the proposed requirements include use of 3 out of the 4 methods proposed in ISO 26262 by manufacturers:

- Validation of the software tool (covered by 3.5)
- Evaluation of the tool development process (covered by 3.4)
- Developing confidence in the tool from use (inferred in 3.4.5.7.2)

The method "developing the software tool in accordance with a safety standard" is not taken forward, there being no established standards in the field to use as a benchmark.

It is advised that regulators continue to monitor guidance and draft standards produced by the European Commission, UNECE and ISO.

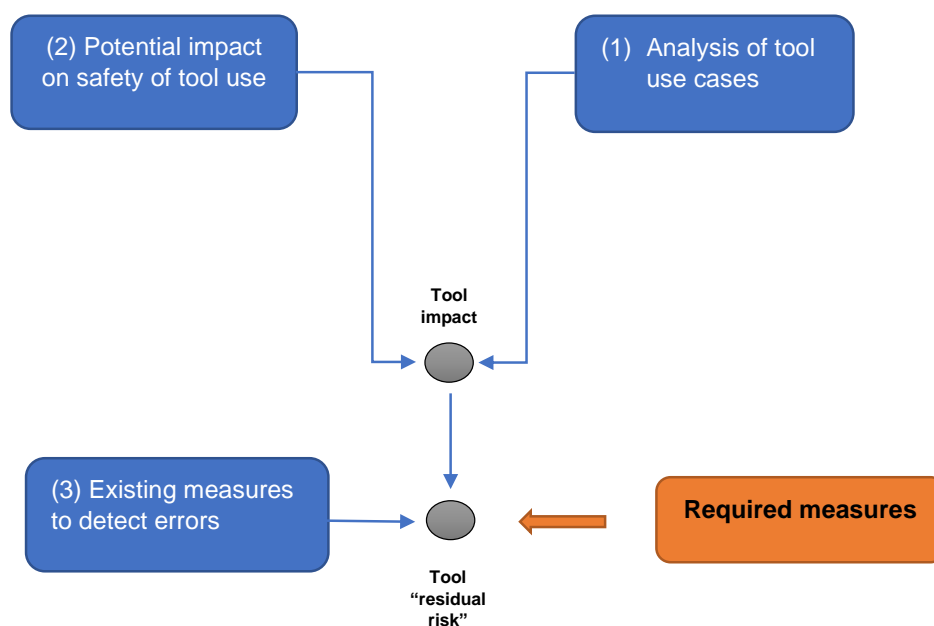


Figure 22: Software tool impact and residual risk

### 5.1.2.1 Scenario-Based Testing in Simulation

Scenario-based testing of the complete vehicle, including the ADS, within simulations of realistic scenarios, poses particular challenges that are not addressed by traditional toolchains. Whereas it is feasible to, for example, consider a C++ compiler to have been 'qualified' on the basis of assessment of the finite problems space presented by the C++ 17 standard, and then use that tool for a wide range of applications relating to the vehicle, this report proposes that such an approach would not be appropriate for scenario-based testing within vehicle-level simulations.

This is as a result of the sheer complexity of the scenarios, and the attributes that need to be modelled. In the same manner as how the complexity of the 'scenario space' created by the range of permutations possible within the behavioural competency and TOD definitions results in an intractable problem in terms of sampling the vehicle's performance, the exact same sample space needs to be considered when validating vehicle-level simulations. As such, even if it has been identified that the tool provides acceptable correlation against real world data within, say, a scenario that features rain and low lighting conditions, this will provide little assurance that simulations will be similarly representative within scenarios that feature snow in combination with bright sunlight.

When this principle is extrapolated to consider all parameters and ranges that could occur within the deployment, and that are therefore appropriate candidates for simulated scenarios, it will readily be

appreciated that this results in a vast scenario space to be assured. Furthermore, it is a scenario space that will continue to expand as more permutations are identified.

As a result, it would not be reasonable to consider the simulation tool to be 'qualified' such that it is representative in all scenarios. Instead, the aim should be to assess the correlation between similar or identical scenarios performed in simulation and the real world, in order to justify that the tool is representative within that area of the scenario space, allowing a level of confidence to be achieved when similar, but not identical, scenarios are simulated. For simulations in other areas of the scenario space to be similarly trusted, they will need to be similarly validated via assessing the correlation against equivalent scenarios in the real world.

Thus, sampling throughout the scenario space allows progressive confidence to be built up in the simulation such that simulations are able to be used to gain coverage in a wider range of permutations. The use of testing in simulation, the real deployment, and physical mock-ups such as upon a proving ground, is further considered within Section 5.9, which again considers the challenge of gaining coverage to validate the simulation as being a similar problem to gaining coverage to validate the performance of the vehicle itself.

This validation of the simulation software within realistic scenarios representing the scenario space should, in a similar manner to the testing of the AV itself, be augmented by verification testing to assess the individual components/ subsystems, and the complete system, against requirements (e.g. assessing the accuracy of vehicle dynamics, radar or camera lens distortion models in isolation, prior to integration).

## 5.2 Safety of the Intended Functionality (SOTIF)

Whereas product safety for traditional vehicles has evolved over time, and regulations for safety-relevant aspects of traditional vehicle systems exist (with later extensions for complex electronic systems having been added), the introduction of ADAS (advanced driver assistance systems) brought with it a realisation that vehicle level hazards could also be caused both by non-malfunctioning behaviour and as a result of misuse of an EE system (which by definition fall outside the scope of ISO 26262). This led to the development of a new standard called SOTIF (Safety of the Intended Functionality), which describes a process for arriving at a safe nominal functional through a safe specification for such systems, that can subsequently be verified and validated against set criteria. Setting the criteria is part of the process, safety thresholds and targets are not prescribed in the standard.

The guidance for performing SOTIF activities is currently published as a PAS (publicly available specification), with work towards a full standard almost complete, publication being expected in Q3 2022.

The emphasis in ISO/PAS 21448 (2019) is on identifying issues with the specification of functionality that could lead to potentially hazardous behaviour, either because it has been incorrectly specified, because specification content has been missed or because of limitations of the design resulting in it being insufficient to correctly implement the specified functionality. The scope covers both functionality that implements the driving functionality, and also functionality related to interacting with and monitoring of the driver, including potential misuse (through misunderstanding, laziness or mistakes, not though malicious action).

ISO/PAS 21448 makes use of the concept of scenarios and sets out a process that aims to identify as many 'unknown unsafe' scenarios as possible (i.e. to uncover previously unidentified unsafe scenario permutations), in order to be able to argue that the system is safe when introduced into the operational environment. The overall goal of ISO/PAS 21448 is to reduce the risk presented by an ADAS/ADS functionality to an acceptable level through ensuring that safe behaviour is specified for all required scenarios.

SOTIF is part of product safety but also not sufficient to ensure overall safety of an automated vehicle product on its own. The standard does provide guidance on how to address safety in use during development of an automated driving system, but also acknowledges that what might be considered mature "best practice" for higher automation driving systems is still not fully established.

The full standard will contain mainly updates to the PAS that add clarification in the terminology, more guidance on the proposed processes and activities, and more specific descriptions on the outcome and output of each activity. It also acknowledges that full assurance at the time of release or approval of an ADS system is not possible, and that safety assurance needs to continue after deployment via monitoring of the achievement of safety during operation. This requires processes in place to observe and evaluate the performance and react appropriately, when required.

SOTIF will also contribute to system safety and the assurance of automated driving safety. As a discipline and standard, it is not yet as well established as Functional safety and ISO 26262, but nonetheless regulators should require that hazards arising from non-malfunctioning behaviour of an automated driving system are shown to be adequately addressed as part of type approval. As per requirements for ISO 26262, compliance with ISO/PAS 21448 should not be mandatory, but manufacturers should be required to evidence via either ISO/PAS 21448 or some equivalent means that the intended functionality has been shown to be acceptably safe with regard to non-malfunctioning behaviour. This must be evidenced within the safety case report that the manufacturer submits.

Evidence from the SOTIF assessment will be a key contribution to the evidence that a manufacturer would submit when demonstrating compliance with the performance requirements of the Safety and Security Scheme. The safety goals defined in Section 3.3 aimed to capture universal safety requirements for the Road Traffic Ecosystem as a whole. In order to be applicable to a particular automated vehicle type, these are refined into more specific technical performance requirements for the ADS of an automated vehicle within Section 5.4. A manufacturer will be required to show that their ADS design and implementation meets these technical performance requirements, and a SOTIF process can be applied to support the claim that the intended functionality of the ADS is acceptably safe.

The key output of a SOTIF process as described in ISO/PAS 21448 is the “SOTIF release argumentation” and its documented evaluation. This contains the evidence that all required activities have been carried out and their objectives fulfilled and should form part of the vehicle safety case report that is submitted as part of the regulatory process. As part of the safety case review, the regulator should consider this argumentation and the evidence behind it.

An additional output of a SOTIF process is the setting off acceptance targets and deriving of validation targets for hazardous behaviours. These criteria and targets should contribute to the vehicle safety case by providing the argument and associated evidence that a manufacturer submits to demonstrate an acceptably safe ADS functionality that meets the requirements and substantiates the claims with quantitative metrics.

### 5.2.1 Nominal Functionality vs Malfunctions

The precursor to the guidance in ISO 26262 is that the nominal functionality of any item being analysed is free from unreasonable risk. It is the guidance in ISO/PAS 21448 that first supplemented the available best practice guidance with a process to achieve freedom from unreasonable risk for the intended functionality, while acknowledging that no system will 100% free from risk.

The model used in the ISO/FDIS 21448 in Appendix A describes the categories of contributing issues or factors to safety concern shown in Figure 23, and which standard they are covered by

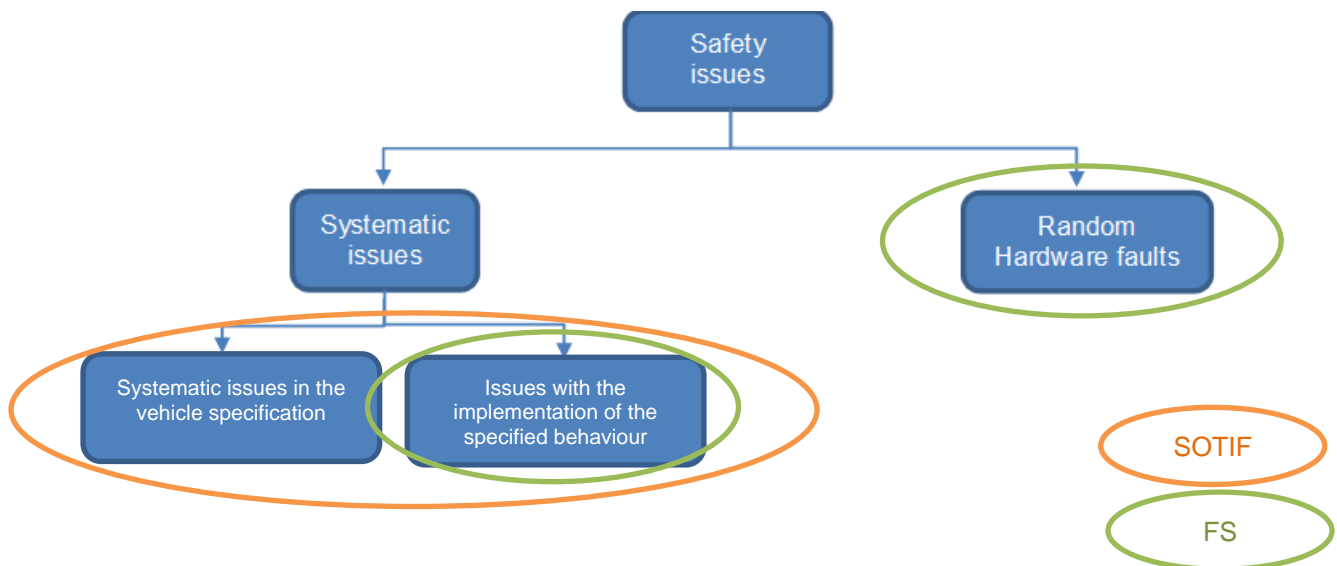


Figure 23 Safety issue causality classification scheme

The safe nominal function that is the goal of SOTIF activities should be specified appropriately as an outcome of the SOTIF process to ensure that the distinction between nominal functionality and fault condition can be determined. When performing functional failure analysis illustrations of malfunction interpretations are often used (see Figure 24).

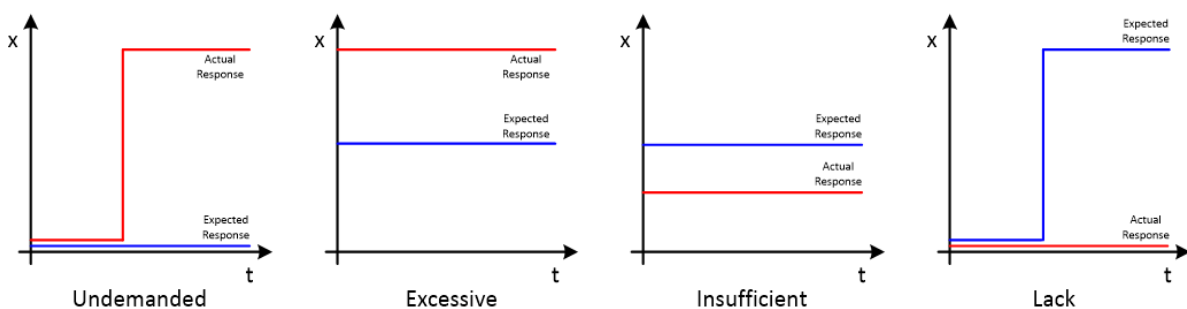
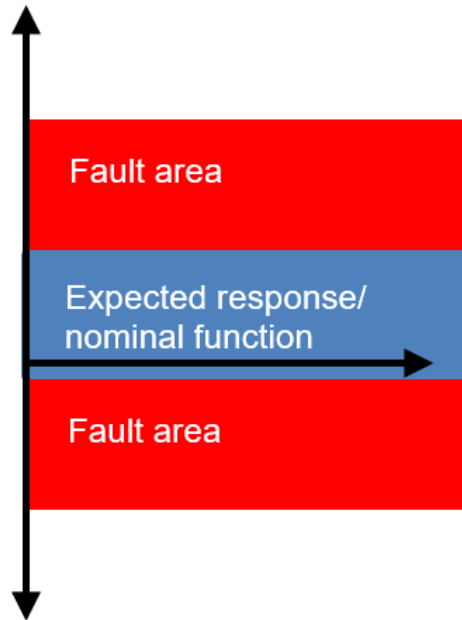


Figure 24: Guideword Illustrations

Rather than “binary” nominal or fault conditions it is more useful to consider a “nominal operating range” of a function that allows for tolerances of components due to factors like ageing, manufacturing and build tolerances, etc. to be accepted for normal operation, but a boundary has to be defined to distinguish the fault condition from the nominal range. Figure 25 shows this as an updated graph from the guideword illustrations shown in Figure 24



*Figure 25: Nominal function including tolerances*

Note: This figure is a simplification as the nominal area and fault area size will vary over time and depending on other conditions.

## 5.3 Cybersecurity and Software Updates

### 5.3.1 Cybersecurity

Automated vehicles will potentially rely on connectivity for critical inputs, and this reliance on external data necessitates that cybersecurity is an integral part of the safety assurance scheme as being able to operate securely is fundamental for safe operation. It is also recognized that security impact extends to additional potential loss event categories beyond safety; in particular, to privacy, financial and operational impacts. As such, requirements for measures to ensure protection against malicious actors need to form part of the approval scheme.

UNECE Regulation 155 addresses cybersecurity at vehicle level, and is applicable to M and N category vehicles, category O (with at least one electronic control unit) and categories L6 and L7 if equipped with level 3 or above automated driving functionality. It requires manufacturers to have implemented a Cybersecurity Management system that needs to be audited and approved before a type approval application can be made. The cybersecurity management system, which requires a cybersecurity lifecycle to be established and followed, assesses and mitigates risk during the design phase and continues monitoring during the operational lifetime of the vehicle. The part of the cybersecurity management process that allocates and agrees roles and responsibility for all cybersecurity activities throughout the cybersecurity lifecycle should also consider the operational phase and the additional organisation(s) that will be involved at this stage and ensure that their responsibilities are considered, allocated and agreed.

It is proposed that the GB scheme carries across the requirements of the UN Regulation, which require at a very high level the following vehicle-level approach:

1. Collect information about cyber-relevant systems
2. Assess cybersecurity risk
3. Implement appropriate measures and verify their effectiveness
4. Detect and Respond to attacks
5. Log and monitor data

This could be expressed in a high-level requirement for the LSAV thus:

*The LSAV (low speed automated vehicle) shall be designed to meet best practices for cybersecurity. The manufacturer shall document how the design ensures safe behaviour in the presence of cyber threats and demonstrate that the effectiveness of any security measures has been adequately tested.*

With the particular use case of these vehicles having no driver available to take over control, it is recommended to ensure that the operating organisation plays a part in mitigating cyber threats. Security for an automated vehicle is the responsibility of all organisations involved in the lifecycle, extending past the point of type approval. Some of the principles included in UN Regulation 155 should be extended to apply to the operating organisation, particularly the requirement for a Cybersecurity Management system. Additionally, we propose to stipulate that security-relevant information is included in the information that the manufacturer provides to the operator, in addition to the vehicle safety case report provided to the regulator. This information pack (the safety and security manual) should specify security requirements for off-board systems and any processes an operator may be required to set up and follow.

The focus at the type approval stage will be the assessment that the design of the vehicle includes adequate consideration of the cybersecurity risk. This risk might depend upon the properties of the ODD and TOD of the vehicle. Another consideration for this type of vehicle, which is likely to be part of a fleet of identical vehicles within a limited area, is the potentially increased attack impact. A number of standards exist that provide guidance on best practice. Akin to the approach for Functional Safety and SOTIF, the application of a particular standard is not proposed to be mandated, but manufacturers may use evidence from a cybersecurity process that is informed by standards towards the submitted evidence at type approval stage. Standards that may be relevant are listed below

- ISO/SAE 21434 Road vehicles – Cybersecurity engineering
- ISO/PAS 5112 Road vehicles – Guidelines for auditing cybersecurity
- ISO 24089 Road vehicles – Software update engineering



- BSI PAS 1885: The fundamental principles of automotive Cyber-security
- BSI PAS 11281 Connected automotive ecosystems – Impact of security on safety – Code of practice

Other publicly available information regarding security for V2X communications includes the European C-ITS standards developed by ETSI (2010, 2012, 2017a).

In-service monitoring plays a major part of the mitigation of cybersecurity threats, and this is already addressed in the regulation and standards listed above. The in-service monitoring required in UN R155 and assessed for type approval looks for evidence of appropriate processes being in place at the time of approval. The monitoring of the actual effectiveness and the execution of those processes should be aligned with activities set out as per the recommendation produced by Work Package 5 of this project.

When reviewing the evidence as a regulator at the type approval stage, it is also recommended that the interaction between security and safety measures is considered, to ensure that where safety measures depend on cybersecurity-relevant systems (e.g., interaction with a remote assistant), these are adequately protected and equally that safety mechanisms do not present potential attack paths. The potential interactions with external systems during operation and potential impacts of a cyberattack need to be considered when performing the vehicle level security analysis, and any assumptions on external interfaces and external security measures documented and reviewed during type approval and confirmed at the deployment approval. If, at the point of type approval, an actual physical deployment domain is known, then the full assessment of the external interfaces and security measures may be possible in one step.

The approach of not specifying particular standards directly, but instead requiring good practice to be shown in applying state of the art documents, enables manufacturers to take guidance from other standards that are currently in preparation, particularly ISO/AWI TS 5083, covering “Safety and Cybersecurity for Automated Driving Systems” (successor to ISO/TR 4804), ISO/AWI PAS 8926 Qualification of pre-existing software products for safety-related applications and ISO/IEC AWI TR 5469 Artificial intelligence – Functional Safety and AI systems. This is particularly relevant as state of the art in the field of automated driving is continuing to evolve.

### 5.3.2 Software Updates

In addition to UNECE Regulation 155, a new regulation specifically addressing software updates (UNECE Regulation 156) was introduced, which is applicable to vehicles within categories M, N, O, R, S and T that permit software updates of their electronic systems. This regulation was introduced to ensure updates that are Type Approval relevant have been notified to and approved by a regulator.

As there are different types of changes to software, ranging from fixes for existing software functionality, patches for cyberbreaches and new functionality added throughout the software lifecycle roadmap, UNECE WP.29 differentiated the need for amended Type Approvals, and hence requires the vehicle manufacturer to notify the relevant Type Approval Authority in case a software update-initiated modification affects the technical performance. On review of the submission, the regulator may request further tests, grant an extension, or consider that the existing approval continues to apply.

In a similar manner to Regulation 155, Regulation 156 requires manufacturers to have implemented a Software Update Management system, which needs to be audited and approved before a type approval application can be made. The regulation also specifies requirements for the vehicle itself regarding the implementation of software updates. These requirements apply regardless of whether the software is updated by wireless (over-the-air) or wired methods. Considerations include security of the update process and protection of the authenticity and integrity of the update, whether the vehicle needs to be stationary for the update to be applied, whether the user of the vehicle needs to be informed or consent to the update, and how to handle failed software updates, which could potentially result in incorrect functioning of an ECU. A new international standard, ISO 24089, is currently under development, which will provide industry- agreed guidance on implementing software updates in line with the requirements of Regulation 156.

## 5.4 Performance requirements

### 5.4.1 Background

#### 5.4.1.1 Summary of Literature Review for Safety Goal and Performance Requirement Development

A literature review was carried out to support the development of proposals for the Safety Goals and Performance Requirements. A number of existing regulations, and ADS-related regulatory drafts and working group outputs, have been reviewed, and their concepts used to influence proposals. A summary of the reviewed documents is included in the Appendices 3 and 4.

The aim of the evaluation for this report was to ensure that the proposed requirements for a GB LSAV Safety and Security Scheme do not diverge grossly from international developments, such that manufacturers do not face additional effort that might deter them from putting their product into operation in Great Britain.

#### 5.4.1.2 Current State of the Art

A key area of focus within the literature review was the current regulatory projects at EU and UNECE level. These contain requirements formulated as objectives that set the goal of safe behaviour in general. These are further elaborated by more detailed requirements describing particular aspects of safe behaviour in relation to the environment, and also separate requirements to handle critical scenarios and fault conditions, showing that the initial requirement is aimed at covering nominal, non-fault behaviours.

Each scheme contains some requirements that are formulated in detail while others require interpretation and evidence to be put forward to demonstrate the achievement. The EU proposal is more detailed in content and expands its requirements through the use of examples.

The draft proposals at UNECE and EU level are reviewed and evaluated in more detail within the remainder of this subsection.

#### Draft UNECE regulation for Automated Driving

Initially, the requirements from the 16<sup>th</sup> session were reviewed with a comparison to the later material available from the 21<sup>st</sup> session.

The material from the 16<sup>th</sup> session proposed a list of the following safety goals.

- ***The ADS should drive safely***

This safety goal is expanded by setting out 8 additional requirements that call for the achievement of safe driving through the performance of the DDT by the ADS, and by detection of the ODD and its boundary, relevant objects and events. Further requirements, giving more interpretation of safe driving, set out that safe interaction with other road users is to be achieved, as well as adherence to traffic rules. Additional requirements set out an expectation that, in order to maintain safety, vehicle behaviour can be adapted to prevailing safety risks and traffic conditions. A separate requirement states that the flow of traffic is to be maintained.

**Evaluation:**

The formulation of the requirement does not specify that the assessment of the performance of the DDT is only possible in the context of the ODD, and it is proposed for the GB scheme to acknowledge this point in the formulation of such a requirement.

The requirements to facilitate conflicts that might arise with traffic rules and through permission to adapt the behaviour (e.g., if safe distances cannot be maintained in order to avoid a potential collision) are considered important, and will be put forward for inclusion in the GB scheme.

- ***The ADS should interact safely with the user***

This safety goal also has 8 supporting requirements addressing activation of the ADS, communication of its status and a number of requirements addressing the interaction with an on-board driver. In case of no driver being available to take control, there is a requirement to facilitate communication between vehicle occupants and a remote assistant.

**Evaluation:**

Requirements for interaction with on-board drivers can be disregarded for the initial phase of the GB scheme, but in case of there not being a person in-charge on board, it is proposed to include the requirement to provide a communication means for passengers. As the activation status of the ADS is critical for liability reasons, this requirement will also be proposed for the GB scheme.

- ***The ADS should manage safety-critical driving situations***

This safety goal has 5 supporting requirements that address functionality that is to be put in place to safely manage failures of the ADS or other vehicle systems. Due to this legislation also covering vehicles that have the potential to hand control over to a driver, which is not a consideration for the GB LSAV scheme, such elements are not duplicated here. Further requirements cover that the ADS performs a Minimum Risk Manoeuvre (MRM) to achieve a Minimum Risk Condition (MRC).

A separate requirement mandates that in case of an accident, the ADS is to stop the vehicle.

**Evaluation:**

Due to the scope of the LSAV scheme being limited to fully-driverless vehicles, it is not proposed to include requirements addressing interaction with on-board users, but in order to ensure that fault conditions are safely managed we propose to include requirements for the ADS to perform MRMs and MRCs. Rather than specifying how MRMs are to be executed and which MRCs are accepted the requirements to be put forward are to be formulated as objectives, with the manufacturer required to provide a description of their solution, appropriate to the context of their ODD.

The requirement to stop the vehicle in case of an accident has been identified as necessary in the WP1 work.

- ***The ADS should safely manage failure modes***

The requirements supporting this safety goal address system malfunctions and require that these are detected, handled, and responded to by the ADS, as well as communicated where necessary. An additional requirement sets out that the ADS should be protected from unauthorised access.

**Evaluation:**

Requirements to manage faults and malfunctions of both ADS and non-ADS systems will be required, with the onus being on the manufacturer to show how they are being addressed sufficiently. The requirement to protect from unauthorised access in the absence of a person in charge on board is considered important.

- ***The ADS should maintain a safe operational state***

The requirements supporting this safety goal address provisions at the design stage for the ADS to support the operational phase by signalling any maintenance requirements, accessibility for maintenance and repair, and continued support for the ADS during the lifetime of the vehicle.

**Evaluation:**

Continued safe performance during operation is important, and monitoring system status and health will be an important part which should be added as a requirement. Whether this is performed with on-board self-monitoring functionality or through operational processes (e.g., regular checks in-service) might depend on the implementation, and left be for the manufacturer to decide and provide evidence of its sufficiency.

In the summary from the 21<sup>st</sup> session of the UNECE working party, the proposed list of high-level safety goals had been reduced to the following

- **The ADS should drive safely**

Although the safety goal has not changed, the supporting requirements have been reduced from a previous draft to only specifying 5 additional objectives.

The requirements to monitor the ODD conditions and to comply with traffic rules have been qualified to rules relevant to the ODD only – which the manufacturer is required to specify in a declaration. This is in line with the conclusion from the previous review, and already included in the GB proposed requirements. The requirements on the ADS to adapt its behaviour to safety risks and surrounding traffic conditions have been removed.

**Evaluation:**

Despite the removal of the requirement for the ADS to adapt its behaviour to prevailing conditions, it is still proposed that the GB scheme should require a manufacturer to provide evidence of their strategies on how the ADS in their vehicle adapts to varying risk conditions that might present themselves within the ODD.

- **ADS interactions with ADS vehicle users**

This section has been removed in this version with only a placeholder for general Human Factor requirements in place.

- **ADS management of safety-critical situations**

The requirements in this section have also been reduced but the requirement for fallback mechanisms including signalling thereof, to be in place is maintained, together with the requirement to stop the vehicle in case of accidents.

**Evaluation:**

The requirements already proposed to be included in the GB scheme are still included in the later UN draft regulation and hence still considered necessary.

- **ADS management of system failures**

The requirements in this section have only been modified slightly to address that the safe execution of the DDT and relevant requirements for fault detection must be in the context of the ODD, and hence have been reworded to address this.

**Evaluation:**

The proposal based on the review of the earlier draft had already proposed the modification contained in the later version.

As an output of the review of ongoing regulatory work at UNECE level, a number of requirements are proposed for inclusion in the GB scheme. Table 13 summarises how these requirements map to those proposed for the GB scheme (see Section 5.12).

The ADS should adapt its behaviour in line with safety risks.	2
The ADS should adapt its behaviour to the surrounding traffic conditions.	2
Activation of an ADS feature should only be possible when the conditions of its ODD have been met	8
Pursuant to a traffic accident, the ADS should stop the vehicle.	5
Requirement for one or more MRCs implementation – this should include appropriate signalling and indication. Obligation on manufacturer to declare MRC and MRM	9, 10, 14, 23

Requirements to manage system malfunctions and faults	22 & 23
The ADS should be protected from unauthorized access.	24
<i>Other requirements facilitating the monitoring of system health and status and indicating maintenance needs.</i>	22 & Ch.0

Table 13: UNECE Requirements identified for consideration in GB scheme.

## Draft regulation at EU level

The latest draft regulation (V8 – published Dec 2021) at EU level was also reviewed for comparison and additional content.

- **Dynamic Driving Task (DDT) under nominal traffic scenarios**

This section of requirements is expanded by setting out supporting requirements that specify the expected capabilities of the ADS concerning the DDT, including some specific requirements on safe speeds, appropriate distances from other road users and the necessary object and event detection capabilities, for which examples are presented. It is stated that the DDT capability and ODD are directly linked. The behavioural competence for driving in reverse is also called out specifically as a requirement to comply with traffic rules of the country of operation, as are signalling requirements to ensure safe interaction with other road users.

Also included are requirements for consideration of traffic conditions that may require adaptation of behaviour. The responsibilities of the ADS are also extended to the activation of other vehicle systems like door opening and wipers.

A specific requirement is given that limits horizontal acceleration to a specific value in case of standing passengers, but with a qualifier that it might be exceeded, without given conditions where this might be possible.

**Evaluation:**

This approach aligns well with what has been developed within WP1 for the requirements of the DDT. A minimum expected ODD specification will be proposed and also the minimum functionality expected for nominal scenarios.

Instead of specifying a maximum permissible acceleration (or indeed deceleration rate), it is proposed to require the manufacturer to specify and justify a value appropriate to their vehicle and ODD.

- **Dynamic Driving Task (DDT) under critical traffic scenarios**

In addition to the requirements on the DDT functionality in nominal conditions, there are requirements on the performance in critical traffic scenarios, requiring the DDT for be capable of managing reasonably foreseeable emergency situations. Collision mitigation functionality is to be implemented to minimise risks to vehicle occupants and other road users, with a requirement to avoid a collision if this is possible without causing another one. Emergency manoeuvres are also required to be signalled appropriately. This section also contains a requirement to stop the vehicle in case of an accident, with reactivation of the ADS only after verification of its operational state.

**Evaluation:**

Similar requirements have already been derived separately as part of the work in WP1, and requirements for mandatory collision avoidance functionality will be proposed similar to those in the EU scheme.

- **Dynamic Driving Task (DDT) at system boundaries**

The EU proposal contains high-level requirements addressing functionality related to ODD boundaries, requiring the ADS to be able to detect and predict, where possible, the ODD conditions and boundaries, ensuring they are met before activation is possible and also requiring safe behaviour to reach an MRC in case they are no longer fulfilled. Manufacturers are required to

establish measurable limits of ODD conditions, and a number of ODD conditions whose detection is mandatory are specified.

**Evaluation:**

Similar requirements have already been derived separately as part of the work in WP1, and requirements for ODD monitoring will be proposed similar to those in the EU scheme.

- ***Dynamic Driving Task (DDT) under failure scenarios***

The requirements in this section address functionality that the ADS needs to implement with regards to monitoring for faults, failures and malfunctions of vehicle systems, and being able to react to them appropriately. Failures and faults are also required to be indicated to relevant persons (e.g., vehicle occupants, operators, or other road users). If the fault prevents the safe execution of the DDT, an MRC is to be achieved, while minor faults may result in continued, potentially restricted operation of the vehicle. A particular requirement addresses steering and braking faults.

**Evaluation:**

Similar requirements have already been derived separately as part of the work in WP1, and requirements for self-monitoring and fault monitoring will be proposed that are similar to those in the EU scheme, covering both ADS and non-ADS systems.

- ***Minimum risk manoeuvre***

The EU proposal contains specific requirements on how an MRM is to be executed by specifying deceleration rates and requirements for signalling, and also the MRC that is to be achieved (vehicle at standstill). There is also a requirement for a positive confirmation before a vehicle can proceed after an MRM.

**Evaluation:**

It is acknowledged that there is a need for the ADS to contain functionality that is activated in case of the automated vehicle's system boundaries being exceeded. The very specific requirements in the EU scheme may restrict manufacturers from putting forward different solutions in certain cases.

The proposal for the GB scheme is suggested to differ by setting high-level requirement for MRMs and MRCs to be part of the system design with the manufacturer providing evidence of the suitability and safety of their chosen MRCs in the context of their vehicle and ODD for assessment. It is also recognised in the GB scheme that there may be multiple ODDs (and, indeed, TODs), with exit from one potentially being addressed by entry to another (e.g. by using degraded functionality to compensate for less favourable environmental conditions) rather than an MRM to reach an MRC.

- ***Human machine interface for vehicles transporting vehicle occupants***

The requirements for HMI in the draft proposal relate to functionality that is needed to address safety considerations concerning the interactions between the vehicle and vehicle occupants, and between the vehicle occupants and "a person in charge". This includes requirements to provide all necessary safety information to vehicle occupants, and means for vehicle occupants to stop the vehicle and communicate with the operator. The draft EU regulation also calls for specific functionality to be present, including a camera monitor system that is relayed to the remote supervisor to be able to monitor the situation on-board, and a remotely operational service door.

**Evaluation:**

Similar requirements have been derived as part of the work in WP1 and requirements for HMI provisions similar to those in the EU scheme will be proposed, covering provision for information, necessary communication means and functionality that allows vehicle occupants to request a vehicle stop to be able to leave the vehicle in an on-board emergency.

- ***Functional and operational safety during the ADS lifecycle***

The requirements in this section require the manufacturer to show due diligence in addressing functional and operational safety through evidence of adequate processes during development. It also extends this requirement to demand that this is shown to be met during the lifetime of the vehicle. There is a requirement setting a minimum overall safety target expressed in fatalities per hour to be achieved in operation, which is required to be demonstrated at the point of type approval.

**Evaluation:**

A requirement to address functional safety is also part of other type approval regulations that involve complex EE systems, and will be part of the proposal for the GB scheme.

The requirement for the overall target to be demonstrated at type approval is considered impractical, as outlined in the section on acceptance criteria (Section 3.4), and hence is not required within the GB scheme.

- **Cyber security and software management systems**

For this topic, the EU proposal calls on the existing UN Regulation No, 155 to ensure that there is adequate protection against unauthorised access, including continued support during the whole lifetime of the vehicle. This is aligned to GB scheme proposals.

- **Specific requirements regarding data recorder for ADS**

Being able to capture safety performance data in use is an important aspect of ensuring that the assessment at type approval remains valid in operation. The detailed content of the recording capability of an ADS is covered by WP5 activities and not reviewed further here.

- **Manual driving for emergency cases or for the purpose of maintenance or similar cases**

The EU scheme contains a number of requirements that address the use case of manual driving, to ensure that any provisions made are in line with current safety regulations and are appropriate to ensure safe manual driving. A speed limitation in remote driving combined with a requirement for line-of-sight control is put forward.

**Evaluation:**

The requirements allow for different solutions from a manufacturer, but with clear constraints to maintain safety. It is proposed to carry over a similar requirement that allows for unique solutions to be developed, but with clear boundaries.

- **Operation manual**

This section requires the manufacturer of an automated vehicle to ensure that any organisation involved in the operation of such a vehicle has access to all information required to ensure safety during operation. This covers limitations on its intended use, instructions for operation and maintenance and guidance, and training needs for any persons in contact with the vehicle during operation (such as owners, vehicle occupants, maintenance staff, remote supervisors, and the general public where necessary).

**Evaluation:**

Communication of safety-related information between organisations involved in different phases of the lifecycle of an automated vehicle is considered to be an important aspect of ensuring safety. Therefore, requirements that ensure a sufficient exchange of information are proposed to be included in the GB scheme.

- **Provisions for periodic roadworthiness tests**

This requirement addresses the need to ensure that the safe condition of the AV can be checked regularly in service, as per the current practise through mandatory roadworthiness ("MOT") tests. To accommodate AVs, the requirement asks for appropriate provisions to be made in the AV design.

**Evaluation:**

The proposal put forward in the EU work is a reasonable and pragmatic approach if existing roadworthiness schemes are expected to be used for AVs. If the specific application of low-speed automated vehicles as fully-driverless vehicles that this scheme is aimed at is supported by an organisation that manages all aspects of operation, there might be different ways of monitoring roadworthiness, but some provision that requires that the low-speed automated vehicle is safe in operation should be made in the overall scheme.

An earlier version (V4.1), which had previously been reviewed, covered the same high level safety goals overall, but combined requirements on handling of ODD boundary and failure scenarios, which have

been separated in the later version with additional requirements added to refine the objectives. It also proposed a more exhaustive description of the DDT functionality required, e.g., for overtaking or lane change situations, which have been removed and instead are set out as objectives to be declared and demonstrated by the manufacturer as part of the approval application. In its previous version V4.1 the overall safety target was also different.

A number of requirements, derived from a review of ongoing regulatory work at the EU-level, are highlighted for inclusion in the GB scheme in Table 14 (including requirement IDs).

Requirements for safe performance of DDT. Obligation on manufacturer to declare DDT performance and behavioural competences on the basis of the targeted ODD.	1, 2, 4
Requirements on ODD detection capabilities	3
HMI requirements addressing: <ul style="list-style-type: none"> <li>- Information for vehicle occupants</li> <li>- Communication facilities</li> <li>- Remote monitoring</li> <li>- "Stop vehicle" mechanisms for evacuation in case of on-board emergency</li> </ul>	11, 14, 15-18, 19
Requirements to specify MRC(s) appropriate for the ODD, how the LSAV achieves the MRC(s) and possible recovery strategies	9, 23
Requirement for a "Safe operating manual" that the manufacturer is expected to provide to support the organisation responsible for operation in maintaining safety.	
Requirements that ensure and support continued roadworthiness during operation.	

Table 14: EU Requirements identified for consideration in GB scheme.

### 5.4.1.3 Conclusions Drawn

From the review of the updated versions of both draft proposals at UNECE level and EU level, it can be seen that each regulation is moving towards specifying more objective-based requirements rather than prescriptive requirements on behaviour for different situations. This aligns with our proposal for the GB scheme in general.

### 5.4.2 Proposed Risk Framework

The safety goals provided in Section 3.3 set out the risk framework for what safe behaviour of the LSAV is expected.

Safe behaviour can only be determined in the context of the ODD, and eventually the target operating domain (TOD), the actual physical environment that the LSAV is expected to operate in. Rather than prescribing required functionality for each possible ODD and TOD, we propose to set out in the requirements that the manufacturer must describe how their implementation achieves the safe behaviour objectives set out in the safety goals, appropriate for their defined target operating domain. This evidence includes any prioritisation of safety goals, handling of exemptions and restrictions that need to be fulfilled during operation.

As the safe behaviour is defined in the context of the TOD, there need to be additional requirements to ensure that the LSAV is able to monitor TOD conditions appropriately and to manage when conditions are not met, as well as monitoring the status and health of all systems involved in the driving task. This



will be reliant upon assurance that the ODD is compatible with the TOD; see Section 4.2 for more information and requirements on this.

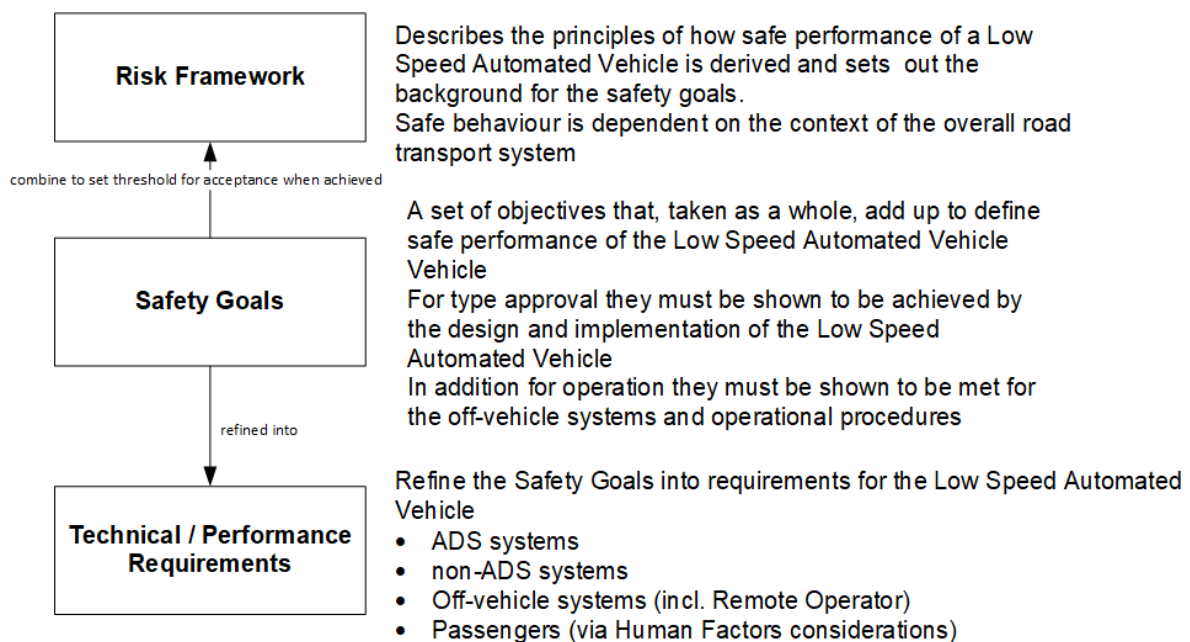



Figure 26 Requirements Structure

The safe behaviour must set the framework for how the LSAV is expected to interact. Figure 26 summarises how the risk framework, safety goals and technical performance requirements are related. Figure 27 shows the LSAV system in context of the overall road transport ecosystem, and also includes the components of the ‘super system’ (SS) – as described in the CertiCAV Assurance Framework report (Connected Places Catapult, 2021) in more detail as the ‘HASS’ (highly-automated supersystem) concept.

To achieve safe operation, the low-speed automated vehicle super system (LSAVSS) is dependent on its interaction with and response to:

- Other road users
- Road layout/infrastructure
- Passengers, with all of these representing potential collision risks. [represented by  ]

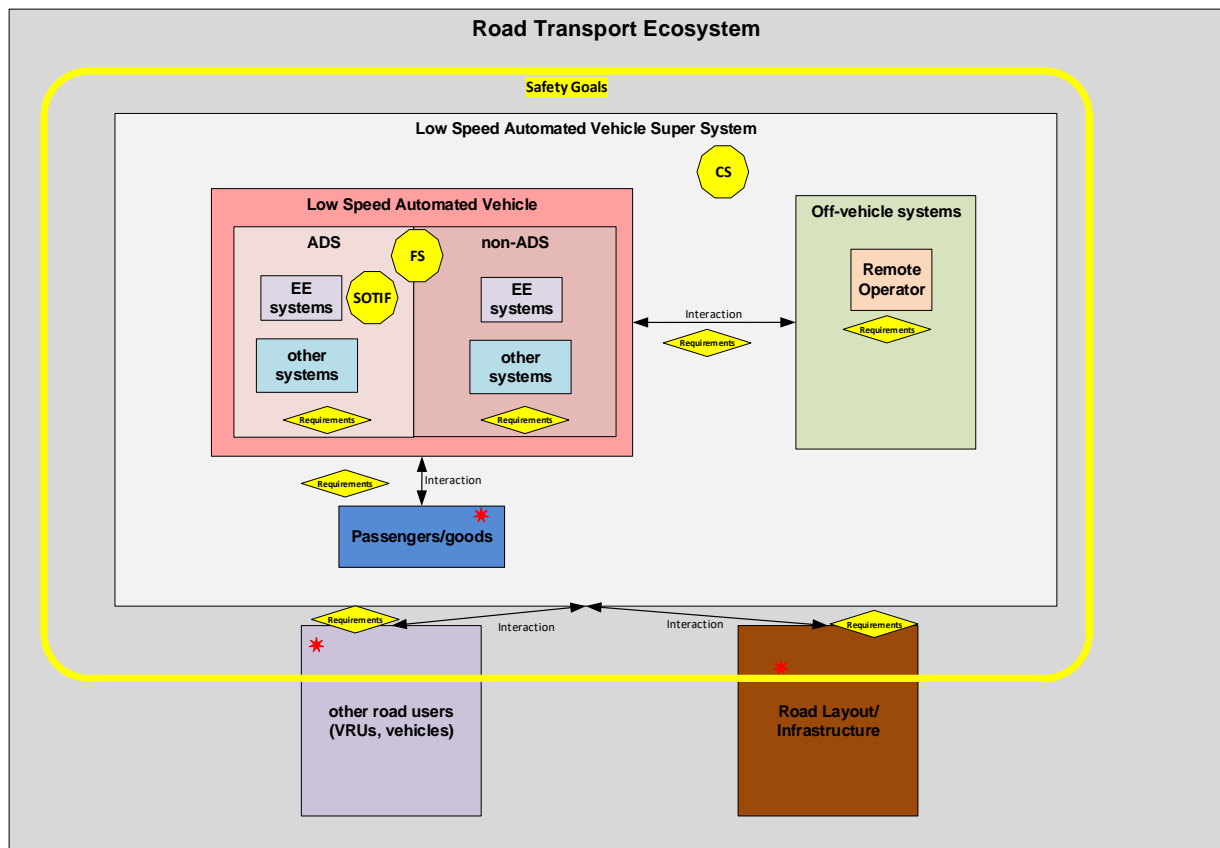


Figure 27 Automated Vehicle Super System and allocation of requirements

It is proposed to set out additional requirements which define the “safety rules” for the assurance scheme in more detail, covering

- safe interactions between elements of the road transport ecosystem (ADS, automated vehicle, passengers, operator(s), off-board systems)
- safe implementation of the ADS

At the point of type approval, it is the design and implementation of the LSAV that is assessed for its suitability for use on public roads, based on the assessment of documentation, the product, and audits of the organisational processes.

The diamond shapes in the diagram indicate what the different sets of performance requirements proposed for the scheme address. Additionally, they show the components of the super system that are covered by functional safety, SOTIF, and cybersecurity requirements.

The requirements sets are structured in the following way

- Addressed jointly by behavioural competencies
- Requirements on how the LSAV (being controlled by the ADS) is expected to interact safely with other road users -
  - Requirements on how the LSAV (being controlled by the ADS) is expected to move safely within the road infrastructure
  - Requirements on how the ADS shall interact with occupants to ensure their safety
  - Requirements on the necessary interaction between the ADS and remote external overnight

- Requirements on the ADS vehicle system implementation [note these requirements might span non-ADS systems, particularly to ensure that the ADS capability includes safety-relevant non-DDT tasks like monitoring of system health and system states.]
- Requirements on manufacturer and operator and their interactions (if affecting vehicle design)

To maintain traceability to the ongoing work and structure at EU and UNECE level, Table 15 shows a mapping between the proposed GB scheme outline and the High-level Safety Goals of the latest EU and UNECE drafts.

As well as requirements on what performance the design and implementation of the ADS must achieve for safe behaviour, it is proposed that the GB Safety and Security Scheme should place requirements upon the manufacturer and the operator such that they need to meet certain obligations for ensure safety of the automated vehicle. This includes following best practice guidance throughout all lifecycle phases and includes the guidance described in 5.1 (Functional Safety), 5.2 (SOTIF) and 5.3 (Cybersecurity and software updates).

The summarised proposed technical requirements for the GB approval scheme are presented in Section 5.12.

EU High Level Safety Goal			GB	SGs	UNECE High Level Safety Goal	
2 Dynamic Driving Task (DDT) under nominal traffic scenarios.	3 DDT under critical traffic scenarios (emergency manoeuvre).	4 DDT at system boundaries	Requirements on how the LSAV (being controlled by the ADS) is expected to interact safely with other road users	1,2,4,5,6,7,8,9,11,12,13,14,15,16,17,18,19,20	The ADS should drive safely	The ADS should manage safety-critical driving situations
			Requirements on how the LSAV (being controlled by the ADS) is expected to move safely within the road infrastructure			
7 Human machine interface for vehicles transporting vehicle occupants			Requirements on how the ADS and vehicle shall interact with occupants to ensure their safety	3,10, 20	The ADS should interact safely with the user	
7 Human machine interface for vehicles transporting vehicle occupants [and across 3, 5, 6, 10, 11, 12]			Requirements on the necessary interaction between the ADS and remote assistants			
5 DDT under failure scenarios			Requirements on the ADS and non-ADS vehicle systems	n/a	The ADS should safely manage failure modes	
6 Minimum risk manoeuvre						
8 Functional and operational safety during the ADS lifecycle					The ADS should maintain a safe operational state	
11 Manual driving for emergency cases or for the purpose of maintenance or similar cases						
8 Functional and operational safety during the ADS lifecycle	9 Specific requirements regarding Cybersecurity and Software-Updates	10 Specific requirements regarding data recorder for ADS			The ADS should maintain a safe operational state	
					Requirements on manufacturer and operator	
12 Operation manual					<i>not addressed in 21st FRAV session document</i>	

13 Provisions for periodic roadworthiness tests			<i>not addressed in 21st FRAV session document</i>
---	--	--	--

*Table 15: UNECE Requirements identified for consideration in the GB scheme.*

## 5.5 Minimal Risk Manoeuvres and Conditions

In case of any failure or condition that prevents the ADS from performing the DDT safely, there must be a means to achieve a safe state for the vehicle. This might be one single state, e.g., immediate vehicle deceleration to stationary, or there might be different safe conditions that can be achieved through different manoeuvres. The trade-offs that must be considered when deciding on safe states and the risks associated with them have been further developed in the SafeMRX project. The conclusions are currently being documented and the resulting report is recommended here as a future reference, but is not within the public domain at the time of writing.

### 5.5.1 Proposed Requirements

The manufacturer shall describe the different MRCs (minimal risk conditions) that the LSAV can enter, together with the description of the MRMs (minimal risk manoeuvres) that achieve them. The conditions that trigger each MRM and MRC shall also be documented. It shall be mandatory to have at least one MRC and one MRM.

Data on the activation of MRMs and achievement of MRCs shall be collected as part of the data that is monitored during operation, such that the cause and the outcome of its activation can be evaluated.

Each declared MRC shall be evaluated in the context of the TOD for its acceptability with respect to the risk involved in achieving the condition. For example, stopping in lane could be acceptable if the LSAV can travel in a dedicated lane or if alternative lanes are available for the other road users to pass safely while recovery can take place in a timely manner. If stopping in lane causes other traffic users to perform potentially dangerous passing manoeuvres in the presence of oncoming traffic this, would be considered unacceptable, particularly if the likelihood of the MRM occurrence was judged to be higher.

For the assessment of MRM functionality that a manufacturer sets out, it is pertinent to ensure that system faults, limitations and malfunctioning behaviour are taken into account in their design. Where possible, it is preferable for backup systems to be available to provide sufficient fail-operational performance such that the need for performing MRMs and reaching MRCs is minimised. Where this is the case, it would typically be expected that the system would transition to a degraded mode to allow continued operation, e.g. by providing a more restricted set of behavioural competencies to compensate for poor weather or subsystem faults. The interaction between the ODD/TOD, behavioural competencies and MEL was considered within Section 4.1.

Examples of possible implementations include:

- Remote assistance (e.g., 'ok to go' decision)
- Continued 'degraded' operation
- Discontinuation of operation including reaching an MRC
  - Different MRC options present different risks. Considerations need to include what safe locations can be achieved ('MRM capability'), which may depend upon the event (ADS failure / vehicle failure / TOD exit) that caused the MRM. Examples of safe locations for an MRC include:
    - Stopped in lane
    - Stopped at side of road
    - Stopped at designated safe area
    - Stopped at destination / depot

## 5.6 External Inputs

### 5.6.1 Background

#### 5.6.1.1 Definition of Problem Addressed

Whilst external inputs to the system, such as wireless 'V2X' communications, have the potential to expand the capabilities of an automated transport system by supporting live updates for traffic and fleet management, relaying safety information, supporting the role of a remote assistant and allowing system updates, nonetheless it raises considerable challenges relating to cybersecurity, robustness, human factors and version control. The cybersecurity and system update aspects are considered specifically within Section 5.3, and the human factors aspect in Section 5.7. Performance requirements relating to the interaction between the LSAV and remote assistance are considered in Section 5.4.

In contrast, the objective of this section is to examine external inputs holistically in order to draw together these separate topics and consider what issues may be encountered at vehicle level, and what restrictions and mitigations may be used to support safe operation.

#### 5.6.1.2 Current State of the Art

Automated vehicle trials currently use external inputs such as wireless communications for a number of reasons, with examples including:

- Use of GNSS (global navigation satellite system) to identify position, orientation and speed, thereby enabling or enhancing the performance of the dynamic driving task (Impacars, 2021).
- Use of vehicle to infrastructure (V2I) and vehicle to vehicle (V2V) wireless communications, collectively referred to as V2X, to support the performance of the dynamic driving task such that manoeuvres can be performed in a collaborative manner and hazards beyond the visual line of sight can be foreseen (MuCCA, 2019). Such inputs could be:
  - Essential, i.e. the DDT cannot be performed safely without them (e.g. if the system is only able to determine traffic light colour via wireless signals).
  - Non-essential, i.e. they provide an enhancement to the functionality, but the system can operate safely without them.
- Updates to software or to other data used by the system (e.g. calibration files, onboard HD maps) to improve the operation of the system. These would not affect the real-time performance of the DDT, provided that the update doesn't occur during performance of the DDT, but could affect how the DDT is performed in the future. Such updates could be:
  - Over-the-air (OTA) updates using a secure wireless connection.
  - Updates performed by plugging in a wired connection (e.g. via a laptop or USB memory stick)
- Live updates on traffic flows – these would not directly affect the tactical decisions of the ADS, but could affect strategic decisions such as route planning.
- Communications with a remote assistant or other member of staff with safety responsibilities.
- Communications related to maintenance and system monitoring, such as a request for the system to transmit data from a self-diagnosis – these will not be examined further here as they do not directly relate to the ADS or the DDT, and instead fall under the scope of Work Package 4.
- Receiving instructions from users relating to the journey to be undertaken, e.g. via a phone app that allows start and finish points of journey to be entered.

One example of oversight within the Law Commissions' report is the provision of remote assistance, whereby a remotely located human can support the operation of the ADS by, for example, confirming the suitability of a path that the system is uncertain about. Remote assistance could also include communication with passengers. Oversight duties may involve 'fleet operations', such as authorising a trip or responding to emergencies. However, if the system required continuous monitoring and control from the remote human, it would not be classed as self-driving, and would therefore fall outside the scope of the Law Commissions' proposals.

The report observes that, while there may be some use cases where assistance can be provided by those in the vicinity of the vehicle, it is anticipated that the vast majority of permutations will involve the use of a remote operations centre. This presents many challenges, such as:

- Ensuring the robustness of a comms link, including consideration of time lag, which can make control difficult if this is significant, and especially if the time lag is variable rather than consistent
- Assuring the cybersecurity of the connection
- Ensuring that staff have the appropriate equipment to gain situational awareness and to maintain their performance (e.g. without experiencing motion sickness)
- Ensuring appropriate training is provided
- Ensuring appropriate break periods are provided
- Ensuring there is sufficient capacity – there must be enough staff and equipment to cover peak demand, not merely average demand, and events such as flash floods or civil unrest could result in assistance demands far above average.
- Appropriate means to communicate with passengers

The report recognises arguments in favour of either the manufacturer or the operator being responsible for providing updates to the system; it provisionally favours the operator assuming responsibility, but allows flexibility for this to change in the light of experience. In practice, the manufacturer and Operator may be the same entity, in which case the distinction would assume lower relevance, although it should still be ensured that the particular department or team responsible is identified in the safety case.

UNECE Regulation 79 (2017) includes provision for low-speed automated parking systems; these fall within ACSF (automatically commanded steering function) category A, and are restricted to a maximum speed of 10km/h. Within the provisions for ACSF category A systems are requirements relating to RCP (remote controlled parking). For RCP, it is required that:

- "The parking manoeuvre shall be initiated by the driver but controlled by the system. A direct influence on steering angle, value of acceleration and deceleration via the remote control device shall not be possible". As such, it will be noted that RCF stops short of full remote control; it requires that the ACSF provides the vehicle control, with the remote signal merely acting as an input to the system. This may be expected to mitigate concerns about human factors and situations awareness, and potentially also cybersecurity.
- "A continuous action of the remote control device by the driver is required during the parking manoeuvre" – in effect, the remote control device is required to function as a 'dead man's handle' such that it is failsafe against any loss of input from the driver, or indeed any loss of connectivity.
- "If the continuous activation is interrupted or the distance between vehicle and remote control device exceeds the specified maximum RCP operating range or the signal between remote control and vehicle is lost, the vehicle shall stop immediately" – it is further elaborated that "the specified maximum RCP operating range shall not exceed 6m".

In the absence of a more objective means to identify an optimal maximum speed, the UNECE Regulation 79 maximum speed for RCP, 10km/h, seems to be a reasonable and pragmatic benchmark to follow.

Attention is drawn to the study undertaken within the Endeavour project (TRL, 2021b), which examined the steps that would be required to permit remote operation to take place without the safety driver that was used within the project. The report on this work included a roadmap for how to progress to full



remote driving, with the caveat that the timescales depend upon the rate of progress achieved within each prior step. It is therefore recommended that regulators continue to monitor progress in this field such that they can identify if and when a point is reached in the future that remote driving can be achieved safely and securely.

UNECE Regulation 79 and the Endeavour project make an interesting comparison, which sums up a wider industry trend; whilst there is ambition within the research domain to develop solutions that allow full remote operation, this technology remains immature, and has not yet been demonstrated to be capable of being made sufficiently robust and secure to allow direct remote control of the vehicle on public roads without a safety driver. Automotive regulations, therefore, adopt a far more cautious approach, requiring solutions that are inherently more robust with regards to cybersecurity and human factors.

### Stakeholder Feedback

One of the stakeholders within the first round of consultation, who is an ADS developer, stated that they are 'starting their journey' towards using remote external inputs but that it comes with many challenges such as cybersecurity, delays in the communication, robustness, and the need for operators to have specific skills. Their stance overall reflected interest in the potential, but awareness of the pitfalls and of the challenges that would need to be overcome to use wireless external inputs safely. Similarly, a representative of a vulnerable road user group expressed concern at the use of wireless inputs to control the vehicle and whether this can ever be made acceptably safe, even with failsafes incorporated into the system. A local authority representative suggested that it would be a concern if systems relied entirely upon wireless inputs for safety, and that they should be able to maintain safety in the absence of the wireless inputs.

Overall, a large majority of respondents expressed strong reservations about external inputs being required to perform any safety-critical functionality, and felt that it is not feasible for a system to rely upon wireless communications for the foreseeable future, unless a safety driver is present to mitigate hazards. Multiple partners expressed their view that it is possible to engineer a system to operate without relying on connectivity, and that this would be good engineering practice. However, one respondent was happy for wireless inputs to be needed for safety critical tasks provided that some suitable means could be applied to assure that the security and robustness are acceptable.

One respondent expressed general concern about remote operation, with a preference instead for remote inputs to only be used to approve 'discrete command requests' (e.g. where a bin bag in the road confuses the ADS, and the remote assistant is asked to approve the vehicle's suggested action). However, they pointed out that an exception to this may be where the remote operator needs to move a vehicle out of a live lane; they stated that there might be use cases that justify this, but that it would depend upon aspects such as whether other actors are in the vicinity of the vehicle, and that they see remote operation as being more appropriate for closed sites than the public road. Interestingly, the same stakeholder went on to say that they couldn't envisage automated vehicles relying on wireless communications for safety critical tasks, and that they should always be able to perform an MRM safely in the event of a communication drop-out. There was strong concern from multiple stakeholders about the concept of remote operators being required to take over rapidly while the vehicle is in motion, with a preference instead for the vehicle always being able to perform an MRM.

Taken as a whole, the stakeholder feedback from the first consultation suggested that, if any wireless inputs are to be used to aid performance of the DDT, the system must still be able to function safely if they are absent or corrupted (e.g. by detecting the issue and performing an MRM or providing degraded functionality), and inputs from remote operators shouldn't be relied upon as a means to respond safely to a situation that arises whilst the vehicle is in motion. There was no clear consensus upon whether remote operation where the operator has a long period of time to take over (e.g. if the vehicle has already stopped) would be acceptable, but if used, it would need to be designed in a failsafe manner.

The second round of stakeholder consultation took place late in the project, once initial proposals were in place, using a survey format to present a summary of key proposals which were expected to be more divisive, and then solicit feedback. In the case of the external inputs section, it was tentatively proposed that wireless inputs should be allowed to be used to support the ADS (e.g. a remote assistant approving an action that the ADS has proposed), but that wireless inputs shouldn't be allowed to directly control the actuators. This was put forward in order to address concerns about cybersecurity and human

factors. The proposal allowed wired connections to control the actuators directly, e.g. for manoeuvring in a depot.

The approach proved controversial, with many stakeholders recognising the concerns around remote operation, and how the initial proposal could help address this, but with significant concern that such a blanket approach would be overly restrictive by preventing remote operation. It was suggested that a more solution-agnostic approach should be adopted where remote inputs of any type can be used, provided that they are made sufficiently safe and secure (and evidenced accordingly). Furthermore, it was pointed out that there may be some use cases where not allowing remote control presents a higher risk, e.g. if the vehicle is stuck in a live traffic lane (although it should be noted that it would be possible to move a vehicle in such circumstances without requiring real-time remote control, e.g. by selecting paths suggested by the system, or by plotting a suitable path).

Overall, there was a moderate consensus against the proposal, and in favour of proposals that offer more flexibility to allow full remote operation. However, it should be borne in mind that these requests to accommodate remote operation primarily came from developers of automated vehicle platforms and/or ADSs – as noted previously, there is a wide disparity between the permissive approaches favoured by those operating within the automated vehicle research domain, and the more conservative approaches favoured by those within the more 'traditional' automotive industry who are concerned with the realities of bringing safety-critical systems to market. Furthermore, it should be noted that the appetite for remote control within the second stakeholder consultation contrasted with the findings from the first round, where a smaller proportion of participants were ADS or automated vehicle platform developers.

Therefore, whilst it has been agreed that the original proposal will be made more flexible, nonetheless a cautious approach should be adopted, with robust assurance required before full remote operation is permitted.

### 5.6.1.3 Conclusions Drawn

Whilst the cybersecurity of LSAVs is considered elsewhere within this report, and will not be duplicated here, nonetheless the challenge of achieving acceptable cybersecurity has direct implications upon how external inputs can be used by the system, as it is vital to ensure that, given the inherent vulnerability of remote communications links, the system does not rely upon external communications in a manner that poses unacceptable risks.

It is therefore recommended that the type approval process should, for the foreseeable future, require systems to be able to perform safely in the absence of any external inputs such as V2X communications, and where such inputs provide incorrect data. The vehicle may offer reduced functionality without the external inputs available, such as performing an MRM to reach an MRC as soon as their absence is detected, and may be prohibited from commencing a journey if external inputs are unavailable, but it is essential that, should the external inputs be missing or incorrect (through errors or spoofing/ tampering), safety is be maintained.

As such, the use of external inputs should be seen as optional, but the assurance of safety in the absence of external inputs must be mandatory. In the future, as technology and assurance processes evolve, it may become possible to relax this requirement and allow external inputs to perform a real-time safety critical role within the performance of the DDT; however, this would need significant evidence to show that it is practicable to ensure acceptable safety and security.

It is expected that external inputs will be used by some manufacturers to update their systems, or the data held and used by the system, in order to correct errors, make improvements and adapt to changes. Such updates should not be permitted while the ADS is active, to prevent risk of the updates compromising safety while the vehicle is in motion. Good practice for cybersecurity should be applied to authenticate the updates prior to them influencing system behaviour.

Wireless inputs that directly control the actuators, such as remote driving, should not be absolutely prohibited, in recognition of the ambition of a range of stakeholders within this field, but should be treated with caution, subjected to extremely thorough safety oversight, and only used if there is a clear justification in terms of how the inherent risks will be mitigated and how they will be compensated for by risk reductions in other areas (e.g. due to the enhanced ability to move the vehicle out of a live traffic lane rapidly).

The human factors of external inputs, such as those by a remote assistant, must also be considered; however, this aspect is addressed within Section 5.7, and therefore will not be duplicated within the requirements here.

## 5.6.2 Recommendations

### 5.6.2.1 Proposed Requirements

The ADS shall be able to maintain safety in the event of the absence or corruption of any external inputs via wireless communications. This may be achieved through normal operation being possible without external inputs, through reduced functionality being provided when external inputs are unavailable, or through journeys being inhibited or terminated safely via an MRM when external inputs are unavailable or corrupted. Appropriate cybersecurity measures shall be in place to validate and authenticate wireless inputs; these must be proportionate to the threat level identified within a threat assessment.

The safety case shall therefore include consideration of every wireless remote input to the system and a justification for why the absence, spoofing, tampering or corruption of the input will not result in a hazard. This shall include consideration of at least two simultaneous faults (e.g. loss of GNSS data and corrupted V2X data occurring simultaneously), and also consideration of common cause failures (e.g. electromagnetic interference resulting in sustained interruption to wireless signals across a wide frequency band).

Wireless external inputs are permitted to indirectly affect the DDT, by acting as an input to the ADS in order to support strategic or tactical driving tasks. Examples of this include a remote assistant proposing or approving a possible path when the ADS would otherwise be unable to proceed, or an input from a passenger via an app requesting that the vehicle pull over at the next safe opportunity. Such inputs shall not directly control actuators such as traction motors, brakes or steering, but instead provide information to the ADS that the ADS can then act upon, meaning that the ADS remains responsible for performing the DDT.

Wireless external inputs that directly control the DDT such that the motion of the vehicle is able to be directly controlled via the wireless connection are not absolutely prohibited by these requirements, but special caution should be taken to ensure that appropriate cybersecurity and human factors measures are in place to allow such remote operation to be done safely; this should include rigorous analysis and testing conducted by the manufacturer, and a detailed audit supplemented by testing conducted or witnessed by the regulator, to ensure that the regulatory process provides appropriate safety assurance. The safety argument for allowing such direct remote control of the vehicle should include justification for how the inherent risks have been mitigated, including means to validate the wireless inputs and selection of a fail-safe architecture, and should show that it allows an overall reduction in risk to be achieved (e.g. by reducing the length of time for which vehicles remain stopped in line traffic lanes).

Regulators should monitor the state of the art with regards to remote control (i.e. remote operation that directly controls the actuators, rather than acting as an input to the ADS), and not permit it to be used within commercial deployments without a safety driver present until such a point is reached where it is demonstrated to be able to be made safe and secure to a level of robustness that is compatible with production systems.

In any situations where the system is subject to direct remote control or is proceeding upon the authority of a remote assistant having approved a requested course of action proposed by the ADS, the speed of the vehicle shall be limited to 10km/h.

It is not required that remote assistance or direct remote control should be limited to situations where the operator is close to the vehicle and within visual line of sight. This is because there may be many operational scenarios where a person outside the vehicle has an inferior view of the complete surroundings, compared to someone in a remote control centre, and because it is anticipated that many business models will rely on remote operation being able to be performed beyond visual line of sight. The safety case shall, however, justify how the situational awareness aspect of human factors has been addressed such that the ability of the remote assistant to perceive the vehicle's surroundings is at least as good as that for a conventional driver in a conventional vehicle.

Wired external connections are permitted to directly control the vehicle, e.g. plugging in a controller that allows low speed manual operation within a depot. Where this is the case, security mechanisms shall be provided to prevent improper use, such as the necessary connection point(s) being protected within a locked compartment or through electronic means to verify authority.

Wired or wireless external inputs are permitted to allow updates to the system (e.g. software updates, map updates). However, the system shall have means to validate the authenticity and integrity of such updates, in line with the cybersecurity requirements within this report, and updates shall not be made while the system is engaged in performing the DDT.

The system shall not require continuous monitoring by a human supervisor to ensure safety; as such, human oversight and inputs shall only be necessary for performance of the DDT when requested by the system. The safety case shall demonstrate that the system, including offboard elements, is failsafe such that the failure to receive such an external input when requested does not result in hazardous behaviour. It shall also be demonstrated that any latency associated with external inputs being received by the system, and any fluctuation in the latency, does not compromise safety.

The system may make use of external inputs from devices not under the control of the manufacturer, operator or a contracted 3<sup>rd</sup> party, e.g. an app on a customer's phone that allows the destination and user preferences to be set. Such external inputs may provide data to the ADS such that strategic planning is affected (e.g. route planning), but shall not influence tactical elements of the DDT (e.g. speed or path adopted to navigate a section of road). All devices capable of influencing the tactical driving task via external inputs shall be managed under the safety assurance processes and safety management systems of the manufacturer and/ or operator.

The 'Digital Information' section within the ODD and TOD definitions (see Section 4.1) shall include a definition of any external inputs that are used by the ADS to perform the DDT, or that could be reasonably foreseen to influence the performance of the DDT by the ADS.

### 5.6.2.2 Supporting Information

There is significant overlap between this section and the cybersecurity section, and the former does not attempt to duplicate the latter. As such, for guidance on assuring the cybersecurity of external inputs, a key factor to ensure such inputs can be relied upon in a safe manner, please refer to Section 5.3.

Similarly, the human factors section considers the human factors aspects of the remote assistant's role, which will not be duplicated here; please see Section 5.7.

The requirements seek to adopt a cautious approach by encouraging solutions where any wireless inputs are not able to directly control the movements of the vehicle, such that the ADS remains responsible for performing the DDT (in line with UNECE regulation 79's requirements in relation to remote controlled parking). This is to minimise the risk of cybersecurity breaches or comms link errors resulting in dangerous behaviour by the vehicle. By providing a level of segregation between the external inputs and the system actuators, it will be more difficult to use the external inputs as an attack surface. Ensuring that the ADS ultimately performs the manoeuvre provides a clear demarcation of legal liability and aligns with the Law Commissions' recommendations on responsibility (Law Commissions, 2022), whereas there could be the potential for responsibility to be less clear if an external input could be argued to be directly controlling some aspects of the vehicle.

However, to accommodate stakeholder feedback, the recommendations also seek to facilitate a more permissive approach such that external inputs are not absolutely prohibited from directly controlling the DDT, provided that appropriate safety assurance and justification is provided. It is not expected that such approaches will be able to be made safe and secure within the short term, and therefore regulators should be cautious about approving systems that use this approach, but it must be recognised that there is an ambition to make remote driving a reality, there are use cases where this could significantly enhance the commercial model for deployments, and there are ongoing research and standardisation activities in the field that could result in direct remote control being practicable within the longer term.

It should be noted that even where measures are in place to detect corrupted or lost wireless signals used for remote control, which can never be entirely protected against due to the physics of the transmission, a safe response will be challenging. The vehicle could react in one of three ways when the wireless signal suddenly becomes unavailable:

- Hand control back to the ADS – however, if the ADS is capable of resuming control, this begs the question of why remote control was required in the first place.
- Apply emergency braking – this would result in passenger discomfort, and a particular risk of harm for standing passengers, together with a risk of a following vehicle colliding with the LSAV.
- Apply moderate braking – this risks the vehicle deviating significantly from an appropriate trajectory before reaching a stop, although the restriction of remote control to 10km/h may help justify an argument that a vehicle can maintain safety via this approach.

Whilst the requirements don't prohibit the use of V2X or GNSS, which may allow enhanced functionality, it does require that the system does not rely on such inputs for safety; this is important as they do not provide an appropriate level of robustness to be relied upon for safety-critical purposes. Similarly, it is anticipated that apps on customers phones may be used as an input to the system, e.g. to select the destination, and this is provided for in the requirements, but it must be ensured that there is no means for this to provide control directly to the vehicle's actuators as it will not be possible to adequately assure the cybersecurity, functional safety or compatibility of such devices.

## 5.6.3 Future Considerations

### 5.6.3.1 Areas for Future Work

Significant research is needed into the practicalities of how remote assistance provided by a remote operations centre could work in practice. This would allow more prescriptive requirements and more detailed guidance to be written regarding the safety, security and human factors aspects, thereby supporting manufacturers, operators and regulators.

In particular, research is needed into how to ensure the remote assistant has the appropriate equipment to allow them to have situational awareness and an appropriate level of control, and how a mechanism can be developed that provides the required functionality whilst being fail-safe such that missing or erroneous external inputs do not present a hazard. In order to reach a point where direct remote control of the vehicle actuators is safe and secure, significant research will be needed into the cybersecurity and functional safety of the communications link.

It is desirable that V2X data can be used to more directly influence vehicle behaviour in the future, e.g. by allowing close collaboration on vehicle movements to smooth traffic flows through a junction. However, this is another area where considerable research, development and testing will be required before assurance can be provided that this can be done securely, safely and robustly.

## 5.7 Human Factors

Human factors performance requirements are necessary to support the perceived interactions an ADS will have with key parties affected by its deployment, including passengers, Remote Operators, and other parties in the event of an incident (for example, other road users, the emergency services and vehicle recovery services). Specific aspects of human factors requirements, such as performance, accessibility, monitoring, usability, and training, will vary for each party, and overlap in some cases; therefore, it is vital these are considered in turn to understand where recommendations should be fed into advice, via guidance, or embedded in regulatory requirements. The requirements documented focus on HMI, processes, and the identification of types of information that may be required by the affected parties.

The following recommendations are being proposed to ensure that Human Factors have been taken into account to a sufficient extent for the safe design and operation of a LSAV. The recommendations have been formulated as performance requirements in the Section 5.12 (interaction with vehicle occupants and external oversight).

This report section aims to address a number of aspects of automated driving system (ADS) functionality relating to human factors considerations; these functionalities focus on specific aspects of ADS requirements. Compared to a traditional vehicle with a driver, an AV will need appropriate systems to ensure that the ADS can interact with:

- Passengers
- Remote Assistants
- Other parties in the event of an incident, such as other road users, emergency services personnel, and recovery personnel.

Human factors considerations for maintenance/depot staff and other road users during nominal automated vehicle (AV) operation are out of scope of this report.

Performance requirements describe the requisite behaviour that an ADS needs to be able to achieve in order to be considered safe and secure within a specified operating domain. It defines and validates a “base set” of safety functionality which is required for all vehicles operating in a particular application domain in order to meet the claims set out by the safety case. For the purposes of WP1, four areas of performance requirements are being developed including: nominal performance, human factors, Minimum Risk Manoeuvre (MRM) and external inputs.

This report focusses on human factors performance requirements only and draws on current best practice and experience from other industries to create high-level requirements to support the perceived interactions with the parties outlined above. The requirements centre around human factors considerations for safe deployment, with specific focus on performance, accessibility, monitoring, usability, and training. To be concise, where considerations overlap significantly, information has been integrated between these areas. How the recommendations in this study relate to guidance, regulations or standards needs to be considered.

### Principles of human factors design

The use of human factors design principles and processes could contribute to effective<sup>3</sup>, efficient<sup>4</sup> and satisfactory<sup>5</sup> development of Human-Machine Interfaces (HMIs) and communications designs for ADSs. This could be achieved by using methods and processes to enhance the usability<sup>6</sup> of interfaces and

---

<sup>3</sup> Effectiveness can broadly be described as the ability of the user to achieve the goals set out in a usability evaluation (Barber, 2015).

<sup>4</sup> Efficiency can be defined as the effort required by the user to achieve the goals set out in a usability evaluation (Barber, 2015).

<sup>5</sup> The concept of satisfaction is complex, however it can be described as “...all aspects of the user’s experience when interacting with the product, service, environment, or facility” (ISO 9241-210, 2008) and includes concepts such as aesthetics, pleasure and hedonic goals (Barber, 2015).

<sup>6</sup> This can be defined as “...the extent to which a product can be used by specific users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” (ISO 9241, 1998).

information processes to ensure that these are fit for purpose and usable. Essentially, the end user plays a key role in the testing of designs throughout an iterative process as illustrated in Figure 28, which is based on the work of Travis and Hodgson (2019). The iterative nature of the design process means that each part does not stand in isolation, but that there is continuous feedback within and between different parts that puts the user, their needs and capabilities, in the centre.

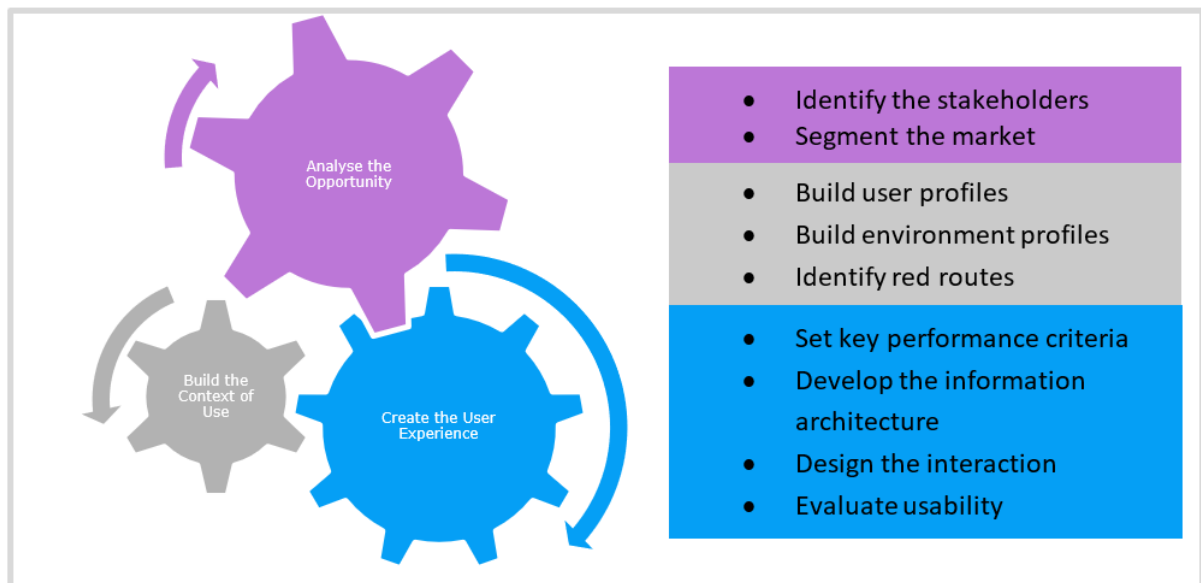


Figure 28: The iterative process of user experience design

## 5.7.1 Methodology

The scope of the human factors performance requirements focusses upon specific aspects of human factors for each user group that were identified as part of the study, including consideration of the HMI, processes and the identification of types of information that may be required by users. The high-level human factors recommendations draw upon current legal requirements for AV testing and trialling, experience from other industries, human factors and usability processes and a review of ongoing research within industry and academia.

The evidence was gained through a 'rapid literature review' consisting of four tasks:

1. Search terms to be used.
2. Assessment of the quality and relevance of identified literature.
3. Review of full texts.
4. Review of texts referenced in literature for relevance.

A list of search terms relevant to the research questions was generated to run the evidence review (see 10.1). These search terms were applied in several research databases (e.g., Google Scholar, ScienceDirect). Multiple searches were conducted within each database through search terms which were used individually and in combination with each other to identify which terms generated relevant results. Additionally, references cited in articles were reviewed and the information added to the report if relevant. This ensured that the review was as in-depth as possible.

## 5.7.2 Results and recommendations

### 5.7.2.1 Onboard occupants

One of the proposed service requirements for LSAVs, or more generally, AVs, is for the transportation of passengers.

Currently, the trialling of AVs in the UK, whether for passenger transportation or freight, has to comply with UK regulations (DfT, 2019). These regulations provide a basis for understanding the range of requirements that may need to be addressed for the safe and secure operation of AVs in the UK. This subsection considers ADS and on-board user HMIs, and information needs, for automated pods.

### Performance, monitoring and training

It is important to consider the interactions between an ADS and on-board users in conditions where on-board occupants:

- May feel that the driving of an ADS is unsafe, or
- Need to instigate an MRM in an emergency.

A distinction can be made between these two conditions in terms of information and HMI needs. In the case of an emergency, passengers will need to act. However, if passengers feel unsafe during normal operation, it could be unsafe to, for instance, instigate an MRM.

#### Perceptions of safety during normal operational conditions

Understanding the contexts in which passengers feel unsafe in an AV during normal operational conditions provides insights into the types of information and feedback that could provide reassurance. The prediction of these types of concerns could reduce the amount of unnecessary Remote Assistant to passenger communications, and instigation of MRMs.

Passenger perceptions of safety and security have been shown to be impacted by:

- Accurate detection of objects in the dynamic driving environment and reactivity to its operational context. In their study, Mouratis and Serrano (2021) found that the majority of participants (88%) felt safe during an AV ride. A contributing factor was that they felt that the AV accurately responded to the objects they could perceive in the environment.
- The low speed at which the AV was travelling (Mouratis & Serrano, 2021). However, some participants also felt that the AV was driving too slowly (in this project the maximum speed was 18 km/h) and that the vehicle speed will need to increase in the future (Mouratis & Serrano, 2021). Therefore, setting a speed for the AV needs to be balanced by service needs of passengers.
- Providing passenger information. Passengers discussed the need for continued feedback about the status of the AV and actions (Dichabeng, Merat, & Markkula, 2021).
- Supervisor/Operator on-board. The presence of a Supervisor/Operator on-board the AV had a positive impact on perceptions of safety and in-vehicle security (Dichabeng, Merat, & Markkula, 2021; Salonen, 2018). However, the impact of having a Supervisor on-board the vehicle could be mitigated by frequent departures, reasonable service costs, safety, speed and travel comfort (Mouratis & Serrano, 2021).
- Hard or abrupt braking contributed to perceptions of unsafe operation. Mouratis and Serrano (2021) reported that participants felt that they could grow accustomed to experiencing hard braking events but that receiving some form of reassuring communication would be useful.
- Sharing an AV at night-time with strangers (Dichabeng, Merat, & Markkula, 2021). Hohenberger *et al.* (2016) suggested the implementation of emergency buttons to give passengers a sense of control and thereby reducing anxiety. Awareness of CCTV surveillance tends to have a moderating impact on anti-social behaviour (Jansen, Giebels, van Rompay, & Junger, 2018; UK Government, 2012).

#### Passenger safety during emergencies

Some instances of emergency contexts in which passengers may need to instigate an MRM or contact with the Remote Assistant include:

- Having a disruptive or ill passenger onboard (Dichabeng, Merat, & Markkula, 2021).
- The AV being involved in an accident or AV failure (Dichabeng, Merat, & Markkula, 2021; Salonen, 2018).



- Passengers falling, tripping or becoming trapped in a door. These represent the majority of accidents that occur on buses and trains in London (Transport for London, 2020).
- Terrorism or acts of violence. The Department for Transport (2018) provides guidance on how organisations could plan for and help to prevent these types of events.

Basic on-board user-Remote Assistant responses to consider include:

- Passenger performs an action (e.g. pressing a button) to request that the ADS performs an MRM, with the Remote Assistant being automatically alerted.
- Passenger communicates directly with the Remote Assistant.
- Passengers do not respond to an emergency.

It is suggested that the following aspects of passenger behaviour is considered:

- Recognise that an emergency has occurred.
- Understand what action/s to take.
- Recognise the interface to use to elicit the action.
- Have the physical and mental capability to complete the action.

### Accessibility and usability

This section of the report addresses consideration of the information needs of users (including visually impaired, deaf, wheelchair users and children) using an AV. The accessibility requirements of public transport vehicles are set out in The Public Service Vehicles Accessibility Regulations (2000). This provides guidance for physical access to bus services; however, it does not cover the information requirements relating to planning a journey, waiting for transport or on-board information. The Public Service Vehicles Regulation (2020) states that:

*“The consequences of this are significant; visually impaired people report regularly missing their stop or being stranded in an unfamiliar place, deaf passengers are unable to confirm their location and wheelchair users travelling backwards in the wheelchair space are prevented from following their journey”.*

In buses, some tasks around accessibility are performed by the drivers, such as providing route and bus stop information, selling bus tickets, dealing with emergencies, managing passenger capacity, providing security for vulnerable passengers and supporting passengers with mobility needs to embark and disembark (DVSA, 2013). Consideration needs to be given to how AVs and the information about services could comply with these regulations.

For instance, Kempapidis et al. (2020) researched how people with visual impairment experienced AV pod rides and found that they experienced higher levels of anxiety compared to the sighted participants. They also found that audible feedback during braking events decreased levels of anxiety (Kempapidis, et al., 2020).

Accessibility also relates to differences between age groups. A number of studies found that age was negatively associated with the likelihood of using AVs (Dichabeng, Merat, & Markkula, 2021; De Vos, Waygood, & Letarte, 2020; Mouratis & Serrano, 2021). At the same time, little research has been done regarding the use of AVs for children; the papers that have been found focus on the use of AVs for children to replace services more closely resembling that of a private vehicle (Koppel, Lee, Mirman, Peiris, & Tremoulet, 2021). However, in their study of the attitudes of US parents to using AVs to transport their children (eight- to sixteen-year-olds), Tremoulet et al. (2019) found that the greatest fear parents reported was that their unaccompanied child(ren) would not be protected by the AV during an unplanned trip interruption. This finding could potentially also impact on the use of AVs by unaccompanied children. Two-way audio communications and video feeds of vehicle interiors, seatbelt checks, automatic locking, secure passenger identification and remote access to vehicles were considered to make the use of AVs more appealing to parents (Tremoulet, Seacrist, McIntosh, DiPietro, & Tushak, 2019).

Dichabeng et al. (2021) found that 86% of their participants referred to convenience as an important factor in influencing their choice of transport. Easy booking and payment, connective ports for charging

phones and onboard internet were mentioned as some of the expected features, while all the participants agreed that reliability is an important factor.

### 5.7.2.2 Remote Assistants

For trialling any level of AV technology on a UK road, a safety driver would typically need to be present within the vehicle, with any trials using an alternative method of oversight such as a remote operator falling under the category of ‘advanced trials’ and requiring permissions to be sought (DfT, 2019). A commercially-deployed AV (i.e. one that is not in a trial) may require a person to provide inputs to support the driving task in situations where the ADS is not able to fully complete it unaided, such as approving the suitability of a proposed path, but they should not be required to continuously supervise the ADS’s performance of the DDT.

As opposed to a safety driver, which refers specifically to safety operators who are located within the vehicle with access to traditional driver controls, if the safety operator is located outside the vehicle, either within or beyond visual line of the sight, they are classed as a ‘remote assistant’ (RA). For the purposes of this report, we consider RAs who are beyond the visual line of sight of the AV only. In the future, having a RA may not be a requirement for the approval of AVs. However, whether an operator or service provider chooses to have a RA or not, the safety and support of passengers throughout their journey needs to be done safely.

#### Performance and usability

According to the Code of Practice: Automated Vehicle Trialling (DfT, 2019), trialling organisations should establish a process to monitor the situational awareness of safety operators and to capture information regarding driver distraction and inattention; whilst not directly referring to RAs within commercial deployments, nonetheless this forms an informative benchmark. UN ECE Regulation 157 (ALKS, 2021) requires that *“if the driver fails to resume control of the DDT during the transition phase, the system shall perform a minimum risk manoeuvre. During a minimum risk manoeuvre, the system shall minimise risks to safety of the vehicle occupants and other road users”*; again, this is not directly applicable to LSAV deployment, but forms an informative benchmark.

This minimum risk condition may vary depending on the target operating domain, and could for example include slowing down, coming to a complete stop, or moving to a place of safety (DfT, 2019). This section reviews findings in terms of situational awareness, the information that RAs require and control room design. Associated technical challenges with the performance and use of RAs should also be considered as part of a robust safety case but are out of the scope of this report.

#### Challenges of the RA driving task

When humans conduct tasks, such as driving, sampling and processing of multiple sensory information (Groeger, 2002; Molholm, et al., 2002), we use cues from different senses to respond safely to the dynamic driving environment and support the anticipation of future events (Stahl, Donmez, & Jamieson, 2014). Assigning RAs the task of providing supporting inputs to the ADS creates challenges as the sensory information delivered through the interface presents a limited view of the entire traffic environment (Linkov & Vanžura, 2021), for example limited depth cues (Fong, Thorpe, & Baur, 2001), limited aural and haptic cues (Gnatzig, Chucholowski, Tang, & Lienkamp, 2013). Time delays in the signal’s transmission, lighting problems at the vehicle’s location, and the vehicle’s interface complexity all present further challenges to the RA task (Linkov & Vanžura, 2021).

RAs may be required to be available to support a number of different vehicles that are operating simultaneously. A RA who has been alerted by an AV may need to “drop in” and assess the problem or make positive inputs (e.g. approve a proposed course of action). This means that they will first need to get an understanding of the remote environment, what is happening in it and what might change (Mutzenich, Durant, Helman, & Dalton, 2021). During the process, there may be a lack of sufficient understanding of their own skill deficit to manage the situation and they may potentially misunderstand the limitations of the ADS (Pattinson, Chen, & Basu, 2020). A careful design of the displays and controls can mitigate these problems, thereby allowing for higher situational awareness (Linkov & Vanžura, 2021).

**Aspects of situational awareness of the RA**

Situational awareness or situation awareness (SA)<sup>7</sup>, is a dynamic process in which elements of the environment are perceived and interpreted, and implications of their future states are assessed (Endsley, 1995). Endsley describes that information from the environment is continuously processed in three levels to create and maintain SA:

- Level 1: perception of the elements in the environment.
- Level 2: comprehension of the current situation.
- Level 3: prediction of future status.

Furthermore, when a remote operation interface is evaluated, consideration of these levels helps to determine any error and improve the design (Krajewski, 2014). When information on an environment is not perceived, or information is interpreted incorrectly, it can lead to responses that do not meet the requirements of the situation.

Table 16 provides an overview of the variety of informational domains that an RA needs to be aware of when overseeing the operation of an AV. The findings are integrated and adapted from a number of research studies.

Domain characteristics	RA understanding	Aspects of domain characteristics
<b>Characteristics of the AV</b>	The AV's current status The quality of vehicle sensors	<ul style="list-style-type: none"> <li>• Including speed, manoeuvres activated, signals and any current faults identified</li> <li>• Including connectedness to RA</li> <li>• Including an understanding of any limitations, e.g., sensor blind spots</li> </ul>
<b>Communications</b>	The quality of vehicle communications	<ul style="list-style-type: none"> <li>• This includes an understanding of vehicle-operator communications as well as vehicle communications to other road users, and includes any limitations, such as latency, between vehicle-operator communications</li> </ul>
<b>The operating environment</b>	Weather conditions Location of objects Status of objects The terrain	<ul style="list-style-type: none"> <li>• Impact on the AV, its sensors and the behaviour of other road users</li> <li>• Awareness of objects that could impact on safe operations</li> <li>• Including the status of traffic lights, vehicles and pedestrians, the road conditions</li> </ul>
<b>RO knowledge and understanding</b>	Traffic rules and regulations Understanding of signage	<ul style="list-style-type: none"> <li>• As applied to the operating environment and impact on the ADS operations</li> <li>• Impact on the operation of the ADS, passengers and other road users,</li> </ul>

<sup>7</sup> Defined as "...the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future." (Endsley, 1995)

	Passenger and cargo status <sup>8</sup>	including vulnerable road users and those with disabilities
<b>RO skills set and capabilities</b>	<p>Application of previous experiences to current events</p> <p>Predicting potential future events based on integration of experience, knowledge and understanding</p>	<ul style="list-style-type: none"> <li>• Being able to apply knowledge gained through experiences to the operational environment</li> <li>• Interpreting the intentions of other road users, seeking a rationale/analysis of other driver’s actions, and predicting elements of the context</li> <li>• Making use of context to expand their comprehension of what was happening in a driving scene</li> <li>• Awareness of the spatial limitations of the vehicle for manoeuvring</li> </ul>
<b>Attitude towards traffic safety or risk-taking</b>	Adherence to safety rules and avoid committing violations	<ul style="list-style-type: none"> <li>• Having personality traits and attitudes conducive to the safe operation of vehicles</li> <li>• Being able to consistently attend to critical aspects of the driving scene</li> </ul>

Table 16: Requirements for the situational awareness of remote operators.

The prospect of RAs having to on occasion make safety-critical decisions relating to the driving task from a separate location makes it essential to identify how their SA needs will be different from those of a Safety Driver (Mouratis & Serrano, 2021). It is therefore important to understand how people use a monitor and video feed to build SA of the dynamic driving environment of an AV. (Mutzenich, Durant, Helman, & Dalton, 2021) found in their research on SA that their participants:

- Interpreted the intentions of other road users, seeking a rationale/ analysis of other driver’s actions.
- Made subjective estimations of the speed the vehicles were travelling at based on indirect cues, even if these were not always accurate.
- Had an awareness of the spatial limitations of the vehicle for manoeuvring, however higher levels of precision would be required to manoeuvre a vehicle safely.
- Made use of context to expand their comprehension of what was happening in a driving scene.
- Both SA comprehension and SA prediction is thought to develop in parallel.
- Elements of prediction are embedded in building up comprehension of a remote scene which could assist with hazard perception.
- Patterns of attention to specific aspects of the context within participants.

---

<sup>8</sup> This was added to provide a more comprehensive overview of the possible considerations regarding AVs and cargo pods. Most of the studies focused on AVs that had a Supervisor/Safety Driver onboard and therefore did not include this aspect of RO.

When considering SA in the context of remote operations, Table 17 summarises some of the factors that impact on how RAs experience their tasks and what the impact of this is on their performance.

Factors that impact remote operation	Influence on situational awareness
<p><b>Workload</b></p>	<p>Automation generally decreases perceived workload (Edwards, Homola, Mercer, &amp; Claudatos, 2017), which could lead to boredom and reduced SA, with reaction times decreasing and direct attention significantly decreasing over time (Cummings, Mastracchio, Thornbury, &amp; Mkrtychyan, 2013). This could lead to longer takeover times (Clark, McLaughlin, &amp; Feng, 2017) and overreliance on the system (Cooke, 2006).</p> <p>Perceived workload is affected by the degree of discrepancy between remote operation conditions and those that would be experienced when driving a vehicle (Mizukoshi, et al., 2020).</p> <p>(De Winter, Happee, Martens, &amp; Stanton, 2014) showed that highly automated driving can reduce workload and can potentially increase SA if operators are motivated or instructed to pay attention to the environment instead of engaging in other non-driving tasks or no instruction is given.</p> <p>(Lu, Coster, &amp; De Winter, 2017) found that participants took longer to gain SA when they had limited time to do so, possibly because of the excessive workload.</p>
<p><b>Type of task</b></p>	<p>Tasks that constantly vary tend to continue to utilise cognitive resources (O'Regan, Faul, &amp; Marnane, 2013).</p> <p>A top-down guidance approach, with drivers being instructed to search for hazards during the takeover period, significantly increased the number of safety checks made (White, et al., 2019).</p>
<p><b>Experience</b></p>	<p>Well-practised tasks tend to be unaffected by workload (Engström, Markkula, Victor, &amp; Merat, 2017).</p> <p>Operators with less experience are likely to underestimate the cognitive resources required to complete the tasks due to the perceived benefits of automation (Stapel, Mullakkal-Babu, &amp; Happee, 2019).</p>
<p><b>Lack of embodiment</b></p>	<p>As information about the environment is mediated through screens and control panels, this could potentially lead to a partial understanding of the conditions of the remotely operated vehicle (Nostadt, Abbink, Christ, &amp; Beckerle, 2020).</p>
<p><b>Motion sickness<sup>9</sup></b></p>	<p>Can be triggered by the exposure to dynamic visual displays and simulated travel scenarios: in this case, it is known as visually-induced motion sickness (VIMS) and can lead to a reduction in task performance (Diels &amp; Bos, 2016).</p>

<sup>9</sup> Motion sickness is the result of the mismatch between the visual system and the vestibular system (located in the inner ear).

	Virtual reality was found to reduce the impact of latency on remote operators, which reduced the experienced motion sickness (Tikanmäki, Bedrník, Raveendran, & Röning, 2016).
<b>Alarms</b>	Over time, false alarms tend to lead to operator distrust in the system (Linkov & Vanžura, 2021).

*Table 17: Considerations of the factors that impact remote operations and their influence on situational awareness.*

## Training and monitoring RA performance

The Code of Practice for Automated Vehicle Trialling (DfT, 2019) requires trialling organisations to consider appropriate Safety Driver and Operator training. Whilst aimed at research trials, it could be seen as indicative of the need for similar training where an RA is undertaking a safety-critical task within a commercial deployment. Pattinson et al. (2020) argues that for an RA's consent to be valid, they must be aware of the risks and responsibilities involved. From this perspective, training could provide RAs with an appreciation of the nature and extent of the risk being undertaken, and as such, support the legal aspect of consent (Pattinson, Chen, & Basu, 2020).

Safety operators supervising public road trials should (DfT, 2019):

- Understand the capabilities and potential limitations of the technologies under trial.
- Be familiar with the characteristics of the vehicle, preferably through extensive experience of trials conducted on closed roads or test tracks.
- Be aware of risks, such as latency and loss of contact with the vehicle.
- Be trained to mitigate and safely respond to any connectivity or control issues.
- Be aware of the situations in which it may be necessary to intervene.
- Be aware of potential hazardous situations that may be encountered and the appropriate action to take when resuming manual control of the vehicle.
- Be fully aware of exactly how control is passed between the safety operator and the vehicle.
- Fully comply with the existing laws regarding driver behaviour.

Training plays a critical part in the knowledge and skills of RAs. For instance, research shows that well-practised tasks tend to be unaffected by workload (Engström, Markkula, Victor, & Merat, 2017) and that Operators with less experience are likely to underestimate the cognitive resources required to complete the tasks (Stapel, Mullakkal-Babu, & Happee, 2019). A step towards providing the right training is doing a needs assessment based on a job analysis, that will identify the potential learners, their necessary prerequisite knowledge/ skills and the instructional objectives (Proctor & Van Zandt, 2008).

There are few standards in the automotive industry that regulate how training should be done; however, the Airline Transport Pilot Licence (ATPL) requirements for commercial pilots and The Train Driving Licences and Certificates Regulations 2010 for training and examination of train drivers detail principles which could be considered within an AV operator's safety case. Research has been done on how adults learn and how to develop effective and learner focused resources and delivery with measurable outputs in the form of formative and summative assessments (Petty, 2009). Similar to the processes in usability, delivery of training is iterative, with continual improvements based on feedback and the outputs form assessments and skills monitoring (Gravells, 2017).

There are a number of methods that could be used to measure the situational awareness of RAs (see 10.2). However, the methods used need to be considered in terms of their potential impact on the performance of a RA, especially when they are engaged in on-road operations. The method of monitoring the RA should not reduce their SA of the AV and its operating environment (Linkov & Vanžura, 2021).

### 5.7.2.3 Other parties in the event of an incident

The range of tasks that emergency services may need to perform at the scene of an incident are numerous, including deactivating the ADS, ensuring that the brakes are applied/released, accessing ADS components for response or recovery, isolating the EV battery, and accessing AV data for investigative purposes. An incident could also be non-reportable, which would require an exchange of insurance details with other road users in the event of a collision. The information requirements and needs of different actors would depend on the tasks that they need to perform and the behaviours that are required. For instance, the information requirements and behaviours of emergency staff treating those injured would be different from passengers trapped in a vehicle. This presents a dynamic complex system with different interacting parts. Additionally, there are elements of time and sequence which means that different actors may well require different informational inputs at different times depending on the unfolding of events within a changing environment.

The complexity of the components, the actors, interactions and feedback loops, makes this a complex system. Given the complexity of this, the implications for, for instance, vehicle design, HMIs or planning, need to be more fully analysed and understood. Therefore, the initial step in this process would be to have a much more in-depth and comprehensive understanding of the system, its components, and their interactions as a whole.

It is strongly recommended that commercial deployment organisations perform a behavioural hazard analysis<sup>10</sup> to understand the dynamic system and the interactions between its component parts. This would then form the basis for understanding how the informational needs of the different actors could be met, and be used to develop usability requirements and testing procedures (see Section 5.7, Principles of human factors design, for an overview).

The Department for Transport (DfT, 2019) requires trialling organisations speak with the road and enforcement authorities, develop engagement plans and have a data recorder fitted. The aim of the Code is to support cooperation between trialling organisations and those responsible for the management of traffic, infrastructure, law enforcement and other areas to support maximum road safety. It is for those carrying out trials to develop plans that are proportionate to the trial and vehicle under trial, as well as being sufficiently capable of capturing data for investigation purposes (DfT, 2019).

In line with this guidance for trials, the safety case for a commercial deployment should:

- Outline how the deployment aligns with legislation and regulation, and provide evidence of engagement with relevant bodies, authorities and other road users.
- Develop plans for police investigators and relevant organisations to readily and immediately access data relating to an incident in a way that maintains the forensic integrity, security, and the preservation of the data. This may include agreement with emergency services prior to trial activity, such as service level commitments for responding to incidents or requests for information.
- Ensure the safety case is regularly updated and continues to be assessed.

Additionally commercial deployment organisations should:

- Engage with authorities that could provide guidance on what to do in the event of a reportable incident and support with public communications and/or media coverage.
- Ensure any reportable incidents are communicated to the police.
- Ensure, depending on the specific incident, police and any other organisation relevant to an investigation are provided with access to relevant vehicle data.

---

<sup>10</sup> One technique that could be used is System-Theoretic Process Analysis (STPA) which supports the identification of points of in complex systems and the interactions between the components in the system (Leveson & Thomas, 2021).

An overview of the DfT's (2019) presentation of data requirements to emergency services includes that:

- The data should be intelligible and not require complex analysis or interpretation techniques.
- Where data is not immediately intelligible, it is expected that trialling and commercial deployment organisations will fully support investigators as part of any requests for access.
- If data is collected that enables individuals to be personally identified, this will amount to the processing of personal data under the Data Protection Act 2018.
- Data storage and use must comply with data protection legislation, including the requirements that the personal data is used fairly and lawfully, kept securely and for no longer than necessary.
- Those conducting remote-controlled trials are required to have real-time supervision of the vehicle and its surroundings. Such safety outcomes may be achieved through two-way, real-time communications links and full processes to deal with any failures. Those conducting remote-controlled trials are still required to have real-time supervision of the vehicle and its surroundings.
- In the event of an incident, such data should also be preserved in full. It is expected that responsible trialling and commercial deployment organisations will cooperate fully with the relevant authorities by providing access to any relevant data.
- Under Section 170 of the Road Traffic Act 1988, drivers have a duty to stop, report, and provide documents and information in the event of an incident. Section 170 covers a number of cases where the duty applies, such as personal injury to a person other than the driver, or where damage is caused to another vehicle, any property, or animals listed in the Road Traffic Act 1988.

For the purposes of information, it is critical to understand the different information requirements of the different actors involved in the incident, including the passengers, any other parties involved in the incident, other road users, emergency services, the operator, recovery services, the media and any RA. Furthermore, the informational needs of actors will be determined by the severity of the incident, impact of the incident on other road users, impact of the incident on passengers, impact of the incident on the integrity of the vehicle, environmental conditions and interactions between actors, to name but a few.

### 5.7.3 Recommendations

The recommendations for human factors performance requirements have been categorised to provide guidance for their implementation. The categories are as follows:

- **Vehicle design** – Recommendations to ensure minimum performance for the base vehicle (such as seats, brakes, etc.) as part of existing, or potentially new/amended, regulations.
- **Passenger safety** - Recommendations to ensure passenger safety that are not required by public service vehicle regulation, but would support good design if taken into consideration.
- **Other** - Recommendations that might be required depending on the scope of the assurance framework.
- **Existing Regulation** – Recommendations where Regulations or guidance already exist, either for AVs or in other domains.

Categorisation of the recommendations is based on subjective assessment therefore it is possible one or more could be categorised in other groups.

#### 5.7.3.1 Onboard occupants

##### Recommendations for supporting passengers' perceptions of safety

By law people have to be made aware that they are under surveillance on public transport (UK Government, 2012). Passengers should be reassured that the AV's safety and security is supported through recording CCTV footage and supervision by a human RA to deter anti-social behaviour and increase feelings of safety. Awareness of the surveillance could be done through visual formats, such



as pictograms, visible cameras, simple audio or visual messages. The recommendations are provided in Table 18.

Recommendation	Vehicle design	Passenger safety	Other	Existing Regulation
Provide passengers with in-vehicle information of what the AVs sensors are detecting during normal operating conditions.		●		
Provide information about trip status and time to board/disembark from the AV.	●			
<p>Communicate with passengers when a hard braking event occurs. This requires considerations about the procedures and responses to the different scenarios in which this could occur. Consider providing passengers with appropriate information to reassure them. It is suggested that procedures are developed to manage passenger behaviour during such an event. Consider feedback that would tell people:</p> <ul style="list-style-type: none"> <li>● What to do, e.g., remain seated</li> <li>● What is happening, e.g., the RA is providing inputs to support the system</li> <li>● What will happen, i.e., the journey will start shortly</li> <li>● How the vehicle will respond, i.e., the pod will be moving in 15 seconds.</li> </ul>	●			
It is strongly recommended that human factors and usability design processes are followed to ensure that the information meets the needs of a range of users and elicit the appropriate passenger responses	.			
It is recommended that passengers are made aware if a RA is monitoring the ADS.		●		
Ensure that passengers are aware that they are under surveillance.				●

Table 18: Recommendations for supporting passengers' perceptions of safety.

### Recommendations for passenger safety during emergencies

To deter passengers from misusing emergency interfaces, clear information and instructions should be provided to guide their use and highlight misuse penalties. In order for passengers to recognise and use emergency interfaces, such as a stop button and communications, it is recommended that these conform to existing good design principles used in other industries which allows for the transfer of knowledge from other domains. Table 19 provides examples of cross industry standards that could be used to inform in-vehicle passenger information. Recommendations for passenger safety during emergencies are provided in Table 20.

Context	Standard/Guide	Reference
London trams	Trams Standard (CR4000): Issue 4	(TfL, 2019)
Across London transport	Pictogram Standards: Issue 4	(TfL, 2009)
Across industries	Safety Signs and Signals	(HSA, 2015)

Table 19: Standards providing guidance for the display of information across different contexts.

Recommendation	Vehicle design	Passenger safety	Other	Existing Regulation
Provision of physical interfaces for passengers to instigate an MRM in the event of a medical emergency, acts of violence or terrorism.	•			
Provision of a communications system so that passengers could initiate contact with an assistant in a help/control centre.	•			
Responding appropriately to emergency situations is critical to the safety of passengers, manufacturers are strongly advised to use an iterative usability design process <sup>11</sup> and usability standards to verify and validate that the procedures, protocols and interfaces used are effective and efficient.			•	
Development of emergency protocols for the RA to follow in response to communications or an MRM. The protocols need to assist the RA to identify the problem, manage the behaviour of passengers, provide appropriate inputs to the ADS, and identify and request assistance from the emergency services, if necessary. The Department for Transport’s Best Practice Guide provides information on what to consider when planning for and responding to acts of violence and terrorism (DfT, 2018).			•	
Development and use of behaviour analysis software that could detect possible behaviour patterns, such as passenger movements in the pod that could alert the RA of an unfolding emergency.		•		

Table 20: Recommendations for passenger safety during emergencies.

<sup>11</sup> As outlined in ISO 9241- 210: Ergonomics of human-system interaction (ISO 9241-210, 2008).

Recommendations for accessibility and usability of information for passengers

Table 21 sets out recommendations for how information to passengers should be made accessible and useable.

Recommendation	Vehicle design	Passenger safety	Other	Existing Regulation
Mechanisms should be put in place to ensure that passengers are fully embarked/ disembarked before doors close; visual and audible cues could be provided to support this. An equivalent of this is the “doors closing” message given at railway stations. A button to keep the door open could also be provided, as is used in lifts.	•			
Mechanisms should be put in place to ensure that a trap and drag incident does not occur. If a door obstruction is detected, the RA should be alerted and ensure all doors are clear from obstruction before allowing the AV to set off. The RA should be able to communicate with people in or outside the vehicle to manage the situation.	•			
Where an AV is for seated passengers only, mechanisms should be put in place to identify someone standing or moving in the pod. This could include in-vehicle CCTV behaviour analysis.		•		
Before setting off, verbal or audio feedback could help passengers to understand when a pod/shuttle is full and prevent overcrowding or overloading.		•		
Where an AV is for seated passengers only, verbal or audio feedback could help passengers to understand the need to sit down before the vehicle moves off, or advise on the need for the engagement of a securing device for wheelchair users.		•		
Information about time to next stop could be displayed and an audio message used to state which stop the pod is coming to and which stop is next. An app and/or in-vehicle monitor could display information about point in journey, next stop and estimated times to arrival.		•		
From a safety perspective for AVs with provision for seating only, it is recommended that passengers are provided with prompts to remain seated until the vehicle comes to a full stop to reduce falls or trips during transit. The design needs to consider how wheelchair users will be able to safely use a pod.			•	

Where applicable, provisions should be made for young children to travel safely in the pod with an adult. Consider seating options to secure young children and the provision of places to store buggies securely.			•	
Where applicable, consideration should be given to how unsupervised/ unaccompanied children could travel securely in the AV. For instance, RAs could be alerted through age identification software when children are boarding a pod.		•		
It is also suggested that an assistant in a help/ control centre could communicate with passengers if passenger behaviour needs to be moderated.			•	
It should be ensured that passengers are aware that they are being monitored. Jansen et al. (2018) found that when people are aware that their behaviour is being monitored, they are more likely to moderate anti-social types of behaviour.		•		

*Table 21: Recommendations for the accessibility and usability of information for passengers.*

### 5.7.3.2 Recommendations for the provision of information to passengers

Providing information in a variety of accessible formats that meet the requirements and needs of a range of user groups could support the safe use of AVs for passengers. To fully understand these needs and how to meet them, it is strongly advised that behavioural and usability methods and processes are used. In terms of the use of information to support safe and secure travelling on an AV, some of the tasks that will need consideration are captured in Table 22.

Passengers' tasks	Recommendations related to information	Vehicle design	Passenger safety	Other	Existing Regulation
Planning the journey	Where and how to get help and support			•	
	Where and how to access the service			•	
	How to buy a ticket			•	
Waiting on the kerb for the AV	Where to wait safely		•		
	If there is capacity on the vehicle, for instance seating or a wheelchair space		•		
	When they can board		•		
	Confirmation that they are boarding the correct service			•	
Boarding	Where to embark		•		
	How to embark safely		•		
	What to do if there is a problem while trying to embark			•	
	Confirmation that they are embarking on the right service			•	
	Potential hazards to be aware off, comparable to the "mind the gap" message on the London Underground		•		
During the journey	Where seating is available (for a visually impaired person)		•		
	Which service it is			•	
	Journey time to next stop			•	
	Point in overall journey			•	
	What the next stop is			•	

	How to keep safe during the journey		•		
	What to do in an emergency	•			
Disembarking	How to disembark safely		•		
	Hazards during disembarking			•	

Table 22: A breakdown of the tasks passengers may need to perform for an LSAV journey, and the foreseeable information they may require.



### 5.7.3.3 Remote Assistant

#### Recommendations for retaining situational awareness and considerations for control room design

These recommendations are summarised in Table 23.

Recommendation	Vehicle design	Passenger safety	Other	Existing Regulation
<p><b>Stanton et al. (2001) summarise the following recommendations to maintain high SA, based on the work of Endsley (1995), when designing interfaces controlling any safety-critical system:</b></p> <ul style="list-style-type: none"> <li>○ <b>Reduce the requirement for people to make calculations.</b></li> <li>○ <b>Present data in a manner that makes level 2 SA (understanding) and level 3 SA (prediction) easier<sup>12</sup>.</b></li> <li>○ <b>Organise information in a manner that is consistent with the persons' goals.</b></li> <li>○ <b>Indicators of the current mode or status of the system can help cue the appropriate SA.</b></li> <li>○ <b>Critical cues should be provided to capture attention during critical events.</b></li> <li>○ <b>Global SA<sup>13</sup> is supported by providing an overview of the situation across the goals of the operator.</b></li> <li>○ <b>System-generated support for projection of future events and states will support level 3 SA.</b></li> </ul>		<ul style="list-style-type: none"> <li>•</li> </ul>		

<sup>12</sup> See Section 5.7.2.2 for an introduction to the three levels of SA.

<sup>13</sup> Global SA refers to situational awareness over every aspect of the task at hand and the environment.

<ul style="list-style-type: none"> <li>○ System design should be multi-modal and present data from different sources together, rather than sequentially, in order to support parallel processing of information.</li> </ul>				
<p>In a remote assistance context, RAs could be instructed to look for hazards after a period of low workload in order to regain SA, in line with the suggestion of using gamification to increase engagement and motivation.</p>		<ul style="list-style-type: none"> <li>•</li> </ul>		
<p>A reduction in discrepancies between the remote assistance conditions and those that would be experienced when driving in a vehicle (e.g., lighting conditions, perception of depth, headway).</p>		<ul style="list-style-type: none"> <li>•</li> </ul>		
<p>Alarms should warn the RA only when something crucial happens (e.g., ADS fault, network drop-out), to avoid displaying false warnings that lead to distrust (Linkov &amp; Vanžura, 2021).</p>			<ul style="list-style-type: none"> <li>•</li> </ul>	
<p>To improve interaction between an RA and the vehicle, and enhance SA, the following could be considered:</p> <ul style="list-style-type: none"> <li>○ Gamification of remote operational interfaces to promote sustained attention (Steinberger, Schroeter, &amp; Watling, 2017).</li> <li>○ Tactile feedback, for instance the use of haptic tables for navigation (Luz, et al., 2019) and alerts (Mohebbi, Gray, &amp; Tan, 2009).</li> <li>○ First person video and audio feedback to increase the experience of embodiment (Aymerich-Franch, Petit, Ganesh, &amp; Kheddar, 2017).</li> <li>○ Immersive interfaces that replicate real environmental conditions (Almeida, Menezes, &amp; Dias, 2020).</li> <li>• Head mounted displays (HMDs)<sup>14</sup> allow the Operators to observe the environment by moving their heads and are becoming increasingly popular when considering remote operation (Shen, et al., 2016)</li> </ul>		<ul style="list-style-type: none"> <li>•</li> </ul>		

<sup>14</sup> HMDs, as opposed to virtual reality, do not recreate the environment artificially, but they display it through a headset that controls the movement of a stereoscopic camera.

<ul style="list-style-type: none"><li>○ <b>The use of haptic input in the form of road sounds to provide extra situational awareness cues to aspects such as speed, possibly creating a more immersive experience (Mutzenich, Durant, Helman, &amp; Dalton, 2021).</b></li></ul>				
--	--	--	--	--

*Table 23: Recommendations for retaining situational awareness and considerations for control room design.*

### Recommendations for the training of Remote Assistants

These recommendations are summarised in Table 24.

Recommendation	Vehicle design	Passenger safety	Other	Existing Regulation
<p>The operators should (DfT, 2019):</p> <ul style="list-style-type: none"> <li>○ Ensure that RAs undergo continued development and training.</li> <li>○ Consider how to appropriately measure RA performance and availability.</li> <li>○ Ensure RAs have received the appropriate training.</li> <li>○ Develop robust procedures to ensure that RAs are sufficiently alert to perform their role and do not suffer fatigue.</li> <li>○ Have in place clear rules regarding RA behaviour and ensure that these are known and understood (such as alcohol or drug use).</li> <li>○ Take into account the impact of seeing a driverless vehicle on other road users.</li> </ul>				<ul style="list-style-type: none"> <li>•</li> </ul>
<p>Regarding transition between automated and manual (human-controlled) modes, the transition system should:</p> <ul style="list-style-type: none"> <li>○ Make the RA aware with an audible warning which may be accompanied by a visual warning in the event of a malfunction or failure of the system. operators should also consider the need for other methods of making an RA aware of a fault, such as haptic feedback.</li> <li>○ Ensure demands for assistance are audible, visible, and/or haptic as appropriate; operators should consider the practicalities of how an alert might work.</li> <li>○ Ensure that the RA is given a clear indication of what mode the vehicle is in (e.g. ADS enabled, standby).</li> </ul>				<ul style="list-style-type: none"> <li>•</li> </ul>

<ul style="list-style-type: none"> <li>○ Ensure that the RA is given sufficient time to gain situational awareness before making inputs that could be safety critical when necessary. Operators should consider potential hazards, and the parameters of the deployment area.</li> <li>○ Allow the RA to make inputs quickly <b>and easily</b> to the ADS if necessary. This should be developed and proven.</li> </ul>				
<p>Since the delivery of training is a critical aspect of how an RA will interact with the system, it is strongly suggested that the development of training material, delivery and evaluation of outputs are supported by professional trainers and evidence-based research in the field of adult education.</p>		•		

*Table 24: Recommendations for the training of Remote Assistants.*

The vehicle’s automated braking and steering systems, and other systems, should be designed such that in the event of failure the vehicle can achieve a minimal risk condition explained in earlier paragraphs, which may include manoeuvring to a safe(r) location.

### 5.7.3.4 Recommendations for information and data requirements in the event of an incident

Recommendations	Vehicle design	Passenger safety	Other	Existing Regulation
That a behaviour hazard analysis is conducted to fully understand, not only the foreseeable extent of tasks and actors involved, but also the interactions between actors. This provides an understanding of the failure points within and between the components of a dynamic system. This could support the provision of data and information to different actors.			•	
DfT (2019) provides guidance for the planning, development, and maintenance of procedures to be followed in the event of an incident. Coordination with emergency services and local authorities is critical to the development of safety and security procedures. It is strongly recommended that planning also involves coordination between different emergency and recovery services and the trialling or commercial deployment organisation.				•
That data and informational interfaces are developed using usability processes and principles or that they make use of existing formats, such as the Rescue Cards provided for electric vehicles. Consider what the impact of the information would be, and on whom; for instance, if it is for passengers trapped in a vehicle, the aim may be to reassure them and support them to remain calm.		•		
It is suggested that consideration is given to the informational needs regarding the assessment of an incident; for instance how an RA would be able to distinguish between a reportable and non-reportable incident, and how images would be gathered of any damage if it is not in the line of sight of the pod’s cameras.		•		
Procedures and HMIs should be tested and the various actors trained to ensure that they are fully aware of different scenarios and what is expected from them.		•		

Table 25: Recommendations for information and data requirements in the event of an incident.

### 5.7.4 Summary

Based on the recommendations outlined in Section 5.7.3, and the safety benefit and practicality of implementing them, it is considered that the priority recommendations in Table 26 are mandatory for any commercial deployment to follow in all cases. This is not to say that other recommendations in Section 5.7.3 should not be considered, but they could be considered as being lower priority and more flexible for implementation.

Group	Topic	Recommendations that should be mandatory for commercial AV deployment
<b>Passengers</b>	Supporting passengers' perceptions of safety	Provision of information about trip status and time to board/disembark from the AV.
		Communication with passengers when a hard braking event occurs and what will happen next.
		Informing passengers whether a RA is monitoring the ADS.
	Passenger safety during emergencies	Provision of information regarding what to do in an emergency.
		Provision of physical interfaces for passengers to instigate an MRM.
		Provision of a communications system so that passengers could initiate contact with an assistant in a help/control centre.
	Accessibility and usability of information for passengers	Mechanisms in place to ensure that passengers are fully embarked or disembarked before doors close.
		Mechanisms in place to alert a RA if a door obstruction is detected.
		Where an AV is for seated passengers only, communication with passengers on the need to sit down before the vehicle moves off.
<b>Other parties in the event of an incident</b>	Information and data requirements in the event of an incident	Development of data and information interfaces for the emergency services to utilise during emergency scenarios.

Table 26: Recommendations that should be mandatory for commercial LSAV development.

## 5.8 Safety of Machine Learning

In this section, guidance is provided on the activities required to be undertaken to assure the safety of machine learning (ML) components used in the vehicle. ML components consist of models, created using training data, that are used to make predictions or decisions without being explicitly programmed to do so. These activities are carried out in order to satisfy the ML safety case objectives described in Section 3.2. This section covers the following ML lifecycle stages as illustrated in Figure 29: ML safety assurance scoping, safety requirements elicitation, data management, model learning, model verification and model deployment.

From the overview of the process shown in Figure 29, it can be seen that the assurance activities discussed in this section are undertaken in parallel to the development of the ML component. The vehicle system safety requirements that relate to the ML component are taken as input, and a safety case for the ML component is created. The ML assurance process is iterative, so any stage could trigger the need to reconsider information generated or consumed by other stages. The ML assurance activities may therefore be performed multiple times throughout the development of the ML component. For example, verification activities may reveal that ML safety requirements are not met by the ML component under some conditions. Depending upon the nature of the findings, this may require that stages such as model learning or data management must be revisited, or even that the ML requirements themselves must be reconsidered.

For each stage, a safety argument pattern is presented that can be used to explain how, and the extent to which, evidence generated from carrying out assurance activities at that stage supports the relevant ML safety claims, explicitly highlighting key assumptions, trade-offs and uncertainties.

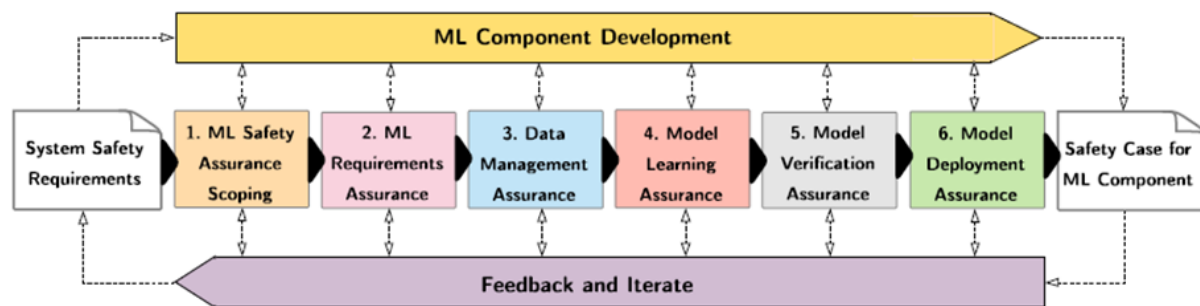


Figure 29 ML safety assurance activities

Further details of the ML assurance activities and the ML assurance case patterns can be found in AMLAS (2022), and will also be reflected in a “Base Document” to be produced as part of the BSI work on standards for AVs, supported by CCAV.

The argument patterns presented for each stage are represented using Goal Structuring Notation (GSN). GSN is a graphical notation for explicitly capturing safety arguments that is widely used in many industries for documenting safety cases. It represents the structure of the argument by showing how claims are broken down into subclaims, until eventually they can be supported by evidence. The strategies adopted, and the rationale (assumptions and justifications) can be captured, along with the context in which the goals are stated. The notation is summarised in Figure 30, and a detailed description of the notation is available in the publicly available GSN standard (GSN, 2018).



Key to GSN notation used to represent assurance argument:

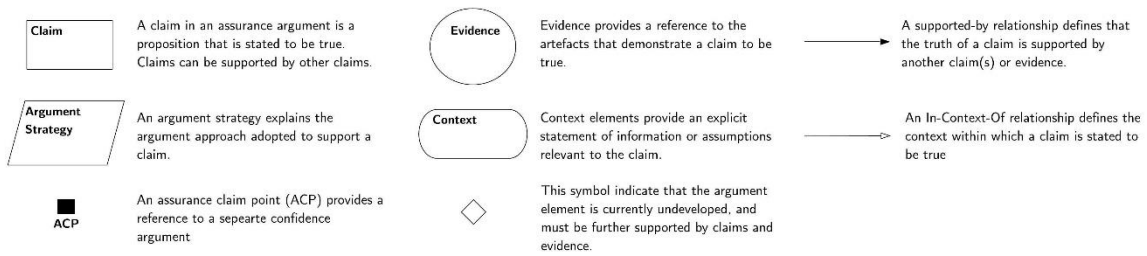


Figure 30 Key to GSN notation used in this re ML Safety Assurance Scoping

### 5.8.1 Inputs to the ML Assurance Process

The scope of the safety case for the ML component must be explicitly defined. This includes the system safety requirements arising from the vehicle-level safety assessment process, descriptions of the vehicle and its operating environment, a description of the role and scope of the ML component within the system of which it is part, and the inputs to which it is exposed from the vehicle sensors during operation. These artefacts are then used to determine the safety requirements that are allocated to the ML component.

The safety requirements allocated to the ML component are defined to control the risk of the identified contributions of the ML component to system hazards. This must take account of the defined system architecture and the operating environment. At this stage, the requirements that are identified from the vehicle safety process will not be defined specifically for ML, but instead will reflect the need for the component to perform safely with the vehicle, regardless of the technology later deployed. For example, consider an automated driving application in which a subsystem may be required to identify pedestrians at a crossing. A component within the perception pipeline may have a requirement of the form “When Ego vehicle is 50 metres from the crossing, the object detection component shall identify pedestrians that are on or close to the crossing in their correct position.”

The allocation of safety requirements must consider architectural features such as redundancy when allocating the safety requirements to the ML component. Where redundancy is provided by other, non-ML components, this may reduce the assurance burden on the ML component, which should be reflected in the allocated safety requirements.

The argument pattern relating to this stage is shown in Figure 31. This provides the top level of the safety case for the ML component that will link into the vehicle safety case. It can be seen in the figure that this argument links to the rest of the argument (via the ML safety requirements argument and the ML deployment argument). Further description of the argument pattern can be found in AMLAS (2022).

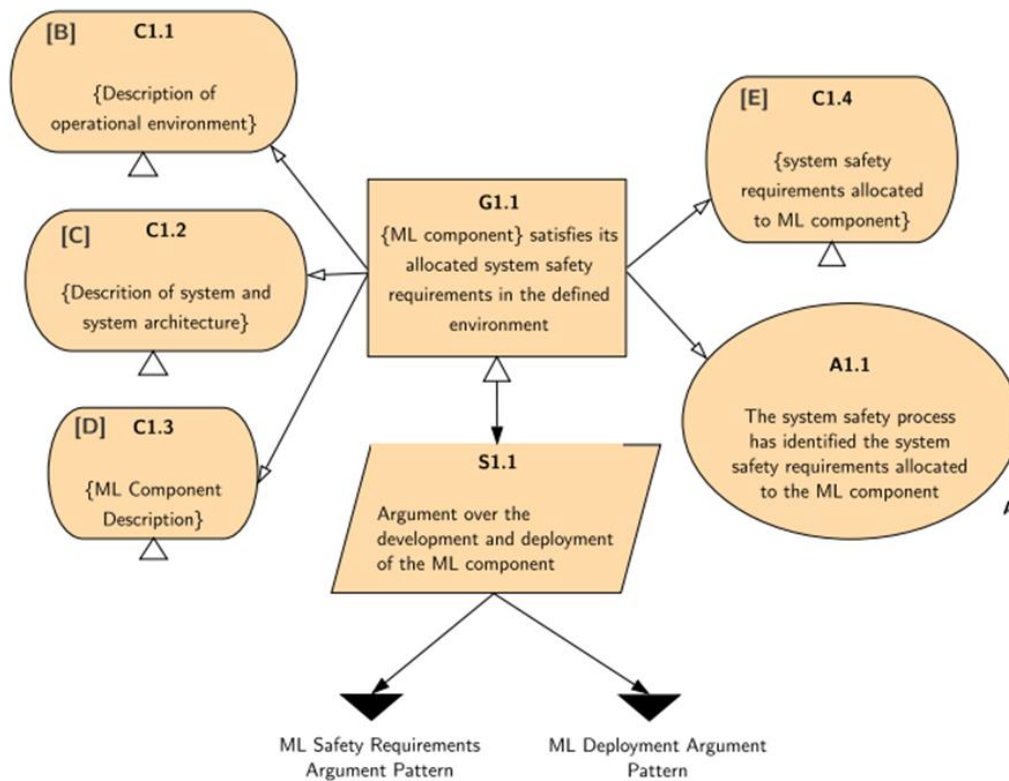


Figure 31 Argument Pattern for ML Safety Assurance Scoping

### 5.8.2 ML Safety Requirements Assurance

At this stage, assurance must be provided in the ML safety requirements. The scope of this stage is limited to the ML model, e.g. the mathematical representation of the neural network, that produces the intended output. This requires as input the system safety requirements allocated to the ML component at the previous stage (Section 5.8.1). ML safety requirements must be defined to control the risk of the identified contributions of the ML component to vehicle hazards, taking account of the defined system architecture and the operating environment of the vehicle. This requires translating complex real-world concepts and cognitive decisions into a format and a level of detail that is amenable to ML implementation and verification (Rahimi 2019).

The safety requirement allocated to the ML component discussed at Section 5.8.1 (for example, the concept of identifying a pedestrian) at the system level must be translated into something meaningful for the ML model. Again, using the pedestrian example, assume it is possible, using knowledge derived from the overall system architecture, to define the ML safety requirement as “all pedestrian bounding boxes produced shall be no more than 10% larger in any dimension than the minimum sized box capable of including the entirety of the pedestrian” (Gauerhof et al, 2020). This requirement helps to ensure that the position reported by the ML component is sufficiently close to the actual position of the pedestrian so as not to result in an unsafe decision being made by the vehicle when approaching the crossing.

In ML, requirements are often seen as implicitly encoded in the data. As such, the process of defining explicit requirements for ML components can be especially challenging. This particularly the case for open environments such as are encountered with AVs where there exists a potentially significant ‘semantic gap’ between the real-world and the defined ML requirements (Burton et al, 2020).

The ML safety requirements should always include requirements for *performance* and *robustness* of the ML model. From a safety assurance perspective:

- ML performance considers quantitative metrics, e.g. classification accuracy and error, whereas;

- ML robustness considers the performance of the model when the inputs encountered are different but similar to those present in the training data. Examples include environmental variability, e.g. flooded roads, and system-level variability, e.g. sensor failure.

The performance of a model can only be assessed with respect to measurable features of the ML model. An ML model does not generally allow for us to measure risk or safety directly. Hence safety measures must be translated to relevant ML performance and robustness measures such as true positive count against a test set or robustness to perturbations. Indeed, not all misclassifications have the same impact on safety, e.g. misclassifying a speed sign of 40 mph as 30 mph is less impactful than misclassifying the same sign as 70 mph.

There is rarely a single performance measurement that can be considered in isolation for an ML component. For example, for a classifier component, one may have to define a trade-off between false positives and false negatives. Over reliance on a single measure is likely to lead to systems which meet acceptance criteria but exhibit unintended behaviour (Amodei et al, 2016). As such, the ML performance safety requirements should focus on reduction/elimination of sources of harm while recognising the need to maintain acceptable overall performance (without which the system, though safe, will not be fit for purpose). Performance requirements may also be driven by constraints on computational power, e.g. the number of objects which can be tracked. This is covered in more detail in Section 5.8.6 on ML deployment.

One useful approach to defining *robustness* requirements is to consider the dimensions of variation which exist in the input space. These may include, for example:

- variation within the domain, e.g. differences between pedestrians of different age groups
- variation due to external factors, e.g. differences due to limitations of sensing technologies or effects of environmental phenomenon
- variation based on a knowledge of the technologies used and their inherent failure modes. For example, the deterioration in the performance of sensors over time may lead to changes in the inputs to the model during operation.

The example ML safety requirement example regarding pedestrian detection used previously may now be refined into performance and robustness requirements (Gauerhof et al, 2020). Example performance requirements may include:

- The ML component shall determine the position of the specified feature in each input frame within 5 pixels of actual position.
- The ML component shall identify the presence of any person present in the defined area with an accuracy of at least 0.93

Example robustness requirements may include:

- The ML component shall perform as required in the defined range of lighting conditions experienced during operation of the system.
- The ML component shall identify a person irrespective of their pose with respect to the camera.

The activity of developing the ML safety requirements is likely to identify implicit assumptions about the system or operating environment. Assumptions that are made should be made explicit either as part of the description of the system environment or through defining additional safety requirements. These are sometimes referred to as derived safety requirements.

The argument pattern relating to this stage is shown in Figure 32 Assurance Argument Pattern for ML Safety Requirements. Further description of the argument pattern can be found in (AMLAS 2022).

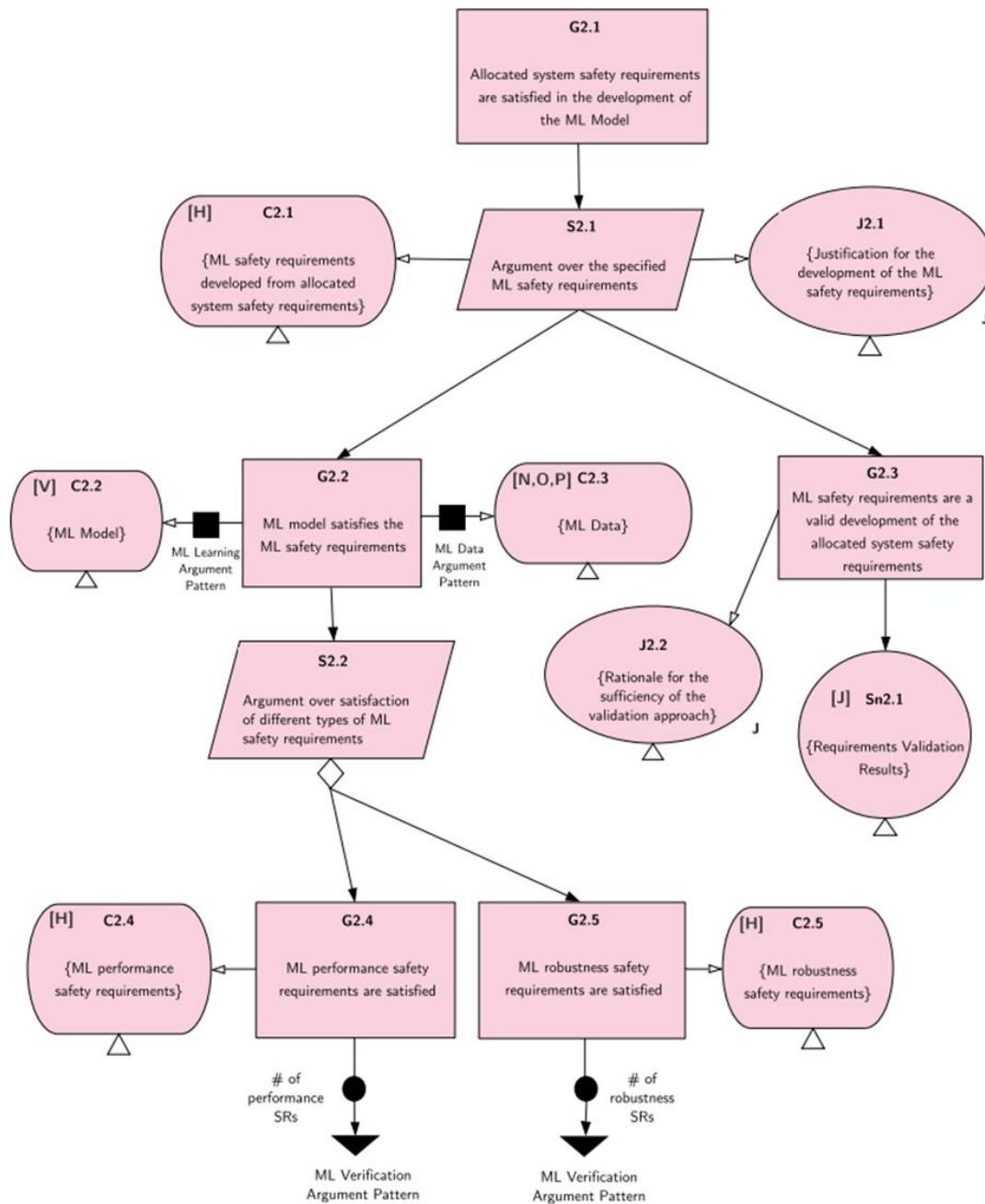


Figure 32 Assurance Argument Pattern for ML Safety Requirements

### 5.8.3 ML Data Management

Data plays a particularly important role in machine learning, because it is the data used to develop the model that predominantly determines the behaviour of the resulting model. It is therefore crucial that explicit data requirements are specified and that the data collected is demonstrated to meet those data requirements. This activity requires as input the ML safety requirements as discussed in Section 5.8.2, from which the data requirements are generated. Of particular interest in the development of data requirements are those safety requirements which pertain to the description of the vehicle operating environment.

### 5.8.3.1 ML Data Requirements

The ML data requirements must specify the characteristics that the data collected must have in order to ensure that a model meeting the ML safety requirements may be created. ML data requirements should include consideration of the *relevance*, *completeness*, *accuracy* and *balance* of the data (Ashmore et al, 2021). These requirements should explicitly state the assumptions made with respect to the operating environment and the data features that characterise the operating domain.

ML data requirements will often focus on specifying data which is necessary to ensure the robustness of the model in the context of the operating domain. This should relate to the dimensions of variation anticipated in the target operating domain (TOD) as described in the ML safety requirements.

ML data requirements relating to *relevance* must specify the extent to which the data must match the intended TOD into which the model is to be deployed. For example, for an ML component used for object detection on a vehicle, the following may be defined as an ML data requirement relating to relevance: “each data sample shall assume sensor positioning which is representative of that to be used on the vehicle”. This requirement is defined to ensure that images with a very low or very high viewpoint of the road (such as an aerial view) are not used in development.

ML data requirements relating to *completeness* must specify the extent to which the development data must be complete with respect to a set of measurable dimensions of the operating domain. This can be done through reference to the anticipated dimensions of variation stated in the ML safety requirements or defined by the TOD definition. For example, if the TOD for an automated vehicle indicates that the vehicle is to operate at all times of day and that the ML component should be robust to changing light levels. An ML data requirement for completeness may state: “Data samples should be gathered at all times of day and under the following light conditions: sunlight, cloud, rural with headlights and urban street lighting”.

ML data requirements need to include requirements that specify the required *accuracy* of the development data. For example, consider the previously defined ML safety requirement that all pedestrians should be identified within 50cm of their true position. Given that the pedestrians are not point masses but instead represented as coloured pixels in the image, an accuracy requirement must clearly specify the required position of the label including the positioning of labels for partially occluded objects. An example accuracy requirement may state that: “When labelling data samples, the position of all pedestrians shall be recorded as their extremity closest to the roadway”.

ML data requirements relating to *balance* should specify the required distribution of samples in the data sets. Consider a classifier which is designed to identify one of  $n$  classes. A data set which is balanced with respect to the classes would present an appropriate number of samples for each class. Note that this does not necessarily mean that an equal number of samples is required for each class; rare classes may require fewer samples in order to be balanced. More generally, however, balance may be considered with respect to certain features of interest, e.g. environmental conditions, gender, race etc. This means that a data set which is balanced with respect to the classes may present as biased when considering critical features of the data.

### 5.8.3.2 ML Data Generation

Data must be generated that meets the ML data requirements. This includes three separate datasets: development data, internal test data and verification data. Here we use the term development data to include training and validation data as it is normally referred to in the ML literature. Development data is used to create a model which is then tested by the development team using the internal test data. Once a model is deemed fit for release by the development team, only then is it exposed to the Verification data. The first two of these sets are for use in the development process (see Section 5.8.4) whilst verification set is used in model verification (see Section 5.8.5).

The generation of ML data will typically consider three sub-processes: collection, pre-processing and augmentation.

*Data collection* is undertaken to obtain data from sources that are available to the data collection team which sufficiently addresses the ML data requirements. This may involve reusing existing data sets where they are deemed appropriate for the context, or the collection of data from primary sources. It may be necessary to collect data from systems which are close to, but not identical to, the envisioned

system. Such compromises and restrictions should be stated explicitly, with a justification of why the data collected is still valid. For example, a vehicle gathering video data may be an experimental variant of the proposed vehicle where variation in vehicle dynamics is assumed to have no impact on the video data with respect to prediction of distance to leading vehicles.

Where it is impossible to gather real world samples, it is common to use simulators. These may be software or hardware in the loop. Where data is collected using such simulators, the configuration data should be recorded to allow for repeatable data collection and to support systematic verification and validation of the simulator within the operational context. Such simulators might need to be subjected to a separate assurance or approval process, such as discussed by Sargent (2010).

*Data Pre-processing* may be required to transform the collected data samples into data that can be consumed by the learning process. This may involve the addition of labels, normalisation of data, the removal of noise or the management of missing features. Pre-processing of data is common and is not necessarily used to compensate for failures in the data collection process. For example, pixel values for image data in the range [0,255] may be pre-processed to ensure they are represented instead as floating point values in the range [0,1].

A common pre-processing activity is the addition of labels to data. This is particularly important in supervised learning where the labels provide a baseline, or ground truth, against which learnt models can be assessed. Whilst labelling may be trivial in some contexts, this may not always be the case. In such cases, a process to ensure consistent labelling should be developed, documented and enacted.

*Data Augmentation* allows for the addition of data where it is infeasible to gather sufficient samples from the real world. This may occur when the real-world system does not yet exist or where collecting such data would be too dangerous or prohibitively expensive. In such cases, the data sets can be augmented with data which is either derived from existing samples or collected from systems which act as a proxy for the real world.

The field of computer vision provides methods to augment data using sophisticated models of environmental conditions (Zhang et al, 2017). For example, an image collected of an object under one controlled lighting condition can be augmented to produce many versions of that object under many different simulated lighting conditions.

Verification data is gathered with the aim of testing the models to breaking point. A different mindset is required for the team engaged with collecting data for verification. They are focused not on creating a model but finding realistic ways in which the model may fail when used in an operational system. Furthermore, the nature of ML is that any single sample may be included into the training set and a specific model found which is able to avoid the failure associated with the sample. However, this does not mean that the resultant model is robust to a more general class of failure to which the sample belongs. It is recommended therefore that the verification data is collected independently from the team developing the model and that the type and details of the verification data is not shared, to ensure the models generated are robust to the whole class of failures and not just to any specific examples present in the verification data.

Variation in the different components of the dataset may not be independent, and combinations of difficult situations are less likely to be included in the development dataset, which aims to represent normal operating behaviour. For example, consider the following situation where a vehicle using its high beam in foggy conditions on a rainy day where ice is present on the road and a vehicle is approaching on the incorrect side of the carriageway. This case, although within the operating domain of the vehicle, is unlikely to be found in the development dataset. A good verification team should try to focus on this type of challenging conditions and include such cases in the verification dataset.

### 5.8.3.3 ML Data Validation

ML data validation should be used to check that the three generated data sets are sufficient to meet the ML data requirements. Data validation should consider the relevance, completeness, and balance of the data sets.

Validation of data relevance should consider the gap between the samples obtained and the real-world environment in which the system is to be deployed. Validation should consider each of the sub-activities undertaken in data generation and provide a clear rationale for their use. Validation should demonstrate that context-specific features defined in the ML safety requirements are present in the collected

datasets. For example, for a pedestrian detection system for deployment on European roads, the images collected should include road furniture of types that would be found in the anticipated countries of deployment.

Validation of *data completeness* demonstrates that the collected data covers all the dimensions of variation stated in the ML safety requirements sufficiently. Given the combinatorial nature of input features, validation should seek to systematically identify areas where there are gaps in the coverage. Consider a system to identify road signs into 43 separate classes. Dimensions of variability are: weather, time of day and levels of partial occlusion up to 70%. Let us assume that we have categorised each dimension as:

- Time: early morning, mid-morning, noon, late afternoon, evening, late evening, night [7 classes]
- Weather: clear, rain light, rain heavy, fog light, fog heavy, snow light, snow heavy [7 classes]
- Occlusion (%): (0, 10, 20, 30, 40, 50, 60, 70) [8 classes]

Validation may show that there are samples for each of the  $43 * 7 * 7 * 8 = 16,856$  possible combinations. A systematic validation process will identify what the data sets are missing, e.g. no samples containing a 40mph sign in light rain with 50% occlusion early in the morning. Although for most practical systems, completeness is not possible, this process should provide evidence of those areas which are incomplete and why this is not problematic for assuring the resultant system.

Validation of *data balance* considers the distribution of samples in the data set. It is easiest to consider balance from a supervised classification perspective where the number of samples associated with each class is a key consideration. At the class level, assessing balance may be a simple case of counting the number of samples in each class; this approach becomes more complex, however, when considering combinational variation and that specific combinations are relatively rare. More generally, data validation should include statements regarding class balance and feature balance for supervised learning tasks. Certain classes may naturally be less common and, whilst techniques such as data augmentation may help, it may be difficult, or even impossible to obtain a truly balanced set of classes. In such cases, the imbalance should be noted and a justification recorded as part of the validation process. This justification must be realistic and not unduly impact the satisfaction of the safety requirements.

Validation of *data accuracy* should consider the extent to which the data samples, and metadata added to the datasets during pre-processing (e.g. labels), are representative of the ground truth associated with samples. Evidence supporting the accuracy of data may be gathered through a combination of the following:

- An analysis of the processes undertaken to collect data.
- Checking subsets of samples by expert users.
- Ensuring diversity of data sources to avoid systematic errors in the data sets.

The argument pattern relating to this stage is shown in Figure 33. Further description of the argument pattern can be found in (AMLAS, 2022).

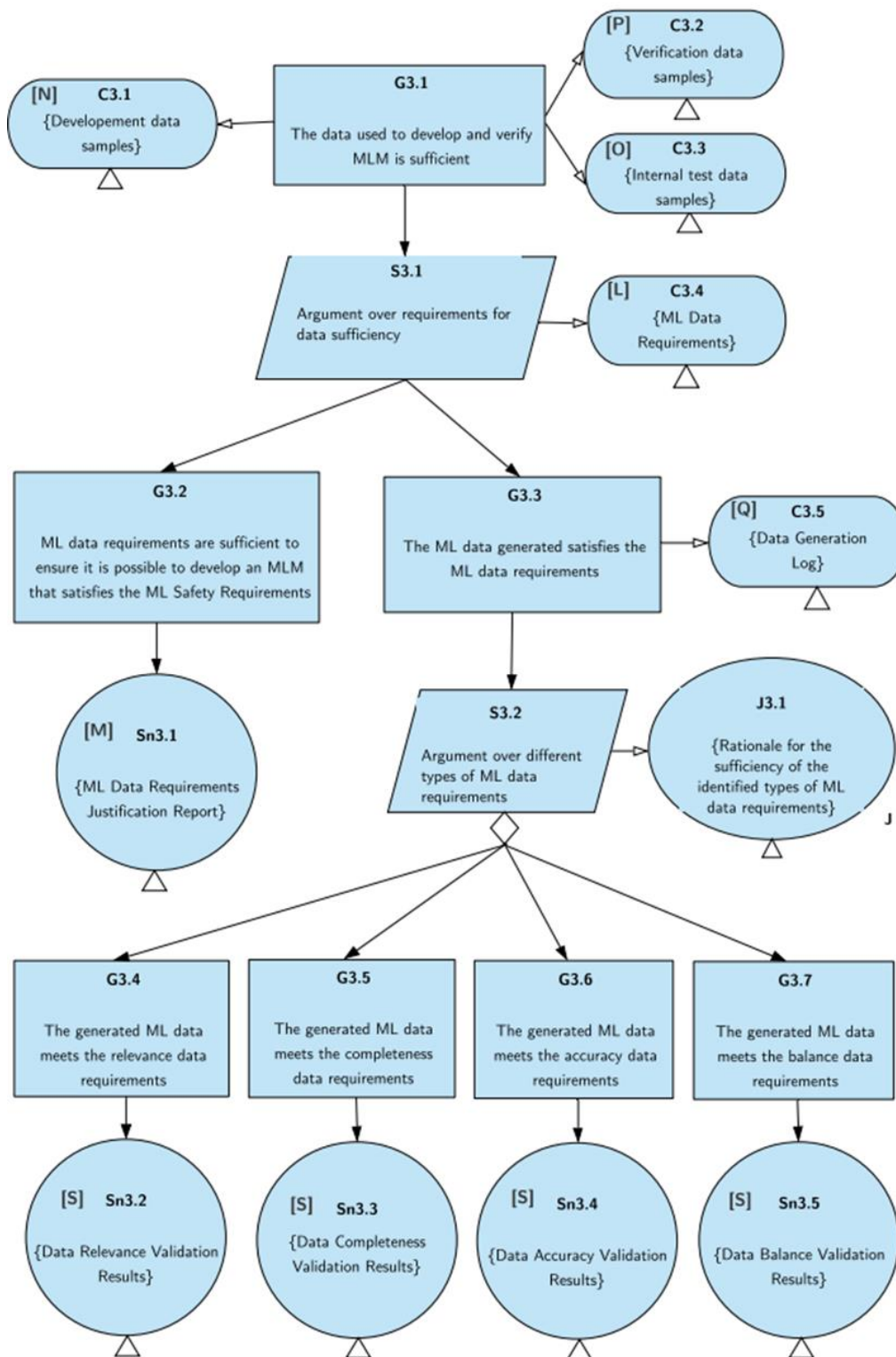


Figure 33 Assurance Argument Pattern for ML Data



## 5.8.4 Model Learning

The creation of an ML model starts with a decision as to the form of model that is most appropriate for the problem at hand and that will be most effective at satisfying the ML safety requirements. This decision may be based on expert knowledge and previous experience of best practice. For example, Deep Neural Networks (DNNs) can extract features from image data. A DNN which receives images as a frame from a video feed has been shown to be capable of identifying objects in a scene and may therefore be suitable for use in an automotive perception pipeline. Typically, numerous different candidate models of the selected type will be created from the development data by tuning the model parameters to create models that may satisfy the ML safety requirements.

A common problem that is encountered when creating a model is overfitting to training data. This happens when the model performs well using the development data but poorly when presented with data not seen before. This results from creating a model that focuses on maximising its performance for the given data, and consequently performance does not generalise. Techniques such as cross-validation (Anguita et al, 2012), leave-one-out (Cheng et al, 2017) and early stopping (Prechelt, 2012) can be used in handling the development data during the creation of the model to improve its generalisability and thus its ability to satisfy the ML safety requirements.

In creating an acceptable model, it is important to note that it is not only the performance of the model that matters. It is important to consider trade-offs between different properties, e.g. the cost of hardware and performance, performance and robustness, or sensitivity and specificity. Several measures are available to assess some of these trade-offs. For example, the Area Under the ROC (Receiver Operating Characteristic) Curves enable the trade-offs between false-positive and false-negative classifications to be evaluated (Fawcett, 2006).

Once a candidate model is created, it must be evaluated using the Internal test dataset to check that it is able to satisfy the ML safety requirements. The Internal test dataset must not have been used in creating the candidate model (allowing the development process to have a view of the internal test data is known as Data Leakage in ML).

The model development stage is iterative, and the model creation and model testing activities may be performed many times, creating different models which will be evaluated to find the best one. If it is not possible to create a model that meets the ML safety requirements when checked using the Internal test dataset, then the data management stage (Section 5.8.3) and/or the ML requirements stage (Section 5.8.2) should be revisited to create an acceptable model. Unlike traditional software testing, it is challenging to understand how an ML model can be changed to solve problems encountered during testing. For example, in testing the model, it might be found that the accuracy is lower than expected, indicating that the model fails to generalise beyond the development data. If an analysis of the images that were incorrectly classified showed that images with bright sunlight have a higher failure rate than other images in the test set, this might indicate a need to return to the data management stage and collect additional images of this class, to help train the model better in an attempt to decrease this mode of failure.

A model is selected from the valid candidate models that have been created. The selected model should be the one which best meets the different, potentially conflicting, requirements that exist. This is a multi-objective optimisation problem where there could be multiple models on the pareto-front and it is important to select the best threshold to satisfy our requirements. This is illustrated by the following example: a model to be deployed in a perception pipeline classifies objects into one of ten classes. A set of ML safety requirements are defined in terms of the minimum accuracy for each class. The model development process returns five models, each of which has accuracy greater than this minimum, but with each performing better with respect to a different particular class. Under such conditions, choosing the 'best' model requires the user to make a trade-off between class accuracies. Furthermore, this trade-off may vary as the context changes, e.g. from a rural to urban environment.

The argument pattern relating to this stage is shown in Figure 34. Further description of the argument pattern can be found in (AMLAS, 2022).

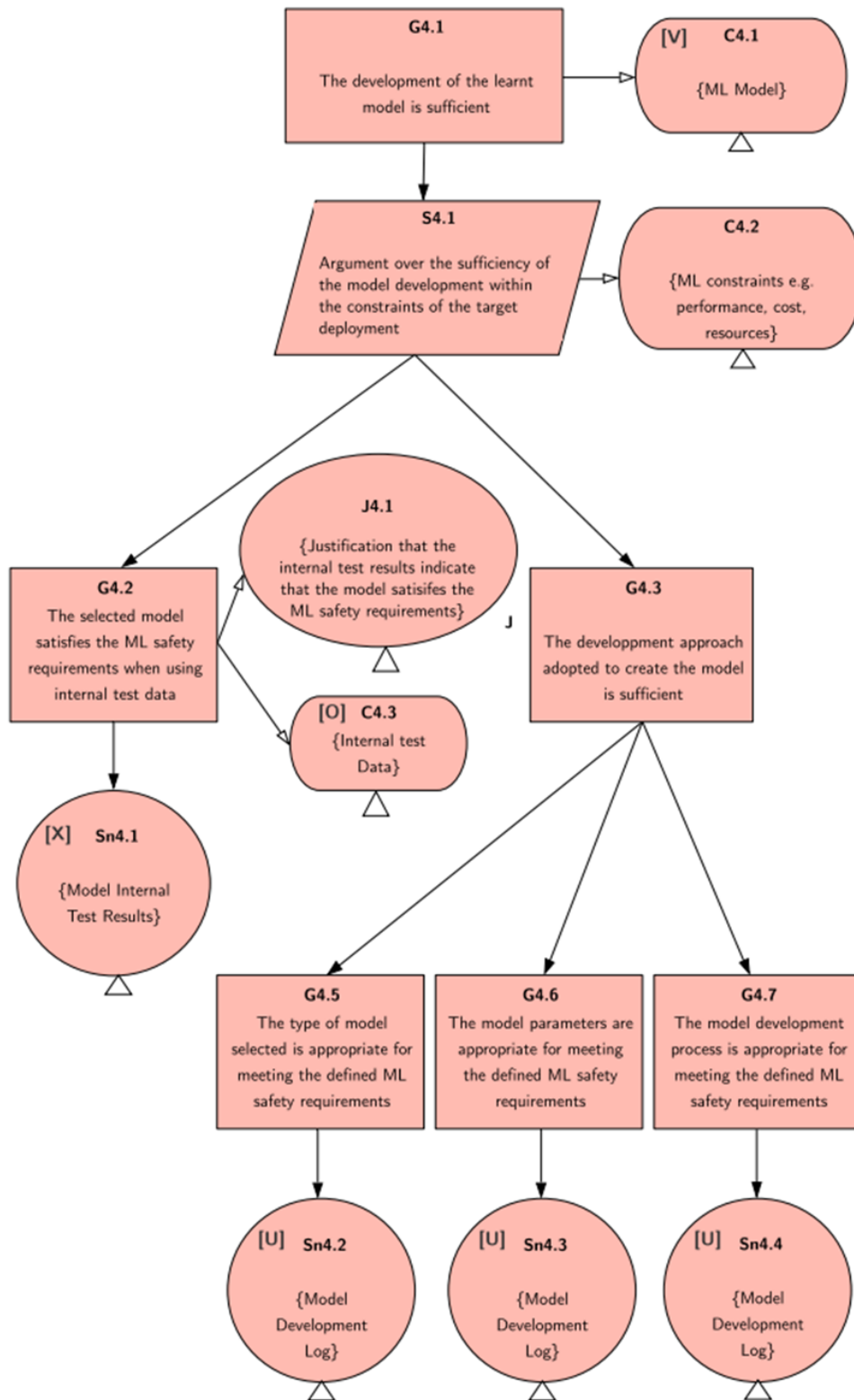


Figure 34 Assurance Argument Pattern for Model Learning

## 5.8.5 Model Verification

This stage requires as input the ML safety requirements, the verification data and the machine learnt model. Model verification may consist of two sub-activities: *test-based verification* and *formal verification*. For every ML safety requirement, at least one verification activity must be undertaken. All verification activities need to be sufficiently independent from the development activities.

One of the aims of model verification is to show that the performance of the model with respect to ML safety requirements is maintained when the model is subjected to inputs not present in the development data. A model which continues to perform when presented with data not included in the development set is known in the ML community as generalisable. Failing to generalise can be due to a lack of feature coverage in the development data or a lack of robustness to perturbations in the inputs to the model that may be considered to be noise.

### 5.8.5.1 Test-based Verification

Test-based verification utilises the verification dataset to demonstrate that the model works for cases not present in the development datasets. In particular, those safety requirements associated with ensuring the robustness of models should be evaluated on the independent verification data set. The aim is to show that the performance is maintained in the presence of adverse conditions or signal perturbations. The test team should examine those cases which lie on boundaries, or which are known to be problematic within the model deployment context.

If the results of applying the verification dataset to the model does not clearly satisfy the safety requirements, it may then be necessary to augment the verification data set and perform further tests which then give a clearer answer as to whether the requirements have been met. This is illustrated by the following example: a neural network which predicts the stopping distance of a vehicle for given environmental conditions is tested across a range of values (temperature, humidity, precipitation). The results show that the model operates safely at all values except three distinct points in the range. Further verification data may need to be collected around these points so that the particular conditions leading to the safety violations can be determined.

It is important to ensure that the verification dataset is not made available to the development team, since if they were to have sight of the verification data, they could utilise techniques at development time which circumvent problematic samples in the verification dataset rather than creating a better model that can tackle the problem of generalisation.

The ML verification process should evaluate test completeness with respect to the dimensions of variability outlined in the ML safety requirements. This is directly related to the desire for data completeness outlined in Section 5.8.3. For example, since it is known that material on a camera lens can lead to blurring of the image, it is possible to make use of 'contextual mutators' (Pezzementi et al, 2018) to assess the robustness of a neural network with respect to levels of blur. In this way, the level of blur which can be accommodated can be assessed and related to measures that are meaningful in the vehicle operating context.

### 5.8.5.2 Formal Verification

Formal verification uses mathematical techniques to prove that the learnt model satisfies formally-specified properties derived from the ML safety requirements. When formal verification is applied, counter-examples are typically created which demonstrate the properties that are violated. In some cases, these may be used to inform further iterations of requirements specification, data management or model learning.

The formally-specified properties must be a sufficient representation of the ML safety requirements in the context of the defined operating environment. An explicit justification must be provided for the sufficiency of the translation to formal properties. The formal models that are used for verification will require assumptions and abstractions to be made, both with respect to the ML model itself, and with respect to the operating environment. The validity of the formal model must therefore be demonstrated.

Having undertaken verification activities, ML verification evidence should be collated and reported in terms which are meaningful to any safety assessors with respect to the ML safety requirements and the

operating environment. The verification evidence must be comprehensive and clearly demonstrate coverage with respect to the dimensions of variability, and combinations thereof, relevant to the ML safety requirements.

One example of a test-based verification technique is Deep Road (Zhang et al, 2018), which utilises Generative Adversarial Network (GAN) based techniques to synthesize realistic and diverse driving scenarios in order to find inconsistencies in automated driving systems. The evidence should enumerate the scenarios examined and the results of the model when presented with these samples, as well as the ground truth labels.

The argument pattern relating to this stage is shown in Figure 35. Further description of the argument pattern can be found in (AMLAS, 2022).

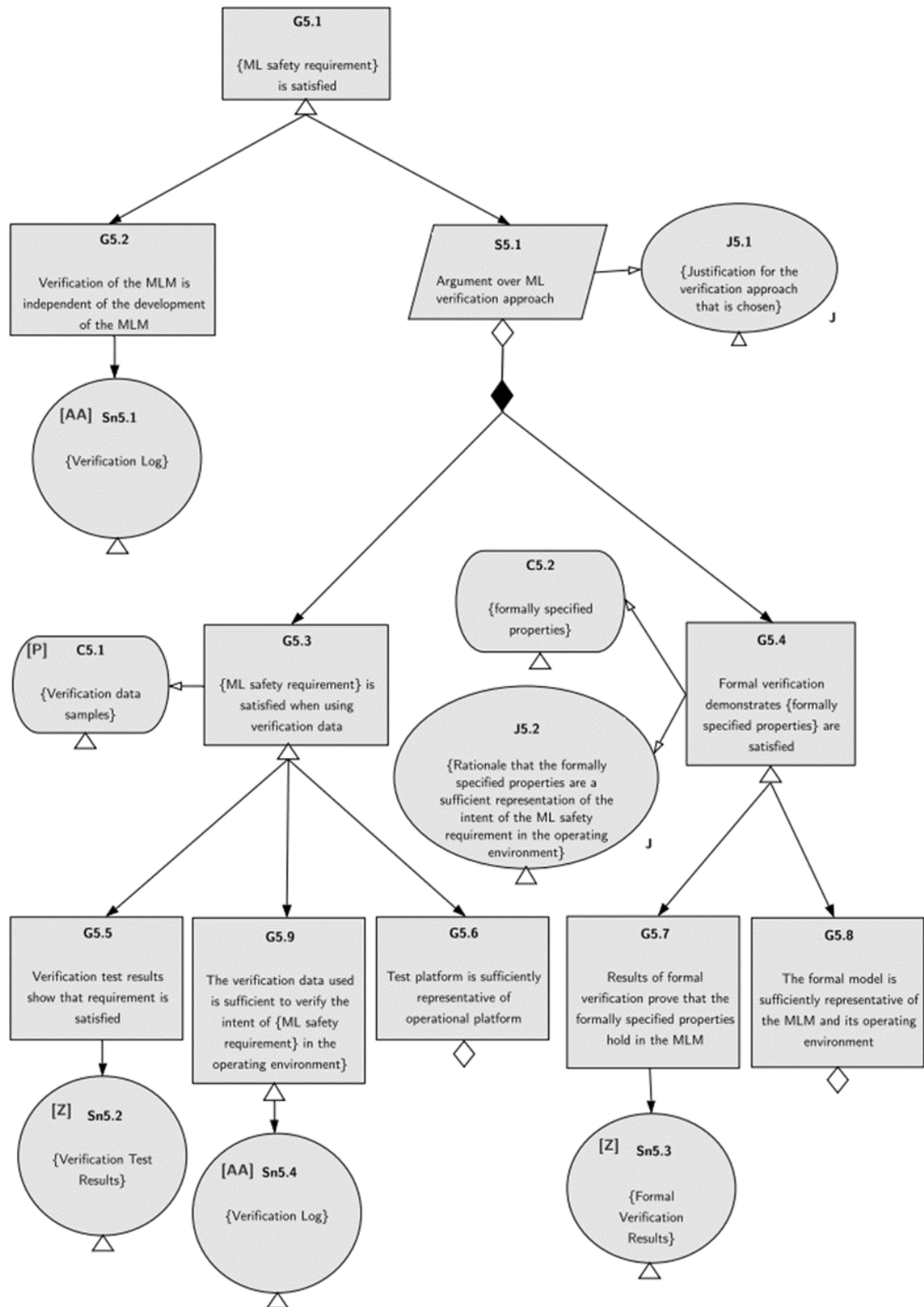


Figure 35 Assurance Argument Pattern for ML Verification

## 5.8.6 Model Deployment

This stage integrates the machine learnt component into the target vehicle in such a manner that the vehicle satisfies the allocated system safety requirements. The component should be integrated in the pipeline linking its inputs and outputs to other system components. It must then be demonstrated that the allocated system safety requirements are still satisfied during operation of the vehicle in the target operating environment. This process should be followed not only for initial deployment of the component, but also for any subsequent deployment if the component is updated.

The ML Model needs to be deployed onto the intended hardware platform and integrated into the broader vehicle systems of which it is a part. Deploying the component may be a multi-stage process in which the component is first deployed to computational hardware which is then integrated at a subsystem level before being integrating with the final hardware platform of the vehicle. The deployment process will include connecting the component's inputs to sensors and providing its output to the wider system. This activity takes as inputs the system safety requirements, the environment description, the system description and the ML model defined in the previous stages, and integrates the model into the vehicle.

The development of the ML model is undertaken in the context of assumptions that are made about the vehicle to which the ML model will be integrated and the operating environment of that vehicle. This will include key assumptions that, if they do not hold during operation of the system, may result in the ML model not behaving in the manner expected from the preceding development and verification activities. For example, when an ML component used for object classification is developed, there may be an assumption that the component will only be used in good lighting conditions. This may be based on the capabilities of the sensors, historic use cases, and the data from which the component is trained and verified. It is crucial to recognise and record that this is a key assumption upon which the assurance of the ML component is based. If a system containing the component is subsequently used at low light levels, then the classification generated by the ML component may not meet its safety requirements. When considering violations of assumptions, this should be linked to the system safety analysis process to identify the impact on system hazards and associated risks.

Measures should be put in place to monitor and check validity of the key system and environmental assumptions throughout the operation of the system. Mechanisms should also be put in place to mitigate the risk posed if any of the assumptions are violated. For example, consider that there is an assumption that the ML component for pedestrian detection deployed in an AV will be used only in daylight conditions. The system monitors the light levels. If the level of light drops below a level defined in the operating environment description, then the vehicle may be required to perform a pre-determined action such as a minimal risk manoeuvre (MRM) or switch to a degraded performance mode. Further guidance on the deployment of components to automated systems may be found in SASWG (2022) and Ashmore et al (2021).

There will always be some level of uncertainty associated with any ML model. This inherent uncertainty can lead to erroneous outputs from the model. The system must monitor the outputs from the ML model during operation, as well as the model's internal states, in order to identify when erroneous behaviour occurs. As well as considering how the system can tolerate erroneous outputs from the ML model, the integration activity should consider erroneous inputs to the model. These may arise from noise and uncertainties in other system components; because of the complexity of the operating environment; or due to adversarial behaviours. For example, the occlusion of a pedestrian in an image due to other objects in the environment may mean that for a (brief) period a pedestrian is not detected. Humans know that in the real world, a pedestrian does not disappear, so the complete system can use this knowledge to ignore and compensate for non-detections of previously identified pedestrians that last for a small number of frames. The system will then expect them to reappear.

When integrating the model into the system, the suitability of the target hardware platform should be considered. During the development of the model, assumptions are made about the target hardware and the validity of those assumptions should be checked during integration. If the target hardware is unsuitable for the ML model, a new model may need to be developed.

The system in which the ML model is deployed should be designed such that the system maintains an acceptable level of safety even in the face of the erroneous outputs provided by the model. For example, an ML model for pedestrian detection deployed in a self-driving car may have a performance requirement where a minimum percentage accuracy must be met. Due to uncertainty in the model, this

performance may not be achieved for every frame. The model assesses images derived from consecutive frames obtained from a camera. The presence of a pedestrian is determined by considering the result in the majority of the frames rather than a single frame. In this way the system compensates for the possible error of the model in predicting a pedestrian in any frame.

Once the ML model has been integrated into the wider system, the integration needs to be tested to check that the system safety requirements are satisfied. The target system containing the integrated ML component should be tested in a controlled setting to allow for safe evaluation of the system. This controlled setting may include additional controls, monitoring, or the use of simulation of real-world scenarios. System testing is discussed in more detail in Section 5.9.

The argument pattern relating to this stage is shown in

Figure 36

Figure 36. Further description of the argument pattern can be found in (AMLAS, 2022).

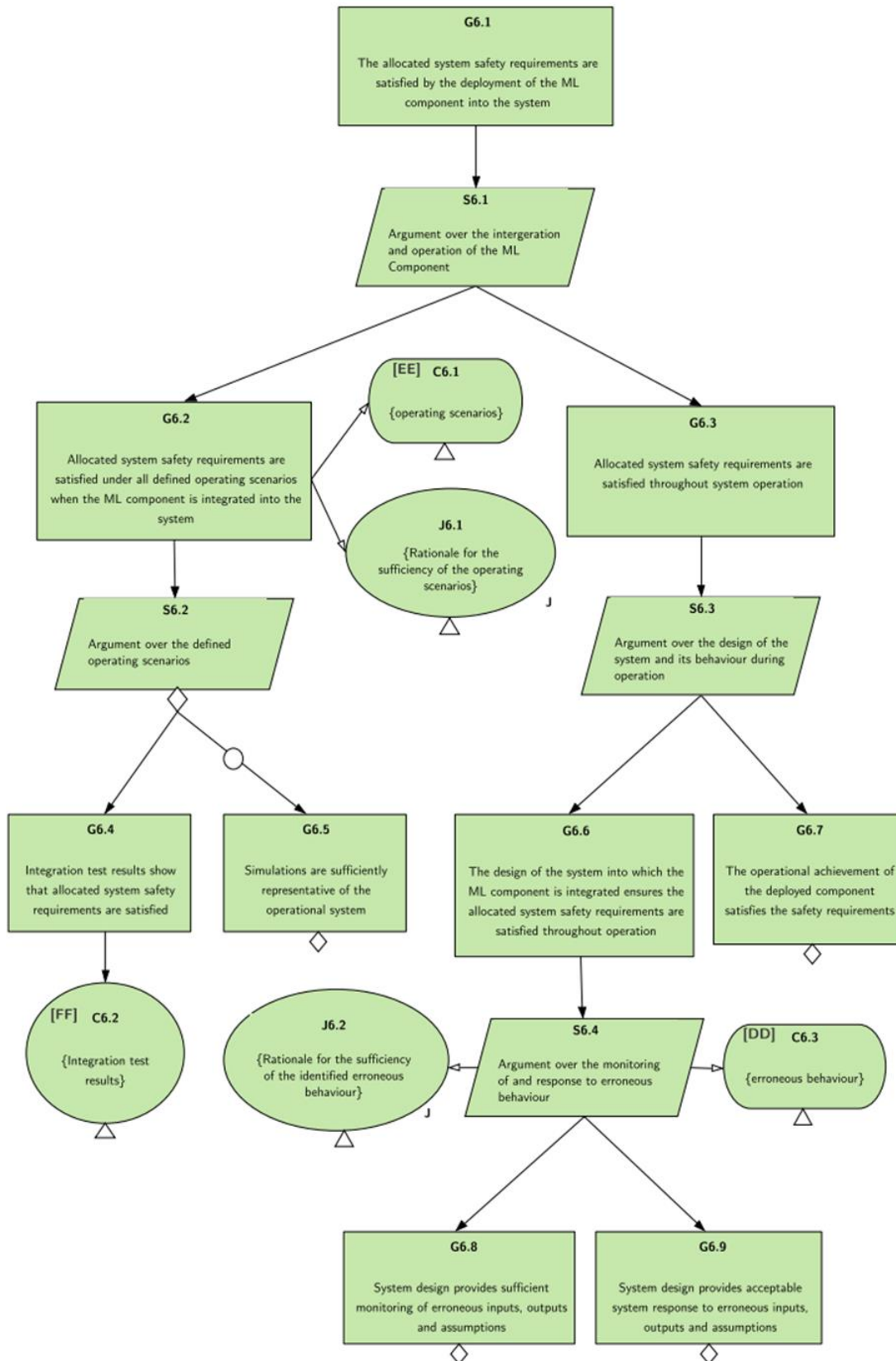


Figure 36 Assurance Argument Pattern for ML Model Deployment



## 5.8.7 Safe Use of ML in low-speed automated vehicles

### 5.8.7.1 Criteria for determining the permissible use of ML in low-speed automated vehicles

The use of training data in place of a detailed specification of the required behaviour, uncertainty in the results of the trained models and the lack of explainability in the calculations lead to specific challenges in the safety assurance of ML-based functionality. Reference to machine learning within safety standards such as ISO/PAS 21448, ISO TR 4804, ISO TS 5084 and UL4600 have so far focused on the class of techniques known as supervised learning, where the trained function is validated before deployment. The initial version of ISO PAS 8800 Road vehicles – Safety and AI, currently under development within ISO TC22/SC32/WG14, will focus on assurance arguments for supervised machine learning techniques which are trained and validated offline.

ISO/IEC TR 5469 Functional Safety and AI Systems is a technical report, currently under development, which focuses on generic issues related to functional safety and artificial intelligence. This work is exploring several concepts relevant to the regulatory requirements for ML in automated vehicles (and will thus be widely referenced within the work in ISO PAS 8800). One key consideration, when discussing the appropriate use of ML within safety-related systems, is the concept of ‘systematic problem complexity’, which can be defined as the difficulty to achieve completeness of all necessary aspects and parameters for the target function and operating scenarios of the AI/ML model. This implies that some tasks, such as the recognition of handwritten numeric digits in black and white images, are inherently simpler to achieve than others (for example detecting all relevant pedestrians within a crowded inner-city environment under all possible weather and lighting conditions).

This complexity manifests itself in several ways related to the Semantic Gap (Burton et al, 2020):

- **Specification insufficiencies:** This includes uncertainty in the definition of appropriate safety acceptance criteria and the definition of acceptably safe behaviour in all situations that can reasonably be anticipated within the target ODD and TOD. This also includes the ability to postulate a sufficiently complete model of the ODD and TOD, which can be used to reason about the completeness of trained and test data.
- **Technical uncertainty and implementation insufficiencies:** The more complex the target function, the greater the probability that the trained model will lead to generalization errors in potentially critical scenarios. This can be due to a lack of representative training data as well as underlying properties and parameterisation of the ML techniques themselves (such as robustness and prediction uncertainties in Deep Neural Networks). Acceptable bounds on these uncertainties must be both defined in the specification and accounted for in the design of the system.
- **Assurance uncertainty:** The systematic problem complexity eventually leads to uncertainties within the assurance process. This is related, on the one hand, to the potentially unclear definition of the required safety properties of the function, but also to a lack of confidence in the created safety evidence for arguing that these properties are met (Burton et al, 2019).

An example of a safety-related vehicle function with systematic problem complexity that could be considered low enough that an adequate safety assurance argument could be developed has been documented by Burton et al (2021).

Ongoing work within ISO/IEC TR 5469 is also considering various usage classes for ML in safety-related systems and include the following categories. Note that the definitions below do not reflect the exact wording of ISO TR 5469, which is yet to be fully finalised and released.

- Usage class A: AI/ML is not part of the safety function and can have no impact on safety due to sufficient segregation and behaviour control.
- Usage class B: AI/ML is used within the development of the safety-related system
- Usage class C: AI/ML is used within the safety-relevant system and has an impact on the safety-relevant decision-making function of the system

This allows for a differentiation in the safety assurance requirements on the function, dependent on its context of use. Furthermore, a differentiation will be made between classes of technologies (and their context of use) as follows:

- Technology class 1: The AI/ML technology can be developed and reviewed using existing functional safety standards and methods.
- Technology class 2: The AI/ML technology cannot be developed and reviewed using existing functional safety standards, but it is possible to identify a set of available methods and techniques for satisfying the safety-related properties of the function.
- Technology class 3: The AI/ML technology cannot be developed and reviewed using existing functional safety methods and it is also not possible to identify a set of available methods and techniques satisfying the properties.

The question of whether ML should be permitted for use in automated driving related functions is therefore very closely related to the systematic problem complexity of the task which the ML-based functions should implement, the context of their use and the availability of standards and best practice for assuring the safety properties of the function. A reflection on these factors is therefore a pre-requisite to the use of ML in automated driving, and can be used to evaluate the appropriateness of use of ML for safety-related functions. A recommendation on how these factors could be used to determine whether the use of ML should be permitted in automated driving applications is summarised in Table 27. Note that the boundary between technology classes 2 and 3 is not “fixed” and will change as, for example, verification technologies change. Nonetheless, the distinction remains valid.

Technology class Usage class	1: Existing standards can be used	2: No existing standards, alternative methods can be identified	3: No alternative methods can be identified
A: Not part of the safety function	No additional regulatory safety requirements required. However, the EU AI regulation act should be considered.		
B: Used in the development of the safety-related system	A conformance to relevant safety standards shall be demonstrated. E.g. ISO 26262-8: Confidence in use of software tools.  The results of the development activities performed using AI/ML shall be confirmed using alternative means (e.g. review, test) and documented in the assurance case. <b>No automated decision making that may impact the safety-related function of the system should be permitted.</b>	Alternative evidence shall be presented in the form of an assurance case.  The results of the development activities performed using AI/ML shall be confirmed using alternative means (e.g. review, test) and documented in the assurance case. <b>No automated decision making that may impact the safety-related function of the system should be permitted.</b>	Strictly prohibited, unless the system architecture enables sufficient safety assurance to be obtained via other means, e.g. by assessment of safety monitors implemented by conventional means.
C: Has a direct impact on the safety-related decision making of the system	A conformance to relevant safety standards shall be demonstrated. E.g. ISO 26262-6	Alternative evidence shall be presented in the form of an assurance case. <b>For tasks with high systematic problem complexity, a single point failure of the ML-based function shall not be permitted to lead to a hazardous event. For example, redundant sensing paths and non-AI algorithms shall be used in addition to the AI function.</b>	<b>Strictly prohibited.</b>

Table 27: Recommendations for the use of ML in automated driving.

In this section, these considerations are explored in more detail by providing illustrative examples of how the guidance described above could be applied to typical use cases and state-of-the-art ML techniques and assurance techniques. The following use cases are evaluated:

- Use Case 1: Supervised learning, e.g. Neural Networks for object recognition as part of the automated driving function, trained and validated off-line before deployment.
- Use Case 2: Reinforcement learning for driving policy optimisation as part of the automated driving function, including continuous optimisation within the vehicle post-deployment.
- Use Case 3: Off-line analysis of driving data for identifying critical scenarios to be considered during development and test of the vehicle.

The purpose of this section of the document is to provide reflection upon the conditions under which the use of machine learning within low-speed automated vehicles can be justified, by evaluating the challenges in fulfilling the assurance objectives outlined above in sections 5.8.1 to 5.8.6.

### 5.8.7.2 Supervised learning for object recognition

ML, and in particular the use of Deep Neural Networks (DNNs), is currently widely considered as the most promising technology for implementing camera-based detection of critical objects such as other vehicles and pedestrians within the path of a vehicle. A major focus of ML research at the intersection with safety assurance has been on improving and demonstrating the performance of DNNs for visual perception tasks, and significant progress is being made. As an example of recent work in this area, see the German publicly funded research project KI-Absicherung (2022).

An abstract representation of an ML-based perception component is shown in Figure 37. The subject of the safety assurance activities discussed here is encapsulated within the *ML component*, which consists of some pre-processing (e.g. normalisation of image contrast, scaling, etc.), the trained DNN itself that performs object detection and classification, and some post-processing. The post-processing can include, for example, non-maximum suppression to combine candidate bounding boxes of detected objects, as well as other means of reducing residual errors such as out-of-distribution detection. The DNN itself is trained and validated prior to deployment based on a set of training data and model parameters.

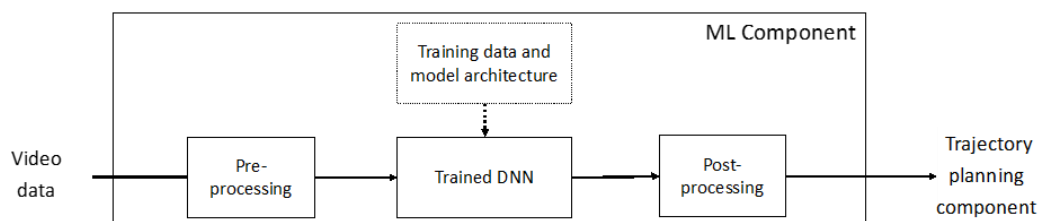


Figure 37 Abstract representation of a DNN-based perception component

An example requirement allocated to the ML component could be formulated as follows:

- Each pedestrian within the *critical range* is correctly detected within any sequence of  $N$  images with a *true positive rate*, *vertical* and *horizontal deviation* from ground truth sufficient to avoid collisions.

Where parameters which must be quantitatively defined specific to the application and context are highlighted in *italics*.

The safety assurance task can be formulated as arguing that, for all inputs that fulfil some set of reasonable assumptions on the operating domain and system context, the output of the ML component must fulfil a set of conditions defined by the safety requirements. Apart from formal verification techniques that are limited to comparatively small models and low dimensional inputs, it is not feasible to “prove” that these conditions hold for all possible inputs. Therefore, the claim will need to be *inferred* in an inductive manner based on evidence that is collected about the design and performance of the ML system, which is the inherent nature of most forms of safety assurance. This leads to the concept of quantitative acceptance criteria and validation targets as proposed by ISO/PAS 21448 to define when

the ML system can be considered “acceptably safe”, a concept further explored within Sections 5.9 and 5.2.

**Summary of the safety lifecycle and scoping of the assurance activities:**

Machine learning is based on statistical modelling techniques, whilst the properties of the environment (triggering conditions) that can lead to failures can also often only be reasoned about in a probabilistic manner due to the complexity of the operating domain and lack of sufficient environmental models. It should therefore come as no surprise that, unlike previous approaches for traditional software-based systems, the safety assurance of machine learning will require statistical arguments regarding the residual failure rates of the system. The assurance activities must ensure that residual errors in the ML component do not lead to an intolerable level of risk of safety-related failures of the control of the automated vehicle.

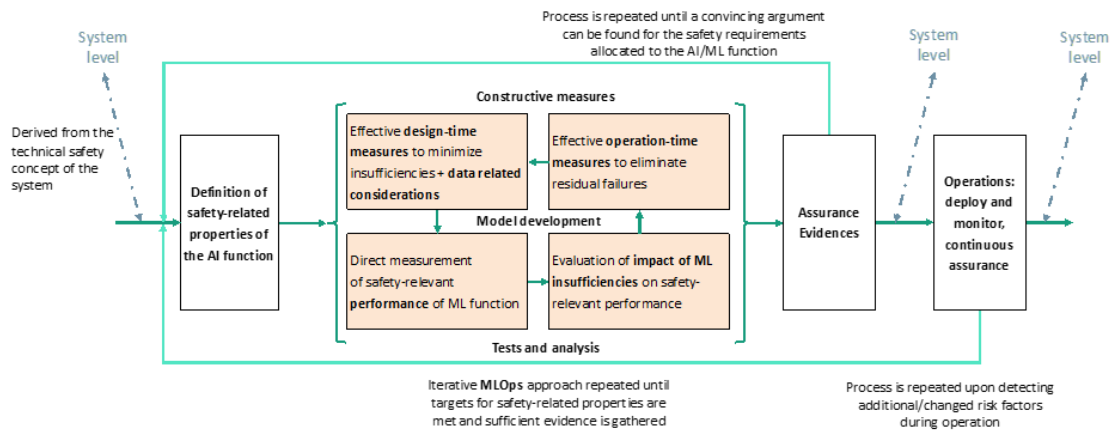


Figure 38 Safety lifecycle for a ML component based on supervised learning

Figure 38 shows the iterative nature of the ML development process, and shows how developments can reach the stage where there is sufficient evidence to support the safety arguments defined by the patterns described in this document. The safety-related properties of the AI function relate directly to the Safety Requirements Assurance (see Section 5.8.2) and the iterative phases generate evidence for the other ML stages. One difference is the greater emphasis on the operational monitoring of the system, including the AI/ML component.

Inductive analyses such as Failure Modes and Effects Analysis (FMEA) can be used to analyse the various types of error classes and underlying insufficiencies in the model that could lead to the violation of a safety goal, and their underlying causes. Alternatively, deductive forms of safety analysis could include evaluating specific safety-related failures discovered during validation to determine the underlying causes and improvements to the ML system required to prevent their occurrence in future. Table 28 includes an excerpt from such an analysis.

Error Class	Insufficiency	Cause	Design-time measure	Metric	Operation-time measure	Metric
Incorrect classification	Lack of generalization	Under-specification, scalable oversight	Balanced training set	Coverage of ODD model	N/A	N/A
Incorrect classification	Unreliable confidence values	Over-confidence due to uncalibrated soft-max values	Temperature scaling	Remaining Error Rate, Remaining Accuracy Rate	N/A	N/A
...	...	...	...	...	...	...

False negatives	Lack of robustness	Instability of DNNs for minor changes to the inputs	Adversarial training	Adversarial and perturbation robustness	Robustness certificates	Certifiable perturbation strength
Sequence of false negatives	Lack of generalization	Under-specification, lack of scalable oversight	Balanced training set	Coverage of ODD model	Comparison with other sensor data	Diagnostic coverage
...	...	...	...	...	...	...
False positives	Clever Hans effect	Spurious correlations in the training data	Diversified training set	Conceptual disentanglement	Plausibility checks	Diagnostic coverage
False positives	Lack of generalization	Distributional shift	N/A	N/A	Out of distribution detection	Diagnostic coverage
...	...	...	...	...	...	...

Table 28: Causal analysis of error classes for a DNN-based perception function

**Summary of considerations related to allowing the use of ML-based perception functions using Deep Neural Networks:**

With respect to the criteria outlined in Section 5.8.1, the task of identifying vulnerable road users and other vehicles based on video data can be considered highly complex. Furthermore, this function can have a direct impact on the safety of the decision-making function. Internationally recognised standards, though under development, have yet to be published. A systematic approach for arguing the impact of safety of errors in ML-based perception functions is outlined above, so this falls into class 2, and it is expected that upcoming standards will closely follow such approaches. Nevertheless, ‘assurance uncertainties’ are likely to remain, and such applications of machine learning will inevitably lead to a generalization error rate several orders of magnitude away from tolerable system-level accident rates. Pre-requisites to the use of such technologies should therefore include:

- a demonstrable understanding of the requirements on the ML component within the system context, including the bounds on decision uncertainty and error rates necessary to achieve a tolerable residual failure rate at system level
- system-level measures (e.g. sensor and algorithmic uncertainty, monitoring and plausibility checks) to limit the propagation of perception errors to safety-related failures,
- a completed assurance case according to the criteria outlined in sections 5.8.1 to 5.8.6 of this report, including confirmation by a qualified third party.

These conclusions are summarised in Table 29.

Technology class Usage class	1: Existing standards can be used	2: No existing standards, alternative methods can be identified	3: No alternative methods can be identified
A: Not part of the safety function	N/A		
B: Used in the development of the safety-related system	N/A	N/A	N/A
C: Has a direct impact on the safety-related decision making of the system	No, but currently under development: ISO TS 5083, ISO/IEC TR 5469, ISO PAS 8800	Initial concepts have been developed for using supervised-learning based safety-related functions. However, current performance of the functions is not sufficient for safety-critical functions and uncertainties still exist within the assurance arguments. The target function has a high degree of systematic complexity.  <b>Therefore, the safety-related properties including safe bounds on residual errors shall be defined, safety-related decision functions shall not rely solely on the ML-based perception function and, performance shall be continuously monitored after deployment and the assurance case shall be rigorously inspected by a third party.</b>	<b>Shall be prohibited in the case that the conditions under column 2 are not satisfied.</b>

Table 29: Summary of consideration related to the use of ML-based perception functions using Deep Neural Networks

### 5.8.7.3 Online reinforcement learning for driving policy optimisation

Reinforcement learning (RL) is a branch of machine learning whereby the algorithm gathers information from the environment and automatically optimises its function based on a pre-defined reward policy. This approach allows for a system to adapt its function during operation based on observations in the field, in comparison with supervised learning approaches where training data must be collected during the development phase. Applications of RL in driving tasks can range from high level functions such as selecting the most efficient route to a destination, trajectory calculation for changing lanes, overtaking etc. and low-level real-time control tasks to execute trajectory paths.

The advantages of RL are increased efficiency during the development of the function, as training data do not need to be systematically collected, analysed and labelled as for supervised learning. In addition, RL algorithms have the ability to continuously adapt to changes in the environment. These advantages have the potential to allow automated vehicles to apply more naturalistic and adaptive approaches to driving.

A key challenge associated with RL is the selection of a reward policy that leads to the required level of performance in the function under all conditions. A common cause of failures within RL is co-called “reward hacking”, whereby naive reward functions, such as “maintain a safe longitudinal distance x to other vehicles”, could lead to unwanted and potentially hazardous scenarios where the vehicle continuously reduces speed in dense traffic as other vehicles fill the gap to the vehicle in front, or where the ego vehicle finds other means to maintain distance, such as driving on the hard shoulder. The relationship between the systematic problem complexity and the choice of the most appropriate reward policy therefore replaces the training data selection problem seen in supervised learning.

RL can be used to develop a function using pre-deployment (off-line) training, for example using simulation environments or recorded field data. However, it also has the potential to be used post-deployment (online) to continuously adapt the function in the target operating environment. This leads to the following considerations with respect to safety assurance:

- Offline, RL:** The safety assurance lifecycle for offline RL mirrors that given in Figure 29, and is amenable to assurance approaches such as those defined in AMLAS. The function is iteratively developed and within each training phase, the performance of the function is measured against the safety requirements, root causes of insufficiencies (e.g. inappropriate reward policy) identified and operation-time measures to reducing residual errors selected. This process is repeated until sufficient evidence for the safety of the function has been collected and the function can be released for deployment. In the offline use case, the function is then “frozen” and no-longer adapted within the deployment environment. If insufficiencies are found, then additional “train, test, analyse” iterations are performed before releasing a new version of the software (e.g. in the form of a fleet-wide over-the-air update).
- Online, RL:** In this case, after the initial offline development and assurance phase, the function is allowed to continuously adapt in the field after deployment. This would lead to each vehicle running a different, untested, version of the function. In this case, it must be ensured that changes to the function over time cannot lead to safety-related failures. This may be achieved, for example, by bounding the changes of parameters of the function, so that they remain within pre-determined, safe ranges based on tolerances calculated during the initial development and assurance phase. Online learning also has the disadvantage that experience gained with the use of the system within the operating domain cannot necessarily be aggregated across all vehicles (although approaches to so-called “federated learning” are possible where data is shared a fleet-wide updates to the learnt model are distributed, see section 5.7.4). This limits the ability to leverage data collected from the entire fleet to improve the quality of the function. This may erode the effectiveness of the continuous assurance activities outlined elsewhere in this document.

In conclusion, the use of online RL should be prohibited unless it can be clearly demonstrated that a failure of the function does not have an impact on safety or that changes to the previously assured function during operation can be restricted to safe ranges that have been pre-determined and validated as part of system development activities. Examples of such permissible uses could include route navigation, adapting to local traffic congestion patterns and optimisation of trajectory parameters to account for limited amounts of sensor drift or actuator degradation. A summary of recommendations for safe online RL can be found in Table 30.

Technology class Usage class	1: Existing standards can be used	2: No existing standards, alternative methods can be identified	3: No alternative methods can be identified
A: Not part of the safety function	Online learning does not have a direct impact on safety-relevant decisions of the function, e.g. online, RL is only used to optimise comfort or efficiency of the driving task, e.g. in the calculation of the driving route. <b>No safety-related requirements are allocated to the function, as confirmed via the hazard and risk analysis (e.g. according to ISO 26262-3).</b>		
B: Used in the development of the safety-related system	N/A	N/A	N/A

<p>C: Has a direct impact on the safety-related decision making of the system</p>	<p>N/A</p>	<p>If changes to the function due to the online RL are bounded by <b>statically defined constraints</b>, implemented according to existing standards. (e.g. ISO 26262-6), then their use could be considered. An argument for the effectiveness of these constraints is required.</p> <p>This argument must be based on a detailed analysis of the classes of residual errors expected in the ML function and the effectiveness of operation-time measures to constrain the impact of such errors.</p> <p><b>Therefore, the safety-related properties including safe bounds on residual errors shall be defined, safety-related decisions shall be bounded by independent SW functions implemented according to safety standards, performance is continuously monitored after deployment and the assurance case shall be rigorously inspected by a third party.</b></p>	<p><b>If no approaches are known for directly ensuring the continuous safety of online learning approaches. Then their use must be prohibited.</b></p>
---	------------	---	--

Table 30: Recommendations for the use of online RL

### 5.8.7.4 Off-line analysis of driving data

Data collected during the deployment of the vehicle can serve several purposes. Firstly, the data could be used to train future iterations of ML-based driving functions. Secondly, the data could be used to detect previously unknown risk factors through off-line analysis. Thirdly, the data can be used to determine the causes of safety-related incidents. The first usage is related to the development of vehicle functions, whilst the second and third are related to monitoring and continuous assurance activities. For these activities, ML techniques themselves could be used, e.g. regression analysis and clustering for identifying correlations between driving behaviour and environmental conditions. Deficiencies in either the data collection or its analysis could lead to inappropriate decisions related to the continuing safety of the system.

This results in the following requirements that must be considered during the offline analysis of driving data:

- **Requirements on the quality of the data itself:** The specification, collection and management of the data should follow the requirements outlined in Section 5.8.3.
- **Requirements on the analysis of the data:** The analysis activities and associated tools shall be evaluated according to the criteria from ISO 26262-8 “Confidence in use of software tools” in terms of the potential *impact* of failures in the activities (e.g. failure to detect previously unknown critical scenarios) and the ability to *detect* such failures in the process. The impact of failures in the data analysis depends on the analysis task itself and its role in the process. Due to the large amounts of digital information involved, it is infeasible to expect that manual review processes could be used to counteract or detect failures in the analysis. Therefore, where ML approaches are used to analyse the data, similar assurance activities such as those described



above should be applied in order to increase the likelihood that tool errors are detected. Furthermore, where a failure in the ML-based analysis could have an impact on safety-related assurance decisions, a diverse set of algorithms shall be applied for the analysis to avoid single-point failures and further increase the likelihood of the detection of errors in individual analysis tools.

In summary, where ML algorithms are used for the analysis of driving data as part of continuous assurance activities, an assurance argument for the use of such tools should be created and the tools should only be used where a failure in the tool cannot impact decisions regarding the safe deployment or continuous safety of the vehicle. These recommendations are summarised in Table 31.

Technology class Usage class	1: Existing standards can be used	2: No existing standards, alternative methods can be identified	3: No alternative methods can be identified
A: Not part of the safety function	N/A		
B: Used in the development of the safety-related system	<p>A conformance to relevant safety standards shall be demonstrated. E.g. ISO 26262-8: Confidence in use of software tools.</p> <p>The results of the development activities performed using AI/ML shall be confirmed using alternative means (e.g. review, test) and documented in the assurance case. <b>Automated decision making that may impact decisions regarding the safe deployment or continuous assurance of the vehicle shall not be based on a single machine learning technique.</b></p>	<p>Alternative evidence shall be presented in the form of an assurance case.</p> <p>The results of the development activities performed using AI/ML shall be confirmed using alternative means (e.g. review, test) and documented in the assurance case. <b>Automated decision making that may impact decisions regarding the safe deployment or continuous assurance of the vehicle shall not be based on a single machine learning technique.</b></p>	<b>Strictly prohibited.</b>
C: Has a direct impact on the safety-related decision making of the system	N/A	N/A	N/A

Table 31: Summary of recommendations for the use of ML in the offline analysis of driving data.

## 5.9 Test Programmes

### 5.9.1 Background and State of the Art

Test programmes is the umbrella term used within this report to refer to the most general form of **evidence-gathering activities**. It will at times be used inter-changeably with **V&V activities** and **V&V programme**. From a scientific perspective, these should be viewed as the suite(s) of all *experiments* that should be conducted in order to demonstrate, or otherwise, that the stated hypothesis, for example “the AV satisfies all the safety goals and acceptance criteria with sufficient confidence so as to be considered by the Department for Transport safe for deployment into its ODD and TOD, which does (not) include UK public highways”, is upheld and supported by the evidence collected therein.

Test programmes, as a suite of experiments to assert a hypothesis, are not new within the automotive domain; processes exist today to provide evidence of safe operating of both vehicles and their human drivers alike (Section 5.9.1.1). In fact, experimentation at large is nothing bespoke to automotive applications, either. This slight generalisation of *testing* to *experimenting* is more than just semantic; it should be philosophical and cultural. Testing for an AV will, and very arguably must, look different from existing automotive test programmes, and not simply seek to extend existing practices and remits. There are fundamental differences, which should be embraced, and it is for this reason that this report frames its recommendations surrounding programmes in a more general context.

#### 5.9.1.1 Test Programmes for Conventional Vehicles

Existing mandatory test programmes for conventional vehicles are driven by the UK’s Type Approval process; optional further testing is often undertaken to comply with industry standards or consumer group testing, such as being EuroNCAP compliant. Both Type Approval and EuroNCAP test programmes ultimately comprise a highly prescriptive set of tests, which is standardised for every vehicle category, and every instance of the programme. Whilst a manufacturer would typically perform their own testing for development and internal sign-off purposes, regulatory testing within GB and the EU is either performed by an accredited ‘Technical Service’ representing the regulatory body, or is conducted within accredited facilities and witnessed by a Technical Service.

Perhaps more relevant to this section, though, is what existing test programmes *are not*. Existing regulatory test programmes for safety have extremely limited variability *by design*; do not employ any random sampling strategy; are very far from exhaustive; and are focussed almost exclusively on asserting the correct function of passive and active safety features. Unsurprisingly, there is no need to test the decision-making, vehicle-level behaviours: this is left to the human driver, who must separately qualify for a Driver’s Licence by passing a theory and practical test. Existing test programmes and safety arguments rely heavily on the (reasonable) assumption that the human driver is the entity which ‘interpolates’ between two familiar situations, to deal with a third, unfamiliar situation.

There is therefore a clear yet crucial difference between what existing test programmes are designed to provide evidence of, and what new, ADS test programmes need to provide evidence of. At the highest level, both forms of test programme provide evidence to support satisfaction of a set of safety-related objectives. Within existing safety regulations, such objectives are practical outcomes that are tied explicitly to the test programme itself (whether Type Approval, EuroNCAP or others); the set of safety goals proposed in Section 3.3 are far more abstract, since the evolved requirement of automated vehicles is to achieve vehicle-level behaviours, properties and competencies (e.g., “do not cause at fault collisions”), rather than test-level, specific outcomes (e.g. forces experienced by test dummies, magnitude of deformation of a survival cell, deceleration achieved by braking systems).

It goes without saying that there are no long-established, state-of-the-art processes or protocols for testing highly-automated vehicles, as is the case for conventional vehicles; similarly, highly-automated vehicles lack the decades of legal precedence that has established for conventional vehicles.

### 5.9.1.2 Handling, Describing and Codifying Driving Scenarios

The field of Scenario-Based Testing (SBT) has established several conventions for handling road traffic scenarios, at several levels of detail. As well as protocols such as that established by the PEGASUS project in Germany (PEGASUS, 2019), where the infrastructure and environmental aspects of a scenario are broken down into multiple “layers”, a familiar hierarchy for the level of prescriptive scenario detail is the *Functional > Logical > Concrete* regime.

- *Functional Scenarios* are the most abstract, taking the form of a formal but high-level description of what occurs during the scenario.
- *Logical Scenarios* are parameterised forms of scenario wherein the relevant variables, whose values can vary over a range while still remaining *functionally* within the bounds of the same basis of events, are identified, and the bounds between which their values can vary are stated (max/min). There is no formal restriction on what constitutes a valid parameter, but common examples are the speeds, separations distances and accelerations (and derivatives thereof) of several actors in the scene, as well as factors like precipitation, time of day (and year) or visibility range.
- *Concrete Scenarios* are the most prescriptive in the hierarchy, and should be thought of as particular instances of a *parent logical scenario*. The exact initial values of all the relevant parameters in the scenario are prescribed and thus a precisely agreeable and reproducible driving “scene” (effectively a freeze-frame or snapshot) becomes the starting point for integrating the scenario forward in time. Thus, a set of concrete scenarios should be selected to sample from within the ‘scenario space’ defined by the ranges of the parent logical scenario parameters.

When reviewing and approving the outputs of a test programme, it will be of vital importance that the scenarios pertaining to each test are **easily traceable**. UL4600 accords with this as a mandatory requirement in its section 12.2.1.1-a.4 [UL4600 citation]. A standardised system for labelling and sub-structuring the network of scenarios explored during the test programme is therefore highly recommended. There might be several ways to achieve this, but the so-called ‘heritage’ of each scenario should be clear and unambiguous. For example,

## Scenario 2.3.8

might mean,

- (i) **Concrete scenario 8**, which is a prescribed instance of
- (ii) **Logical scenario 3 (parent)**, which is itself one parameterisation of
- (iii) **Functional scenario 2 (grandparent)**.

The exact description/parameterisation/prescription of each functional/logical/concrete scenario shall be stored in a database and human-readable records of each should be available upon request from the approval authority.

In order to ensure acceptable coverage of the scenarios that the vehicle could face within its actual deployment, it will be essential to ensure that the suite of functional scenarios provides coverage of the behavioural competencies defined for the vehicle; that the parameters and their ranges as identified in the logical scenarios provide sufficient coverage of the TOD; and that both the behavioural competencies and TOD are accurate reflections of the intended deployment of the vehicle (see Section 4.2). A process to ensure traceability back to, and coverage of, the definitions for the system and its operating environment will therefore be required (HumanDrive, 2020).

### 5.9.2 Reporting results

Throughout this section, ‘*outcome*’ is used to describe the result of any form of test, reported in any format, as simple as a Boolean PASS, or as detailed as a full Ground-Truth record; ‘*output*’ is used to

describe the relevant document or other deliverable through which the approval authority shall actually receive the summary and details of all of the test *outcomes*.

### 5.9.2.1 How to report the outcome of a test

Current arrangements for conventional vehicles are simply that, at the end of a test programme, a summary (the *output*) is generated which lists every test which was conducted, and whether the vehicle-under-test was deemed to have achieved a PASS or FAIL in each (the *outcomes*). The exact language and format differ between different testing streams (e.g., Type-Approval testing versus EuroNCAP), but the *simplicity* of the outcomes is almost universal.

It is acknowledged that some existing aspects of and approaches to vehicle testing may remain unchanged in the AV domain, in the spirit of “if it ain’t broke, don’t fix it”. Standards, best practices, and conventions already exist for reporting the outcomes of, say, software unit tests; at this low, granular, function/ clause/ component level, the distinction between a conventional and an automated vehicle is minor, if not none.

Therefore, the novel and most non-trivial aspects of an AV test programme should be targeted at finding and asserting the safety of **emergent properties, characteristics and behaviours at the vehicle-level**. This is a totally unique aspect of AV testing since the equivalent ‘behaviours’ of a conventional vehicle are in fact those of the human driver, and not those of the vehicle (ADS) itself at all.

There are several ways in which reported ADS test programme outcomes shall differ:

- (a) Boolean PASS versus FAIL is likely to be insufficient to properly capture the outcome of a more sophisticated test (e.g. in SBT) with sufficient fidelity as to be at all useful in the decision-making process of the approval authority. Instead, a **vector of Boolean event status flags** (effectively 0 or 1 for every value in an array) should be used to capture the affirmative/negative status of all relevant events after a test has completed; let this be the so-called **flag vector**. It is recommended that the nominal (desirable) value of all flags in this vector be zero; that way, the (apparently) “perfect” outcome of any test is the zero vector, [0 0 0 ... 0 0].

The set of Boolean events which should be monitored and reported against will not be prescribed *exactly*, since the specific ADS functionality (and therefore responsibility) in question, and its particular ODD, will readily render different flags more or less relevant – or in the extreme, totally irrelevant. Nevertheless, to give a flavour of what is intended by such *events*, take as examples:

- (i) “collided with any object”
- (ii) “collided with a dynamic object”
- (iii) “collided with a VRU”
- (iv) “collided with a static object”
- (v) “departed lane by greater than allowed tolerance”
- (vi) “passed a red light”
- (vii) “exceeded maximum comfortable acceleration/deceleration threshold”

This set of seven events is obviously not exhaustive. Notice that some of these events are conditionally dependent on one another; this is intentional. The first advantage of such a set of events is that it yields both a macro (e.g., “collided with anything”) and a micro (e.g., “collided with VRU”) indication of what happened, both generically and more specifically. The second important motivation for such overlap between events is that it *designs in* an implicit consistency and correlation check, at the small cost of an element of redundancy. For example, if “collided with VRU” returned 1 (true), but “collided with any object” returned 0 (false), then the test outcome can immediately be considered void since it is obviously contradictory/ inconsistent. Similarly, if “collided with any object” returned 1, but all other (more specific) collision flags returned 0, then further investigation is necessary since it is unclear with what the vehicle-under-test has collided. It may transpire that the test outcome is void, or an unforeseen collision (category) may have occurred.

Lastly, it is anticipated that the set of events monitored during each test shall be the union of ‘permanent’ (core) and ‘temporary’ (specific) subsets. The permanent set shall comprise events which are always relevant to all aspects of driving (e.g., collisions, adherence or otherwise to traffic rules, etc.), including safety-critical events where overall incidence rates and ‘big-picture’ patterns/behaviours are being sought by the test programme at large. The (likely smaller) temporary set should comprise events which are only relevant to the driving scenario-under-test.

- (b) To complement the flag vector, a set of gradated, quantitative scores should be ascribed wherever possible according to a validated scoring regime. These are intended to serve as an added layer of detail beyond the ‘record of what happened’ given by the flag vector alone. Notice “a set” of scores is recommended; in the same way that PASS versus FAIL is considered an inadequate outcome for AV tests, so is a single, standalone score. Any individual quantitative score should indicate the magnitude of the severity of any poor or unintended performance, whether safety-critical, or pertaining to passenger comfort, or otherwise.

One proposal for a gradated, quantitative scoring structure/ method (its exact calibration is omitted and left as a decision for policy-makers), is to devise a matrix or matrices of weights to use in combination with the flag vector to generate the quantitative scores.

- (c) The test modality (e.g. simulation testing, proving ground testing) shall also be given and justified. Where the same test was carried out using multiple (two or more) modalities, an alignment/correlation analysis shall be given to demonstrate consistent outcomes, and that no contradictory evidence exists. In rare cases, it may be that the test outcomes are not well correlated; if this is to be acceptable, then adequate reason(s) shall be given.

Some hypothetical test outcomes are shown in Table 32 by way of example.

Test ID	Scenario Heritage	Non-zero flags	Quantitative scores	Test Modality	Comments
TC001	2.1.1	N/A	8.2	Software-in-the-loop simulation	
TC002	2.1.2	(7) Exceeded maximum comfortable accel/decel	26.9	Software-in-the-loop simulation	
TC003	2.1.2	(1) Collided with an object (2) Collided with a dynamic object (7) Exceeded maximum comfortable accel/decel	107.3	Proving Ground	Same test conducted as in TC002 but with alternative modality. Outcomes differ significantly and therefore TC002 is void. Take worst-case TC003 as indicative outcome.

Table 32: Perceived outcomes from three hypothetical test cases, listing the non-zero flagged events, and giving any quantitative ‘Oracle’ scores (VeriCAV, 2021) where relevant.

The same scenario heritage executed in different test modalities should yield reproducible outcomes (in a strictly scientific sense); should this not be the case, then the worst-of-many outcomes shall be taken as the principal outcome, and the lack of inter-modality correlation shall be investigated and explained to rationalise the otherwise contradictory outcomes.

### 5.9.2.2 What are the outputs of the Test Programme?

The proposed approval process for LSAVs invokes testing at two key points as set out in Section 3.1: “independent vehicle assessment and testing” (the activity of scoping and executing the test programme) acts as an input to the VSCR and DSCR. Whilst testing that is undertaken on a ‘generic’ basis (i.e., *not* directly related to a particular deployment, as defined in Section 4.1), such as for sub-systems and components, would fall within the scope of the VSCR, Section 4.1 explores two alternative options for the portion of the testing that is ‘specific’ to the particular deployment route(s) or area(s):

- include in the VSCR – this introduces the complexity of having ‘specific’ and ‘generic’ evidence in the same report, requiring that the TOD as well as the ODD be elaborated.
- include in the DSCR – this could raise challenges in terms of the regulatory body responsible for the deployment approval possessing the required technical skills, and would also mean that vehicles could be said to be ‘approved’ (on the basis of the VSCR) whilst significant safety evidence is still outstanding, which may be grossly misleading.

As set out previously, the decision on this is one that should be taken by the DfT once the regulatory body or bodies responsible for each phase have been identified and the legal framework within which the approval operates has been defined; this report therefore draws no conclusion as to which option should be selected, and merely seeks to set out the technical background to support the making of such a decision.

Whilst the exact design, modality and purpose of each individual test, and the distribution of tests overall, cannot plausibly nor appropriately be prescribed for all possible future ADSs and ODDs, it is reasonable to recommend the ‘meta’ requirements of the portfolio/suite of tests whose outcomes inform the output from the test programme with regard to the ‘specific’ and ‘generic’ portions of the test evidence:

- (i) *Test evidence collected on a ‘generic’ basis*: this evidence shall not necessarily be deployment (TOD) specific, since it is intended to support the hypothesis that the vehicle is safe *from first principles (ab initio)*. A non-exhaustive contents list for this first test programme output might contain the outcomes and conclusions drawn from suites of tests which:
  - are especially relevant to the middle and lower parts of the right-hand-side of the engineering V-cycle;
    - that is to say, component-level and upwards testing, including software-in-the-loop simulation (e.g., unit testing) and hardware-in-the-loop simulation (e.g., isolated testing and calibration for hardware such as RADARs, ultrasonic sensors, LIDARs, etc.);
    - large parts of this aspect of testing may well be conducted by Tier 1/2 suppliers, hardware manufacturers, or similar, and the outcomes simply reported ‘up the approval chain’ to be included, possibly even as an Addendum or an Appendix, in the document(s) to which the approval authority itself will be exposed; it should be remembered that the “test programme” will, in practice, be a highly distributed activity with respect to both (a) time and (b) responsible stakeholders;
  - are derived from the Verification & Validation activities/ programmes within both (c) Functional Safety (ISO 26262, 2018) and (d) SOTIF (ISO/PAS 21448, 2019) analyses;
    - by leveraging existing safety standards and processes, these outcomes will provide evidence that the vehicle is sufficiently safe with regard to faults, and that nominal operation is not infringed upon by any known or unknown performance limitations;

- test individual functionalities in isolation, to demonstrate fundamental performance and safety (before mixing in any complicating or adversarial factors to which the system must be robust);
  - this “divide and conquer” testing strategy has been suggested by previous research projects (HumanDrive, 2020) as one way to achieve basic levels of coverage without the need for excessive volumes of testing; insofar that this first phase is intended to provide evidence of the fundamental ‘building blocks’ of safety operation, this or a similar modular/compartmentalised approach to baseline testing should be adopted, so as to filter out ‘non-starters’ and flag immediate issues as early as possible in the approval process;
- test the functionality of multiple (and all) integrated (sub-)systems in combination/conjunction;
  - ‘integration testing’ should not be confused with ‘equivalent mileage accumulation’ or ‘coverage’ testing, wherein the objective is to maximise exposure of the ADS to many realistic situations within the test programme; instead, ‘integration testing’ is the targeted testing of systems acting in unison (e.g., the sensor fusion software correctly identifying objects given an array of detection hardware and its returned signals, rather than simply asserting positive detection by, and calibration of, the RADAR, HD camera, or LIDAR in isolation);
- seek to expose the ADS to a wide and well-distributed variety of realistic driving situations (roads and traffic), through ‘equivalent mileage accumulation’ or ‘coverage’ testing (see Section 5.9.4);
  - this is where SBT with the complete vehicle (e.g. upon a proving ground, upon the real route, or potentially within a ‘vehicle-in-the-loop’ simulation) is most likely to be employed, and will also present the most ‘opaque’ test outcomes, insofar that the tests themselves will be some of the most sophisticated in the whole test programme;
  - previous research projects (HumanDrive, 2020, VeriCAV 2021) have suggested various ‘fuzzing’ methodologies to achieve coverage of realistic scenario permutations/ combinations;
- strike a balance between achieving high sensitivity and high specificity regarding their outcomes (see Section 5.9.3.45.9.3.3);
  - previous research projects (VeriCAV, 2021) have suggested that there should be two deliberately distinct ‘flavours’ of test that make up an overall testing regime, those which bolster overall sensitivity and those which, conversely, bolster overall specificity;
- test all of the non-ADS facets of the vehicle as a whole, and its super-system;
  - this is the remit of Work Package 4 within this project; please refer to the relevant work products for further details and elaboration in this area.



- (ii) *Test evidence collected on a specific basis*: unlike the ‘generic’ evidence, all evidence presented in this phase shall be specifically intended to support the hypothesis that the vehicle behaves safely and appropriately when operating within its intended deployment, as defined by the OD. A non-exhaustive contents list for this the second test programme output might include the outcomes and conclusions drawn from suites of tests which:
- are especially relevant to the upper parts of the right-hand-side of the engineering V-cycle, with a greater focus than was applied for the ‘generic’ testing on system/vehicle validation;
    - test modalities should evolve between outputs; it is anticipated that the ‘specific’ test evidence will include significant testing upon the actual route(s) or on a providing ground, which may be less relevant to the generic test evidence, for example;
  - focus on establishing the emergent properties, characteristics and behaviours of the ADS at the logical and functional scenario levels;
    - tests should not be overly “compartmentalised” nor modular, instead embracing “fuzzy” factors, the chaotic continuum of inputs, and any edge cases that arise, even if unintentionally;
  - capture ‘dull’ false negative events (see Section 5.9.3.4);
    - a similar idea was proposed in the Digital Commentary Driving (DCD) interim research report document (BSI, 2021) – whilst ostensibly aimed at in-service monitoring, the same principle is applicable to monitoring performance within a test programme;
  - actively seek out and intend to reduce the relative sizes of SOTIF scenario areas 2 and 3 (in ISO/PAS 21448 terminology);
    - that is, actively root out and test all *known hazards* and *unknown hazards*.

### 5.9.3 Arguing ADS Safety robustly and with confidence

To briefly disambiguate some language before going any further: a *statistic* is just a function of some *observed data*, which may have been captured through any modality, be that simulation, proving ground testing, real-world usage post-deployment, and so on; *confidence*, *certainty* and *belief* will be used interchangeably to mean the *likelihood* that a stated *hypothesis* is, in fact, valid or true in reality; a *hypothesis*, then, is a statement of ‘fact’, the belief in which can reasonably and measurably be affected by the gathering and consideration of (a.k.a. conditioning upon) relevant *evidence*.

#### 5.9.3.1 How confident should the parties be?

It is clearly impossible to prove anything with 100% confidence; it is well-accepted that one cannot prove a negated hypothesis either. The natural question then, is what level of confidence in the safety hypothesis is acceptable at the point of deployment (and six months post-deployment etc.)? What are the appropriate sample and population statistics (mathematically and legally speaking) from which this confidence may be systematically and consistently derived?

Confidence intervals are a long-established and easily interpreted measure of *certainty derived from (sample) statistics*. A 95% confidence interval indicates that the *true* value of some population statistic is 95% likely to belong inside the stated range, and 5% likely to lie outside of it. One quickly appreciates that in the context of safety assurance, usually only one tail of the distribution will be safety critical; the other would represent *safer than necessary* values. Another common measure of certainty in a hypothesis is to count standard deviations (*n-sigma*), in accordance with the *Central Limit Theorem* (CLT): in most branches of science, a “scientific discovery” is considered to have been made once the certainty reaches  $5\sigma$ ; medicine by contrast conventionally uses only  $2\sigma$ , forgoing certainty for agility, so that patients can benefit as soon as *some* confidence is established, rather than missing out while waiting for high levels of patient coverage (note that test patients/volunteers can often be sparse

for rare medical conditions). The appropriate level of certainty, then, is subjective and depends on the context of the hypothesis in question.

**This report proposes that the exact confidence level to be used by the DfT, and related organisations such as the VCA, is not prescribed.** Instead, the systems of test execution, assessment and analysis which generate the safety assurance argument shall be agnostic to the chosen certainty, which behaves as a **configurable parameter** in the process.

### 5.9.3.2 Statistical Validation of the Supporting Safety Goals (4) to (21); quantifying compliance and asserting confidence

A set of appropriate and meaningful statistics, in which sufficient confidence must be evidenced, can be recommended, though. Clearly, it is unrealistic and arguably unsafe to attempt to assert confidence in a very vague and sweeping hypothesis such as “*the ADS operates safely within its deployment domain*”. Instead, Table 33 lists some more tractable sample and population statistics which could be evaluated or estimated, respectively, using the evidence (i.e., test outcomes) presented in the test programme outputs (see Section 5.9.2.2). These are worded in a far more representative manner, so as to be as clear as possible what the statistic actually means, and therefore what a regulator or approval authority might reasonably infer from its value, or otherwise.

Table 34 goes further and attempts to partner bespoke population statistics with the Safety Goals (SGs) from Section 3.3.1, wherever this is deemed to be appropriate, technically feasible, and ethical. In some cases, more general statistical approaches and ideas are proposed; a small minority of SGs are deemed to be outside the scope of statistical validation and assertion altogether. In all cases, the inferential implications and caveats for approval authorities are discussed. Some proposals are stated as applications of, or variations on, population statistics PS-1 to PS-3 given in Table 33.

Notice also that most of the statistics are **normalised**: this means that they are not dependent on the size of the sample of evidence from which they are derived; in practice, this manifests as the “per” in “X per operating hour” or “X per operating mile”. Overall, there should be a preference for “per unit distance” statistics over “per unit time” statistics, since the latter can be negatively affected by an ADS spending disproportionate, or in extreme cases entirely unreasonable, amounts of time ‘doing nothing’; this would arbitrarily extend the time over which the same number of (bad) events occurred, thus artificially reducing the *apparent* rate of their incidence. Nevertheless, for the sake of redundancy and consistency checks, a minority of some “per unit time” statistics may reasonably be permitted, and it should be considered that assessing systems using per-mile rather than per-hour statistics may proportionally favour vehicles that operate at higher speeds, or indeed scenarios that inherently lead to lower speeds.

It is important to stress that none of the statistics given in Table 33 or Table 34 should be interpreted as explicit recommendations, never mind requirements, nor quoted as the “gold standard”; there exists no perfect statistic to evidence a single assertion. Instead, the contents should be considered as an **advanced starting point**, requiring review, evaluation, and validation.

In that same vein, then, this report **strongly stipulates a requirement** that,

The Department, the regulator, the approval authority, and any other government or government-contracted organisation or other person(s) who shall either:

- (a) be exposed to the statistical (quantitative) evidence supporting an ADS safety case; or,
- (b) be making determinations on the acceptability, appropriateness, validity, or inferential value of the statistical (quantitative) evidence supporting an ADS safety case;

must possess the appropriate **mathematical**, **scientific**, and **statistical** background and skills. This is self-explanatory at the individual level. At the organisational level, this should be achieved by hiring, or contracting, the **appropriately qualified and experienced staff** to make educated, informed and soundly reasoned judgements on the statistical evidence put before them.

Where necessary, it shall be incumbent upon The Department, the regulator, the approval authority, and others to provide **training** to the same effect, where it is deemed and assessed that existing staff are not adequately qualified and/or trained for appraising and judging the large quantities of highly non-trivial statistical data which are likely to be sought in support of an ADS safety case argumentation.

ID	Population Statistic	Overall Performance Requirement (Associated Safety Hypothesis)	Critical value or bound (indicative, not prescribed)
PS-1	Rate of dynamic (non-stationary) collisions with VRUs per operating mile	<p>“The median estimate of the rate of collisions with VRUs per operating mile shall be no greater than {safety critical upper bound}.”</p> <p>OR</p> <p>“The 95<sup>th</sup> percentile of the estimate of the rate of collisions with VRUs per operating mile shall be no greater than {safety critical upper bound}.”</p>	<p>Median: 10<sup>-6</sup></p> <p>95<sup>th</sup> Percentile: 10<sup>-5</sup></p>
PS-2	Variance from the instantaneous <i>unit</i> speed limit	<p>“The sample variance of the ego speed with respect to the instantaneous <i>unit</i> speed limit shall be no less than {safety critical lower bound} and no greater than {progress critical upper bound}.”</p>	<p>Lower: 0.05</p> <p>Upper: 0.65</p>
PS-3	100-period (100-mile) Simple Moving Average (SMA) of false negative OEDR events per operating mile	<p>“The maximum observed value of the 100-period SMA of false negative OEDR events per operating mile shall be no greater than {safety critical upper bound}.”</p>	<p>Upper: 0.10</p>

Table 33: Example Sample & Population statistics to be evaluated or estimated, respectively, using the evidence presented in the test programme outputs, derived from the set of all test outcomes, post-V&V activities. The critical values stated are for the purposes of illustration, and should not be taken as proposed requirements.

SG-ID	Articulation	Proposed Statistical Approaches, including explicitly defined Sample & Population statistics; Good Data Practices and Considerations; Inferential implications and caveats
4	Follow the rules of the road	<p>Leveraging the “codified rule set” / Highway Code output from Work Package 2 of this project, it should be possible to automatically track and record the number of contraventions of any given “rule of the road” as per the wording of SG-ID4.</p> <p>Recommend two separate statistics: one for breaking “must” rules and one for breaking “should” rules.</p> <p>Clearly the former should be close to zero without good explanation; the latter may reasonably be greater (even by one to two orders of magnitude). This assumes a certain ‘equivalency’ between all “should” rules, which may or may not prove to be a reasonable assumption.</p>
5	Approach intersections with care	<p>Suggested statistics:</p> <p>(i) Kolmogorov-Smirnov <i>style</i> statistic, measuring the maximum unit distance / velocity / acceleration / jerk / jounce (or in general, any <math>n^{th}</math> derivative of displacement) discrepancy from a benchmark ‘ideal’ profile of the same derivative expressed in a canonical co-ordinate (i.e., <math>q, \dot{q}, \ddot{q}, \ddot{\ddot{q}}</math>, and so on).</p>
6	Drive only into clear space	<p>Non-trivial to measure statistically; this would require a formalised, algorithmic definition of “clear space” as per the wording of SG-ID6.</p> <p>Insofar that driving into space which is not “clear” would likely yield a collision or near-miss, PS-1 would be informative/indicative of compliance with SG-ID6, or otherwise.</p>
7	Adjust vehicle speed to prevailing conditions	<p>A variation on PS-2 would be appropriate here, whereby the “instantaneous unit speed limit” comprises a more sophisticated function of the “prevailing conditions” as per the wording of SG-ID7.</p> <p>In other words, consider the variance from an ‘appropriate’ unit speed-limit, rather than simply the ‘legal’ unit speed-limit.</p> <p>“Prevailing conditions” may consider some or all of: meteorological (weather); environmental (urban, rural, pedestrianised); temporal (time of day); situational (within TOD vs. MRX or similar compromised / uncertain state); and/or visibility factors.</p>

<p>8</p>	<p>Prioritise human life while reducing damage</p>	<p>This SG-ID8 presents significant ethical questions if it were to be measured or approved/disapproved according to a quantitative or statistical regime. While all the SGs plausibly correlate with the prevention (or otherwise) of human fatalities and injuries, no others do so in such a direct fashion as SG-ID8.</p> <p>The farthest that this report can recommend is to suggest statistic (i) which finds the ratio of human major injuries and fatalities (considered to be equally undesirable) to cost of damage to property (internal and external to ego). This statistic (i) must be taken in tandem with (ii):</p> <ul style="list-style-type: none"> <li>(i) injuries and fatalities per £100,000 damage to property;</li> <li>(ii) cost of damage to property per operating mile.</li> </ul> <p>(i) should ideally then be as low as possible. There is no lower bound below which the statistic (i) is 'acceptable', as opposed to unacceptable, and this report advises against there ever being such a prescriptive or absolute lower bound imposed in the future, since this would effectively require financial measurement of the value of human life. While this makes interpretation of statistic (i) challenging, especially for early applications, it will over time allow for filtering out of egregiously poor LSAVs against previous approvals and emerging safety standards of the day, by rejecting values of (i) which are judged to be unacceptably high by e.g., VCA.</p> <p>(ii) is a necessary sibling statistic, since (i) could in theory be artificially suppressed by high cost of damage to property (rather than achieving a true minimum of injuries and fatalities, as actually intended). The giveaway sign of this would be a low value of (i) but a very high value of (ii).</p>
<p>9</p>	<p>Drive considerately</p>	<p>The VeriCAV (2021) project yielded a similar recommendation and notionally introduced a metric to quantify 'etiquette', without prescribing any exact calculations or algorithms. It is non-trivial to define 'considerate', or conversely 'rude', driving, since this judgement is highly dependent upon the wider <b>context</b>.</p> <p>A formalised metric of "inconsiderate" driving, in the form of a set of particular (re)actions triggered <i>upon</i> other road users <i>by</i> the ego (vehicle-under-test) would plausibly permit the evaluation of a simple statistic which reports:</p> <ul style="list-style-type: none"> <li>(i) the rate of "inconsiderate" driving events per operating mile.</li> </ul> <p>This statistic may reasonably be reported for different subsets of driving scenarios (e.g., pedestrianised vs. highways), since the definition of "inconsiderate" is likely to be very different in different operating environments, even within the same ODD (superset). Clearly, the ideal value of (i) is low or very low.</p>
<p>10</p>	<p>Provide information to occupants</p>	<p>This SG-ID10 does not fall within the remit of assertion based on statistical evidence and argumentation.</p>

<p>11</p>	<p>Drive smoothly</p>	<p>The VeriCAV (2021) project yielded a similar recommendation and introduced a quantitative metric for ‘comfort’ of its own.</p> <p>Suggested statistics:</p> <ul style="list-style-type: none"> <li>(i) rate of excessive acceleration events per operating mile, whose absolute values exceed some pre-defined maximum acceleration which is considered to be the upper limit of ‘comfortable’;</li> <li>(ii) rate of excessive jerk events per operating mile, whose absolute values exceed some pre-defined maximum jerk which is considered to be the upper limit of ‘comfortable’;</li> <li>(iii) the <math>n</math>-period exponential moving average of acceleration (or jerk), time-averaged over <math>t</math>-second-long intervals;</li> <li>(iv) the maximum value of (iii);</li> <li>(v) the rate of exceedance of (iii) over a pre-defined comfort threshold per operating mile.</li> </ul> <p>The exponential moving average is suggested (instead of simple moving average) since it is anticipated that passengers will consider multiple repeated high-acceleration or high-jerk events to be more discomforting and disconcerting than the same number spread over longer periods of time (e.g., when pulling out into busy traffic, navigating a parking lot, etc.).</p> <p>Indicative values of <math>n</math> and <math>t</math> in statistics (iii), (iv), (v) are:  <math>n = \{3, 5, 10, 15, 30, 60\}</math>;  <math>t = \{0.1, 0.25, 0.50, 1.00\}</math>.</p>
<p>12</p>	<p>Travel only on appropriate lanes / road segments</p>	<p>Provided there exists a mechanism to detect travelling on inappropriate lanes or road segments, then:</p> <ul style="list-style-type: none"> <li>(i) percentage (%) of time spent driving on inappropriate lanes or road segments,</li> </ul> <p>offers a valid but simple statistic corresponding to this SG-ID12. Clearly, its reported value should be very low.</p>
<p>13</p>	<p>Do not hit a road user travelling ahead from behind</p>	<p>An estimated value or percentile of PS-1 judged to be acceptable would provide some evidence of compliance with SG-ID13.</p>

<p>14</p>	<p>Do not obstruct other road users when changing lane</p>	<p>An approach similar to that suggested against SG-ID6 could be adopted, since “driving only into clear space” and “not obstructing others while changing lanes” have some overlap.</p> <p>A formal, codified metric of “obstruction” should be established and validated; perhaps centred on whether a lane-change manoeuvre causes an otherwise unanticipated <i>braking</i> or <i>steering</i> counter-manoevre by a third vehicle.</p> <p>Asserting causality in such situations can be very challenging, but close temporal correlations may be a reasonable metric for capturing all manoeuvres by third vehicles after the ego has changed lane (both those triggered by the latter and those which are not causally related). Since it is impossible to know with any certainty different forward-integrations in time of the same scenario after a critical “branching” event occurred / did not occur, the statistically safe assumption must be that vehicles do not change lane without good reason, and so all counter-manoevres by third vehicles should be considered as causally related to the ego changing lane itself.</p> <p>In certain scenarios, if driver or ADS intent of third vehicles can robustly be asserted, and it is thusly shown that counter-manoevres were made for unrelated reasons (e.g. navigation, changing lane or braking in readiness to depart via a slip road), then those explicitly explained test cases may be excluded from the set of cases wherein the ego may have triggered the counter-manoevre (and by assumption, did).</p>
<p>15</p>	<p>When turning follow right of way rules</p>	<p>Leveraging the “codified rule set” / Highway Code output from Work Package 2 of this project, it should be possible to automatically track and record the number of occurrences wherein appropriate right-of-way rules are <i>not</i> followed.</p> <p>Two appropriate statistics for SG-ID15 would then be:</p> <ul style="list-style-type: none"> <li>(i) rate of failure to obey the right-of-way of other vehicles and VRUs when the ego executes specifically <i>turning manoeuvres</i> (percentage of <i>turning manoeuvres</i> where appropriate right-of-way is not exhibited).</li> <li>(ii) mean / median number of <i>turning manoeuvres</i> executed (elapsed) between violations of other vehicles’ and/or VRUs’ right-of-way.</li> </ul> <p>(i) shall use a sufficiently large sample of scenarios which are explicitly designed to contain <i>turning manoeuvres</i>, and not be aggregated from open-driving alone, since e.g. highway driving may contain very few such manoeuvres.</p>
<p>16</p>	<p>Follow prevailing driving styles</p>	<p>This SG-ID16 does not fall within the remit of statistical validation, nor the use of quantitative supporting evidence to argue compliance.</p>

<p>17</p>	<p>Indicate intentions as per rules</p>	<p>Insofar that illuminating the brake lights when actuating the brake pedal (and thus the brakes themselves) is an existing regulatory requirement, and similar such examples where the indication of intentions is explicitly tied to the action of executing those intentions, the correct and appropriate operation of turning signals is the primary operation within the DDT that falls within the scope of this SG-ID17.</p> <p>All manoeuvres which require turning signals to be shown should be stipulated by the regulator (e.g., turning left/right at a junction; changing lane on a dual-carriageway; showing the hazard lights when braking extremely hard or coming to an unexpected complete stop in-lane). Systematic detection of all such manoeuvres, taken together with the timeseries history of turning signal status (a Boolean for each of the L &amp; R turning signals), especially in the moments preceding such a manoeuvre, should permit simple yet reliable detection of appropriate operation of turning signals, or otherwise.</p> <p>Two sample statistics should be reported:</p> <ul style="list-style-type: none"> <li>(i) a simple percentage of events in which the turning signals were <i>not</i> shown, but should have been;</li> <li>(ii) a rate of occurrence with which the turning signals are shown for no valid reason (“indicating without intention”).</li> </ul> <p>An identical process and similar pairs of statistics could be used to evidence the appropriate usage, or otherwise, of other signalling hardware, e.g., use of the horn, flashing of headlights, etc.</p>
<p>18</p>	<p>Maintain appropriate safety margins</p>	<p>The “appropriate safety margins” should be set and justified by the developer, and audited by the approval authority. These should most likely take the form of conceptual ‘extended lobes’ of safety ‘buffer’ around the vehicle in 3D. In other words, a violation of the “appropriate safety margins” (a near miss) should be considered to be any encroachment by another vehicle or VRU inside the lobes/regions. It would be reasonable but not mandatory to have multiple layers of safety margin: e.g., an encroachment within 300 mm of the actual ego vehicle could be considered a “serious near miss” whereas an encroachment within 1000 mm could be considered as a “minor/marginal near miss”.</p> <p>In general, external violation of the safety margins (e.g., a pedestrian walking close to the vehicle while it remains stationary), as opposed to self-induced violations (e.g., choosing to brake too late and coming to a stop too close to a potential collision object) should be separated into distinct categories of ‘near miss’, although there may be some occasions on which the ego effectively put itself into a position where so-called ‘external’ violations of the safety margin were to be expected. In such cases, these should be considered as self-induced violations, unless satisfactory explanation to the contrary can be provided.</p> <p>Regarding the statistical validation of this SG-ID18, any statistic or family of statistics taken from Table 33 could be used to evidence the sufficiently low (a) frequency (exposure) and (b) severity of any violations of the “appropriate safety margins” over the course of a test programme. E.g.,</p> <ul style="list-style-type: none"> <li>(i) “the median or any percentile of the distance to the nearest object when braking from non-zero velocity to stationary”; or,</li> <li>(ii) “the 100-period (100-mile) Simple Moving Average (SMA) of the number of self-induced violations of the appropriate safety margins (buffers) as defined around the vehicle in 2D (BEV) or 3D”.</li> </ul>



19	Avoid behaviour not expected by other road users	This SG-ID19 does not fall within the remit of statistical validation, nor the use of quantitative supporting evidence to argue compliance.
20	Avoid obstructing traffic flow	<p>Obstructed traffic flow could be detected by flagging any moments in any test case where all the other road vehicles (actors) within a certain radius or proximity of the ego (vehicle-under-test) are moving with velocity ~zero (at or very close to a standstill).</p> <p>Statistic (i) could then be reported:</p> <p>(i) rate of at-fault (causal) traffic obstructions per operating hour or mile.</p>
21	Avoid behaviour not expected by occupants or persons in vicinity of vehicle	This SG-ID21 does not fall within the remit of statistical validation, nor the use of quantitative supporting evidence to argue compliance.

Table 34: Secondary Safety Goals (SGs) listed against their supporting evidential statistics, and/or any statistical and inferential considerations of which the approval authorities should be aware.

### 5.9.3.3 Sensitivity versus Specificity

Blindly considering sample statistics and estimates of population statistics alone may not guarantee confidence in the overall safety of an ADS. To have confidence in the representative nature, and therefore inferential value, of the evidence from which such statistics are themselves derived, the Sensitivity and Specificity of the test programme itself must be asserted somehow. The need to incorporate these concepts was one finding of the VeriCAV project (2021).

Sensitivity is the ability of a test to detect what it is intended and designed to test for; in the context of an AV test programme, that something might be ‘poor performance’, or more likely a particular event or behaviour such as “accurate detection of all relevant objects within range and intended field of view (FOV) at all times”. In practice, a sophisticated and multi-faceted concrete test such as those undertaken during SBT may reasonably be intended and designed to test for multiple things. Nevertheless, the principle of sensitivity still applies.

Specificity is the opposite: the ability of a test to correctly ignore (not falsely detect) outcomes wherein the thing(s) which it is (are) intended and designed to test for are **not** present.

The reader may recognise these concepts as similar to those of *false positives* and *false negatives*. This is absolutely the case; however thinking of **sensitivity** and **specificity** as desirable and intended properties of the test programme itself is often more succinct and useful vocabulary.

Figure 39 visualises what low and high sensitivity versus specificity mean in practical terms. The key message which legislators and approval authorities alike should take away is that judging a test by its ability to detect unintended or unsafe ADS behaviour alone is a false economy. Why? Well, in the extreme case of a 100% sensitive but 0% specific test programme, all test outcomes would appear to yield evidence of unacceptable ADS behaviour. Such an output would genuinely capture and flag all the truly unsafe behaviours; of course, this would *come at the cost of flagging everything* as an unsafe behaviour, though. While no unsafe behaviours go undetected by this hypothetical test programme, its potentially useful safety-critical findings are instead useless, since they are completely obliterated by the noise of false positives.

As a general rule of thumb, there will always be a need to seek compromise between maximising sensitivity and maximising specificity. Consequently, some tests should intend by design to complement one more than the other, rather than trying (and failing) to hedge between each.

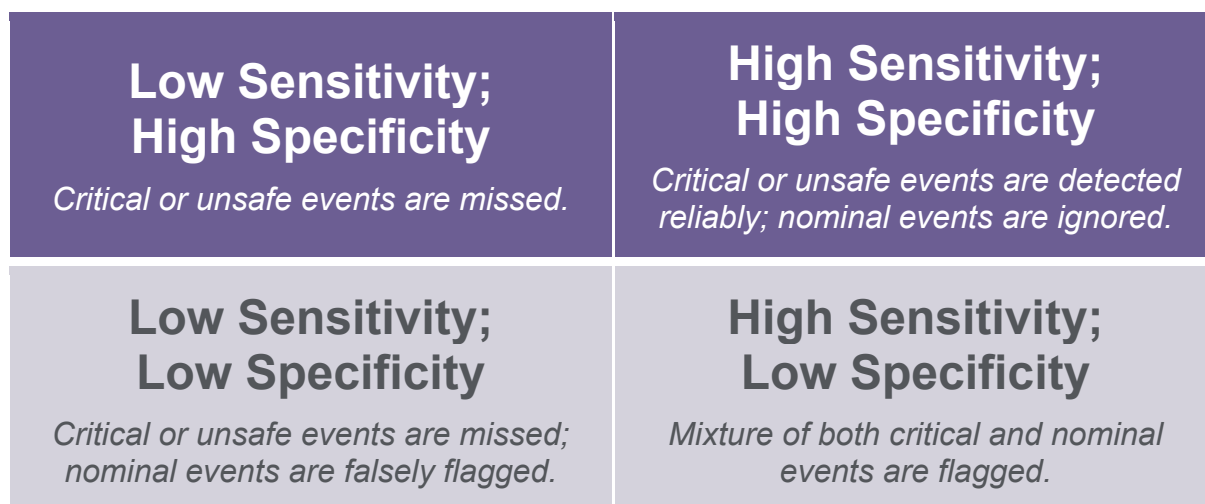


Figure 39: Sensitivity versus Specificity.

### 5.9.3.4 The importance of capturing dull false negatives

It is essential that throughout the entire test programme, consistent consideration is given to the capture of so-called ‘dull’ false negatives. A ‘dull’ false negative is a false negative event (e.g., during OEDR) which does not lead to an explicitly hazardous or harmful outcome.

Take, for example, an AV which fails to detect a pedestrian walking on the pavement alongside a road in an urban environment; the AV passes by, blissfully unaware, but does not deviate from its intended trajectory within its lane. This is a dull false negative, since an object which should have been detected according to the specified functionality, was not. The question is, then: should this be flagged as an issue? This report proposes that the answer to this should be ‘yes’, for the simple reason that, should the pedestrian have wandered into the AV’s lane (intentionally or accidentally), the AV may not have been able to avoid an otherwise avoidable collision.

The OEDR event history in this alternative scenario might involve no detection at all, or plausibly a ‘too-late’ detection as the pedestrian goes from being “undetected, no immediate collision risk” straight to “detected, immediate collision risk”. Clearly, “detected, no immediate collision risk” should appear somewhere in the intended OEDR trace, even if the pedestrian does eventually encroach into the vehicle’s lane, and regardless of whether this encroachment happens too late for the vehicle to reasonably react to avoid collision (i.e., perhaps it would not be a specified or intended behaviour to avoid collision if the pedestrian suddenly and unpredictably moves into vehicle’s lane only a short distance ahead).

The above is only one example, but it does serve to demonstrate the concept of dull false negatives. The benefit to capturing these events properly is that one might reasonably establish a theoretical (or compound) estimate for the rate of incidence of certain unsafe events *at a lower cost*, that is at a lower test burden and effort. The justification for this assertion is that many unsafe events obey a geometric probability distribution; this is sometimes referred to as the “Swiss Cheese” model in safety engineering terms.

If only one in one thousand pedestrian-related OEDR events were false negatives, only one in one hundred of these involved the undetected pedestrian actually encroaching (directly or indirectly) into the vehicle’s path, and only one in one hundred of those led to a collision and therefore harm, then only one in ten million instances of this scenario will yield a harmful outcome. This obviously requires lots of testing to establish even a poor estimate of the rate of incidence of this type of unsafe collision outcome. Please note that observing one event in ten million does *not* mean that the correct estimate for its incidence is 1 in 10,000,000. Rare Poisson-distributed events must be carefully handled, and can be normally approximated under certain conditions, to establish bounds on the true Poisson parameter  $\lambda$ . The numbers given are intended to explain the lessening of test burden by capturing dull false negatives only. However, the same ten million tests (or considerably fewer), would be sufficient to establish very high confidence in the rate at which the OEDR subsystems failed to accurately detect pedestrians as described (principally 1 in 1,000 OEDR events).

The proposals presented here rely on the argument that high confidence in the rate of false detection of pedestrians is far more meaningful, and offers greater inferential value, than low confidence in the rate of incidence of collisions between the ego and pedestrians in this fashion. Although the rate of incidence of collisions (harm) *is* the more indicative statistic, is it not practical to establish sufficiently precise estimates of it. To the contrary, by leveraging a practicably large amount of testing to establish with high confidence precise estimates in ‘root-cause’ statistics (e.g., the 1 in 1,000 false negative pedestrian OEDR events in the above example), and then employing **reasonable worst-case estimates** (upper or lower bounds) for the rates of incidence of subsequent “Swiss Cheese” events, whose intersection is what actually yields harm, one can obtain a tangible, meaningful, and fully-evidenced estimate for that same *most indicative* statistic (the rate of incidence of collisions in this example), but *at significantly lower cost*.

Ultimately, there is a trade-off to be made between the immediacy of test statistics and controlling test burden. By analysing the triggering conditions/events whose intersection yields an unsafe outcome (much like a SOTIF analysis in ISO/PAS 21448), and establishing which of those is either (a) the ‘root-cause’ condition; or, (b) the most frequently occurring condition if there are several at the ‘root’ level, then one can target the test programme towards finding precise estimates of statistics which pertain specifically to those same ‘root’ conditions/events. In combination with reasonable worst-case estimates (or independently verified estimates from elsewhere in the test programme) for the rates of incidence of compounding triggering conditions beyond these ‘root’ events, one can more efficiently and cheaply attain the evidence required to grant or deny an approval. One might think of this overall as “finding the weakest link in the chain, and establishing *with high confidence exactly how weak it is!*”

## 5.9.4 Coverage

‘Coverage’ can seem a bit of a buzzword when it comes to discussions around test programmes. Fundamentally, coverage is the notion of how well ‘explored’ the many different (i) system ‘behavioural competencies’, and (ii) operational situations (scenarios at the validation level) are upon completion of the test programme(s). For approval purposes, the intersection of (i) and (ii) form the relevant ‘problem space’ – it is this space which must be adequately covered.

Coverage is made yet more ambiguous because it is used to refer to proper treatment of test cases at the very bottom, all the way to the very top, of the right-hand-side of the engineering V-cycle. There are consequently and necessarily many different measures of ‘coverage’, both at a single point in the development process, and across the process as a whole. Defining acceptable and robust measures of coverage (a.k.a. coverage metrics) is, for example, less non-trivial at the component-level, and extremely non-trivial at the SBT and system-/vehicle-validation level.

UL 4600 (2020) is one of the few documents which has attempted to deal with the coverage question regarding ADS V&V. Its authors argue that the “V&V [activities] shall provide acceptable coverage of safety-related faults associated with [...]”:

- (i) “the design phase”;
- (ii) “the construction of each item instance”;
- (iii) “the item lifecycle”;
- (iv) “the item structure and intended operations”.

More granular details are given in the various subclauses (of UL4600 section 12.3.x), which broadly fall under the categories of quality assurance; conformance to specification; calibration; and systematic maintenance, migration and handling of defects. This report does not seek to contradict any of these items, which are omitted for brevity here, and anyway are anticipated in large part to form integral parts of existing manufacturing and safety assurance processes (including Functional Safety, ISO 26262, and SOTIF, ISO/PAS 21448, analyses).

The elephant in the room is the language of “shall provide acceptable coverage”. UL4600 does far less to elaborate upon what defines “acceptable” in this context. Although in UL4600 12.2.1.1-a.2, a “description of what work products [V&V activities shall] produce” is mandatory (treated by this report in Section 5.9.4), it goes no further than to mandate that an accompanying coverage strategy be identified. The lone examples given thereafter of (i) peer review and (ii) unit-testing are appropriate only to component and sub-function levels of V&V; they do not scale in any way to SBT and validation test

modalities. Insofar that “acceptable coverage” is thereby *implied* by adherence to, and application of, the coverage strategy/metric identified, this report has considered what those very coverage strategies and metrics should look like at the SBT and validation level of the test programme.

### 5.9.4.1 Sufficiency of test coverage

There is no single ‘correct’ or clearly superior way in which to measure coverage of the problem space as defined in Section 5.9.4. There are several challenges which might impede the inception of a ‘general’ measure of coverage, insofar that:

- (a) Every make/model of AV will have different functionalities, and a unique ODD and TOD, at least in the early-to-mid era of the adoption of AVs at large. This means that the operational, tactical and environmental parameters to which any given ADS is sensitive (with respect to its driving behaviours and therefore test outcomes) will never twice be exactly the same sets.
- (b) A moderate amount of test data and approvals history will likely need to be attained before clear macro-patterns emerge, after which the most inferential, efficient, and indicative measures of coverage will become clearer by their extensive real-world validation. Early vehicle and deployment safety cases should employ several independent measures of coverage to add an element of redundancy/contingency. This is in a similar spirit to reporting the outcome of a test with several quantitative scores, rather than few or a single score; and producing estimates for an array of population statistics in the outputs of a test programme, rather than only few.

Recall that in the Functional > Logical > Concrete scenario hierarchy, logical scenarios identify testing ranges for their  $n$  sensitive parameters. If each  $j^{th}$  parameter is represented by a normalised axis,  $\hat{x}_i$ , then the problem space can be modelled as an  $n$ -dimensional unit hyperspace (where all parameters belong either on the interval [0, 1] or on the interval [-1, 1]). This permits the usage of quantitative, as well as qualitative, and particularly statistical, coverage metrics.

Some potential ‘stem cell’ measures of coverage, which manufacturers and regulators alike may choose to use or adapt, are given in Table 35.

Potential SBT Coverage Metric	Notes and demonstrative example(s)
<b>Maximum <math>n</math>-d Euclidean (Pythagorean) distance between concrete scenario test cases.</b>	Effectively, this statistic $D$ is given by, $D = \max_{\forall k} \left\{ \sum_j \sum_{i=1}^n (x_i^k - x_i^j)^2 \right\}$ Explicit evaluation of $D$ may be very computationally expensive and so <b>numerical methods</b> to estimate its <i>upper bound</i> may reasonably be employed.
<b><math>k</math>-d Trees</b>	$k$ -d Trees are a special type of binary tree, often used for space-partitioning in multi-dimensional problems or creating point clouds in computer science applications. If utilised in a particular fashion, they offer an efficient way to test for proximity between large and highly dimensional data arrays (basically exploiting the point cloud properties). Figure 40 gives a visual flavour of how $k$ -d Trees might be employed to measure coverage. Figure 41 indicates their efficiency versus more basic/crude alternatives.
<b>Maximum Gaussian Process Regression (GPR surrogate model) uncertainty when conditioned on the test cases and their <i>quantified</i> outcomes (VeriCAV, 2021).</b>	Deep & Active Learning applications make use of surrogate data models in several cases in order to establish trends, patterns, contours, and transitional or limiting thresholds in extremely large and often highly-dimensional data sets. The same concepts and tools could be applied as a LSAV test programme coverage metric. In essence, the uncertainty (variance,

	<p>standard deviation) distribution of a surrogate GPR model across the entire problem space could itself be used as a statistic, and its maximum: (a) value could be used to determine whether sufficient coverage has been attained; (b) vector position could be used to determine where more test cases should be executed.</p> <p>This same notion could be extended to a set of top-5 local maxima, for example.</p>
--	--

Table 35: Potential coverage metrics for use or adaptation in SBT.

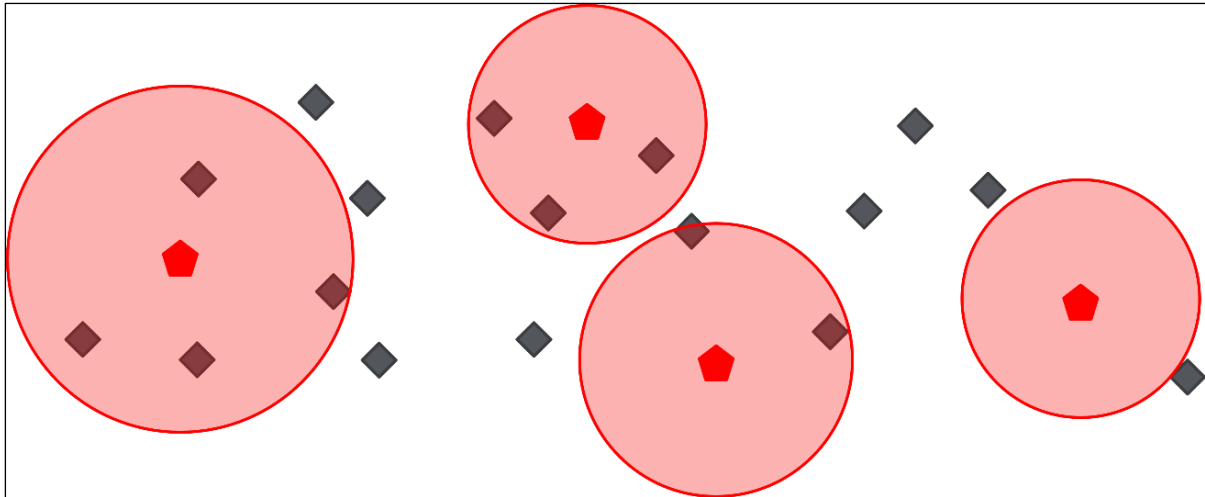


Figure 40: A 2D Visualisation of how k-D Trees can be used, with an array of configurable “radii”, to perform a metricised test of scenario coverage. The array of grey points represents the set of completed test cases (whose k co-ordinates are the initial values of the k logical scenario parameters which uniquely define/identify each concrete test). The array of red points and shaded proximities represent the ‘coverage critical’ domains (i.e., not test cases). The fraction of completed test cases belonging inside the critical domains is the output of the k-D Tree algorithm. This ‘contained’ fraction of all test cases may yield an indicative measure of the extent to which the ‘interesting’ scenario sub-spaces – be them safety critical or otherwise – have been adequately/appropriately sampled during the V&V activity.

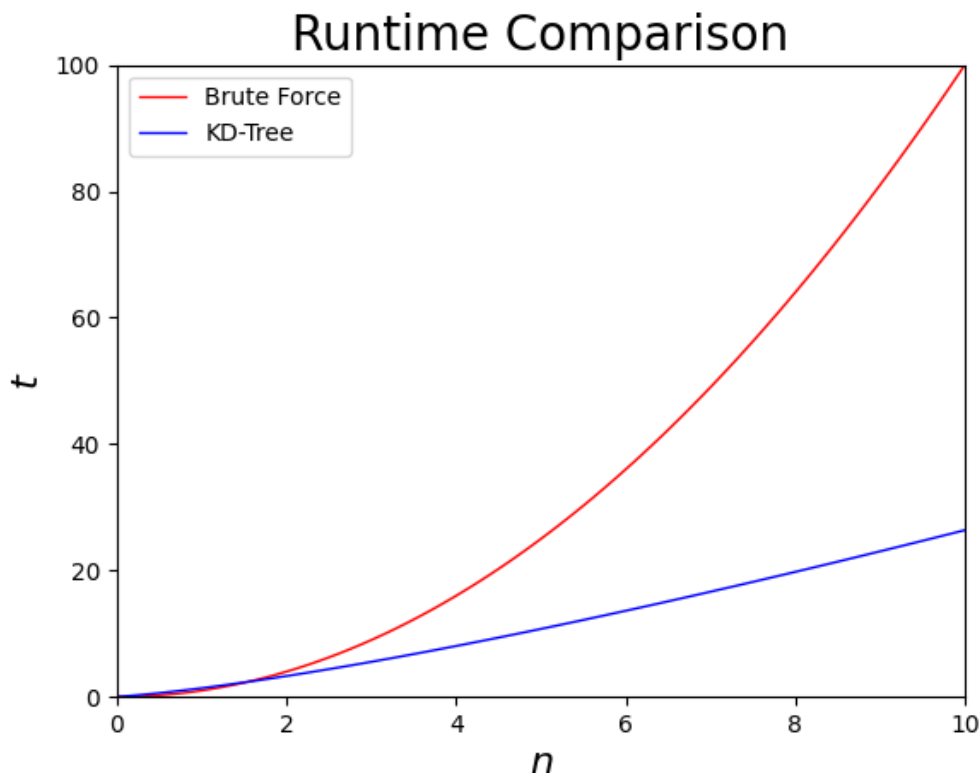


Figure 41: Runtime Comparison to demonstrate the efficiency savings brought by k-d Trees versus a “brute force” calculation approach.

## 5.9.5 Balance between Manufacturer / Developer Testing and Independent Testing

Pragmatically speaking, the majority of the testing burden falls upon developers and manufacturers. There shall be a requirement that an independent, trusted authority conducts or witnesses a random subset of tests within the programme for the purpose of gathering unbiased evidence that the tests are being carried out as otherwise stated by the developer in their test programme outputs (Section 5.9.2.2), and in a manner in-keeping with their established SMS (see Section 7.1). The evidence gathered shall provide assurance that the tests are being appropriately and safely conducted, in an as far as possible reproducible manner, that shall yield valid results. This requirement is intended to uphold and bolster the quality and fidelity of the safety argumentation evidence.

This is perhaps more readily realised for physical test modalities than it is for virtual/remote ones, such as simulation, since regulators may invoke their existing familiarity with witnessed type approval testing. Therefore, unlike existing arrangements for conventional vehicles, this report further recommends that manufacturers should be required:

- (i) to provide, upon request by The Department, the regulator, or the approval authority, a black-box driving model of their ADS which, given certain sensory (mock) input data, returns (a) the OEDR history/trace observed by the ADS; (b) the tactical decisions (dynamic actuation events) demanded by the ADS in response to (a);

and under certain circumstances,

- (ii) to execute safety-critical simulated tests in more than one environment / engine.

These requirements are recommended in the spirit of replicating ‘simulated witnessing’ of tests, without manufacturers being required to reveal the sensitive, inner details and workings of their ADS. Anyone with appropriate permissions and access to this model could then independently test or experiment with the model subject to their own physics-handling. Reproducing the same test *outcomes* via several different test modalities, and demonstrating repeatability intra-modality, is one way to increase

confidence in the fidelity, and therefore argumentative and inferential value, of all the V&V evidence derived from the test programme.

There shall therefore be a further requirement for the test *outputs* to include sufficient evidence of validation of all the test modalities utilised within the test programme (set-up and apparatus; both physical and non-physical). The single exception to this is for physical testing enacted upon the actual deployment route, since the fidelity of the true deployment is clearly self-fulfilled. This does not remove or contravene the requirement for sufficient coverage of the actual deployment (the TOD, see Sections 4.2, 5.9.4 and 5.9.4.1) to be attained; for example, exposure to all weather conditions, varied traffic densities, etc. Rather, it means that the requirement to validate the tools that constitute the test modality apparatus (e.g., software and hardware) applies by default, but that the actual deployment route (the TOD) does not need to have its fidelity qualified, for obvious reasons.

The need for a robust correlation analysis is further elaborated upon in Section 5.9.2.1 (c). The mechanisms by which the fidelity of and correlation between modalities should be qualified and asserted are anticipated to overlap significantly with those same mechanisms as outlined in Section 5.9.3.2; in other words, the same approach may be taken to validating test outcomes and behaviours between modalities as is taken to validate those very behaviours against the performance requirements, acceptance criteria and behavioural competency 'benchmarks', at least within a statistical domain.

Software tool validation and qualification is also expanded upon in Section 5.1.2.

## 5.10 System Monitoring

System Monitoring is already present in current vehicles as the status of critical systems is presented to drivers in order to take appropriate actions. With the advent of more electronic equipment on vehicles, tasks that would previously have required manual checks are now performed by sensors and software (e.g. tyre pressure monitoring).

Without a driver in charge of an ADS, it is not sufficient to present the information to an on-board display and expect the driver to handle the situation. To ensure safe operation of the vehicle, the ADS has to be able to detect failures that would have previously been detected by or notified to the driver, unless there are operational procedures in place to ensure that all components are in the appropriate condition to operate safely for the entire trip duration before each journey. The decision on which systems and components need to be monitored during run-time and what is to be addressed with operational procedures should be determined by the manufacturer based on analysis of its system implementation, available quality or reliability data and technical expertise. This shall be documented as part of the safety and security/ operation manual. The activities that are to be carried out to support these requirements should be documented in a manufacturer's Safety Management System and might be based on guidance from ISO 26262, ISO/PAS 21448 and other appropriate standards and guidance.

Security measures should also feed into the monitoring concept, and care must be taken to ensure that safety and security mechanisms are compatible.

The requirements related to system monitoring functionality required by the ADS are captured as proposed technical requirements 22 and 23 and also covered by item (9) in the Safety Case requirements. The requirements to ensure, during deployment, that adequate vehicle and system checks are performed are captured as part of the Deployment SMS requirements.

Additionally, work that is ongoing in ISO/TC 22/SC 32 WG8 is highlighted for future reference. Although scoped around semiconductors, Technical Report TR9839 is looking to capture best practice for predictive maintenance of EE hardware (semiconductor) to provide guidance on safety mechanisms to protect against intermittent faults. This concept could be extended to higher level vehicle systems – for example, the system could monitor wear and tear of mechanical steering components, resulting in degraded functionality (transition to a 'lower' MEL), or inhibition of operation altogether, coupled with notification being provided to maintenance staff, if the response characteristics drift outside tolerance. Such strategies will be vital to protect against risks resulting from faults remaining undetected ('latent faults'), and should be considered and documented within the functional safety analysis. Work Package 4 of this project considers safety of non-ADS aspects of the vehicle in further detail.



## 5.11 System Updates

The vehicle approval process set out in Section 3.1 sets out requirements to monitor the AV behaviour in operation and have mechanisms in place in order to react appropriately, either:

- in case of incidents with remedial action or
- to implement planned changes to the functionality.

Responsibility for a remedial action that results in an update to the vehicle design (software or hardware) must be assigned to the vehicle manufacturer.

A key activity when determining updates is to ensure that they are aligned between operator and manufacturer such that any design changes do not negatively affect operational procedures. The updates also need to be assessed for their impact, and, if determined necessary, iterated through the relevant Type approval review.

UNECE Regulation 156 addresses software updates in type approval, including over-the-air technology that facilitates a way of modifying vehicle performance that could be considered 'invisible' and go unnoticed by vehicle owners or authorities. The approach behind this regulation requires manufacturers to notify Type Approval authorities if a software update affecting a technical requirement for type approval regulations is performed, triggering a review by the Type Approval authority which could result in further evidence being requested. Based on the overall evidence provided, the authority might either grant an extension to the type approval or consider the existing approval to maintain its validity.

For an automated vehicle, the recommendation is that any system that forms part of the ADS (including sensors, actuators, telematics, HMI or vehicle body controller) that interfaces with vehicle occupants should be managed under the SUMS framework.

Additionally, the regulation requires manufacturers to be "in control" of their software, meaning to be able to identify appropriate software versions with respect to particular vehicle and system versions, architectures and variants. This is to be ensured by having a software update management system (SUMS) implemented, which is audited and has to be approved before a type approval application may be made. As part of the regulation, the manufacturer is also required to declare their compliance with the rules set out. Already considered in the regulation is the strong link between software updates and cybersecurity, and also to system and functional safety, in order to ensure that faulty, inappropriate or missed software updates do not compromise safe operation of the vehicle.

For an automated vehicle, the recommendation is that any system that forms part of the ADS (including sensors, actuators, telematics, HMI or vehicle body controller) that interfaces with vehicle occupants should be managed under the SUMS framework.

In extension to the existing Regulation 156, it is recommended that both a software and hardware/component update management system is part of the safety management system (SMS) that both the manufacturer and also any operator have to have in place— see Section 7.1 for more information on how an SMS should be implemented. This report assumes that the operator's responsibility includes identification of changes in the TOD that require a modification to the AV while the manufacturer is responsible for the development and approval of any update. The responsibility for performing the update could be either the operator or manufacturer. It is recommended that it is confirmed by an approval authority that this responsibility has been agreed between the involved organisations – as a cross-check, this could be performed as part of the pre-deployment type approval, and again at the deployment approval.

Separate from the UNECE regulation, there is guidance being prepared for best practice for software update engineering, with an international standard currently at DIS stage (ISO /DIS 24089). Its scope covers guidance for software update engineering for road vehicles on both organisational and project levels, as well as providing guidance on the deployment of software update packages to road vehicles. The organisational and project level guidance might be referred to when defining a SUMS, and is also recommended to be used as assessment criteria during a type approval assessment.

The requirements set out in ISO/DIS 24089 cover the following topics that, in combination, support an organisation in managing and controlling their products during operation:

- Data Sharing Policy

- Data Management Process
- Continuous Improvement Process
- Document management
- Requirement management
- Configuration Management
- Quality Management
- Change Management

It can be seen that these topics overlap with management system requirements defined in other standards (e.g. ISO/SAE 21434 or ISO 26262). It would be advisable for an organisation to ensure that these supporting processes listed above are fit for purpose to underpin all aspects of organisational risk management (system safety, functional safety, operational safety and cybersecurity), and it is proposed that these requirements are assessed jointly for that reason.

It is noted though that certain aspects of the guidance on deployment of software updates might be tailored for the vehicles targeted by the approval scheme, taking into account a specific update infrastructure or update route that is in place for the duration of the operation of these vehicles.

## 5.12 Proposed Technical Requirements for GB Approval Scheme

Table 36 shows the requirements proposed in WP1 for the GB approval scheme, based on the high-level hazard log and safety goals from Section 3.3 and refined for the abstract high-level system architecture assumed for a LSAV super system (Section 3.4). Additional requirements identified from the review of the UNECE (2022) and EU (2022) proposals have also been added.

These requirements aim to set out the behaviour and characteristics that an ADS has to achieve and demonstrate for type approval. Additional requirements will need to be specified to give more detail to manufacturers on what is expected to be submitted for type approval. These requirements have been further developed Section 7.2.

The numbers in [] brackets indicate the Safety goal ID in Section 3.3. It can be seen that Safety goals 5, 13, 14 and 15 are not explicitly listed. This is because they are only appropriate when particular road layouts are present. Rather than specifying the expected behaviour separately for each behavioural competence in detail, the manufacturer will be required to declare what is required within their target ODD and demonstrate that their ADS is able to perform the required functionality safely. So, for example, in case of intersections within the ODD and a route incorporating a right turn, this would call on Safety goal 15 – When turning follow right of way rules.

Proposed Technical Requirement on ADS		Comments and Explanation
<p>Requirements on how the LSAV (being controlled by the ADS) is expected to interact safely with other road users</p> <p>Requirements on how the LSAV (being controlled by the ADS) is expected to move safely within the road infrastructure</p>	<p>1</p> <p>The ADS controlling the LSAV shall perform the DDT such that the LSAV</p> <ul style="list-style-type: none"> <li>- does not cause collisions [1]</li> <li>- is able to avoid foreseeable collisions [2] and</li> <li>- protects all persons within and in the vicinity of the vehicle [3]</li> </ul> <p>when operating within its ODD.</p>	<p>This is proposed to be assessed by reviewing submitted documentation (as part of the safety case) describing the design of the ADS and the elements of the ODD and the validation evidence (see further details on assessing V&amp;V evidence in Section 5.9). The acceptance will be determined through evaluation of whether the defined functionality achieves the required behavioural competencies in the context of applicable scenarios that are expected to occur in the ODD and ultimately the TOD, considering</p> <p>(A) Have sufficient capabilities been declared, e.g. are lane changes required, are there crossings or intersections?</p> <p>(B) Is the design appropriate to perform the required functionality with respect to the</p> <ul style="list-style-type: none"> <li>- Sensing functionality : This needs to allow for innovation but enable the regulator to make a judgement that the perception system is (1) appropriate to the requirements of object detection within the ODD, (2) an appropriate type of technology, e.g. cameras for object classification/ colour perception to be fitted where required and (3) includes appropriate sensor ranges and field of views.</li> </ul> <p>This also includes having identified additional functionality to support perception systems (e.g. camera cleaning systems, adjustment and monitoring of the perception system functionality). Any applicable specific requirements for perception systems using ML can be found in Section 5.8.</p> <ul style="list-style-type: none"> <li>- Planning functionality: The planning functionality needs to be able to determine:                     <ul style="list-style-type: none"> <li>(A) the appropriate behavioural competences to execute in each situation,</li> <li>(B) the required lateral and longitudinal control for each of the individual capabilities and</li> </ul> </li> </ul>

Proposed Technical Requirement on ADS		Comments and Explanation
		<p>(C) also, in case of conflicting events, which action to prioritise.</p> <ul style="list-style-type: none"> <li>- Actuation functionality: The steering, braking and propulsion system must be able to support the lateral and longitudinal control functionality required for the DDT including maximum and minimum acceleration and deceleration rates and achievable steering radii. This also includes ensuring the capabilities of the actuators are appropriate and have considered the MRMs that have been defined, even in the case of failures of any actuator component, e.g., by the design of safe state provisions or emergency operating provisions built into their design.</li> </ul>
2	<p>While performing the DDT in nominal traffic scenarios the ADS shall</p> <ul style="list-style-type: none"> <li>- Travel in a stable lateral position within the appropriate lane [12]</li> <li>- Maintain appropriate safety margins to other road users [18]</li> <li>- Drive considerately and follow the rules of the road/traffic rules [9] and [4]</li> <li>- Adjust vehicle speed to prevailing conditions and for occupant safety [7], [11] and [21]</li> <li>- Follow prevailing driving styles by avoiding behaviour that is not expected by other road users [19] and that avoids obstructing traffic [16], [20]</li> <li>- Interact safely with other road users including providing appropriate signalling of intentions and providing information [10] and [17]</li> <li>- be able to drive in the reverse direction (reverse gear) [6]</li> </ul>	<p>Acceptance will be based on evidence of whether the vehicle is kept inside its lane of travel (outer edge of front tyre to outer edge of lane marking) at all times, unless there is compelling reason not to, e.g. vehicle is changing lane, or an evasive manoeuvre is taking place.</p> <p>The manufacturer shall declare the minimum distance that the ADS maintains to objects travelling ahead and in parallel lanes to the LSAV. For lead vehicles, travelling ahead in lane, UNECE Regulation 157 (ALKS) requires 1.3s time gap or 10.8m @30 kph; this should be considered the minimum, but to ensure that harsh braking can be avoided, a larger distance may be necessary. ISO 22737 defines a maximum deceleration of 4.9m/s<sup>2</sup> for MRMs in case of standing passengers, which would require a larger minimum gap of at least 1.7s.</p> <p>Part of the evidence submitted should also be reviewed for the manufacturer's approach to prioritisation of safety objectives in case conflicting constraints develop within a scenario (e.g., presence of a overtaking vehicle laterally close and pedestrian walking along road edge).</p>

Proposed Technical Requirement on ADS		Comments and Explanation
3	<p>The ADS shall detect and respond to all objects and events declared as part of the specified ODD that are required for the execution of the DDT. Objects and Events might include but are not limited to [VRU, vehicles, precipitation, road markings...]-</p> <p><b>For information on ODD specification see Section 4.1.</b></p>	<p>The manufacturer must demonstrate an appropriate combination of sensor technologies, considering the performance limitations of each and ensuring they complement each other. Considerations also need to be given to external factors affecting detection, e.g., occlusion of objects by other objects or through environmental factors depending on the parameters of the ODD.</p> <p>Perception guidance should facilitate continued innovation and application-specific solutions, but enable the regulator to make a judgement that the perception system is appropriate to the requirements of object detection within the ODD.</p>

Proposed Technical Requirement on ADS		Comments and Explanation
4	<p>When encountering critical scenarios, the ADS performing the DDT shall</p> <ul style="list-style-type: none"> <li>- Be able to detect collision risks with other road users or unexpected obstacles and</li> <li>- Perform an appropriate emergency manoeuvre to minimise risks to safety of the vehicle occupants and other road users. [2]</li> <li>- Prioritise human life while reducing damage and losses [8]</li> </ul>	<p>This is also expected to be assessed by reviewing submitted documentation (as part of the safety case) describing the design of the ADS and the elements of the ODD. The acceptance will be determined through evaluation of whether the defined functionality achieves the required behavioural competencies in the context of scenarios that are expected to occur in the ODD and ultimately the OD.</p> <p>Minimum behavioural competencies to be demonstrated include</p> <ul style="list-style-type: none"> <li>- being able to detect VRUs travelling in lane or approaching the lane either with an intention to cross or join, while considering occluded areas. As part of this the manufacturer must declare the assumptions made about behaviour of VRUs, e.g. ISO 22737 (2021) defines an approach speeds pedestrian 8 kph/ cyclist 25 kph, and show its appropriateness in the safety argument. Note that this should consider the full range of permutations possible within the TOD – as such, it is strongly recommended that much higher speeds than those listed in ISO 22737 should be catered for, the ISO standard being significantly inadequate to assure safety in this regard.</li> <li>- avoidance of collision risk obstacles in lane (i.e. objects that, in case of a collision, result in harm to occupants like stationary vehicles, large animals, large lost cargo).</li> </ul> <p>Meeting the objective of this requirement could be achieved with a different implementation if a particular vehicle was specifically designed with only goods transport or only passenger transport in mind. A passenger vehicle would need to consider the safety of passengers onboard while a goods vehicle would always prioritise VRU and other road users' safety.</p> <p>This requirement should also be considered in combination with requirement 2 as the prioritisation of safety objectives in case conflicting constraints applies to both nominal and critical scenarios.</p>
5	<p>The ADS shall be able to detect when a collision has occurred and stop and secure the vehicle. The ADS shall be deactivated until it is verified that the vehicle (including ADS) is able to proceed safely.</p>	<p>The manufacturer shall describe and show evidence of their implementation to show that it meets the intent of this requirement, and a demonstration of this functionality could be included in potential witnessed testing or during initial supervised trial operation.</p>

Proposed Technical Requirement on ADS		Comments and Explanation
6	<p>The ADS shall detect all reasonable events where safety goals have been violated or that could lead the violation of a safety goal.</p> <ul style="list-style-type: none"> <li>- activation of collision avoidance mechanisms that successfully mitigate a collision risk</li> <li>- minimum distances / safety envelopes not maintained</li> <li>- Highway code violations</li> <li>- situations where the ODD has been exited (particularly for ODD parameters that are of binary nature)</li> <li>- situations where a MRM has been performed</li> <li>- situations where a fault in a vehicle system implementing functionality related to the DDT has been detected</li> <li>- situations where the technical oversight was required to intervene</li> </ul>	<p>The manufacturer shall describe and show evidence of their implementation to show that it meets the intent of this requirement, and a demonstration of this functionality could be included in potential witnessed testing or during initial supervised trial operation.</p>
7	<p>The maximum speed at which the ADS is permitted to operate the LSAV is 20 mph/ 32 kph.</p>	<p>The manufacturer shall describe and show evidence of their implementation to demonstrate it meets the intent of this requirement, and a demonstration of this functionality could be included in potential witnessed testing or during initial supervised trial operation.</p>
8	<p>Activation of the ADS shall only be possible when the conditions are compatible with the System Deployment Capability Definition (see Section 4.1.2.1), i.e. the surrounding conditions and fault status of the vehicle are compatible with a permitted combination of TOD and MEL respectively. Activation status shall be recorded.</p>	<p>This requirement results from the necessity to know that the ADS is in control of the vehicle in order to comply with the recommendations of the Law Commissions' report regarding ADS activation. The manufacturer shall describe the implementation that ensures the intent of this requirement is met, and a demonstration of this functionality could be included in potential witnessed testing or during initial supervised trial operation.</p>

Proposed Technical Requirement on ADS		Comments and Explanation
	<p>9 The ADS shall be able to monitor the parameters of its COD and react safely to reaching and exceeding conditions of the TOD by performing an MRM to reach an MRC.</p>	<p>The manufacturer shall declare the strategies taken in case TOD boundaries are reached or exceeded, or fault conditions affecting monitoring of TOD conditions occur. These requirements and criteria for TOD and deployment domain specification are discussed in more detail in Section 4.</p> <p>There shall be at least one MRC and one MRM, but a manufacturer may specify a more differentiated approach for different conditions. These need to be presented with evidence that they are suitable and appropriately safe in the conditions where they are triggered.</p>



Proposed Technical Requirement on ADS		Comments and Explanation
10	<p>The ADS controlling the LSAV shall use appropriate signalling functionality to</p> <p>(1) indicate its intention to other road users where required</p> <p>(2) ensure the vehicle is visible to other road users, including signalling its position, orientation, and current status.</p>	<p>The manufacturer has to show the following functionality, tailored to their application, is implemented such that</p> <p>(1) is achieved through the use of the applicable turn indicators. The vehicle shall indicate its intention as required by the Highway Code and with appropriate timing to ensure other road users can anticipate the LSAV’s intentions.</p> <p>Necessary situations where indication is required (as per UK Highway code, other situations may be appropriate):</p> <ul style="list-style-type: none"> <li>- changing lane / merging into or exiting a lane</li> <li>- when temporarily crossing a lane boundary (e.g. to pass a partial obstruction of the lane)</li> <li>- turning (left or right) (including roundabouts)</li> <li>- when leaving a carriageway (e.g. to enter a stop)</li> <li>- when setting off from a stop</li> <li>- when making way for emergency vehicles</li> </ul> <p>Appropriate timing can be based on time or distance before a particular manoeuvre is initiated either</p> <ul style="list-style-type: none"> <li>- before a change of velocity (before braking is initiated)</li> <li>- at least 30 m before a junction, more at higher speed (up to 250m at motorway speed – not relevant for application) [consideration for road layout with multiple turnings)</li> </ul> <p>((2) is achieved through front headlights, rear lights, reversing lights and warning lights and headlights in daylight condition in reduced visibility condition. If applicable within the ODD, the ADS shall turn on dipped headlights in conditions where visibility is reduced. Reduced visibility situations may include the presence of rain, fog, snow, sand, dust, or ash. Although it is expected that operation in fog would not be part of initial applications of these types of vehicles, rear fog lights should only be activated if conditions of severely restricted visibility conditions are met.</p>

Proposed Technical Requirement on ADS		Comments and Explanation	
		<p>The ADS should also not use full beam (if fitted) in case of oncoming vehicles – detection of oncoming vehicles, cyclists or VRUs should result in dipped beam.</p> <p>In case of a single failed headlight, the LSAV might continue while aborting its journey (e.g. drop off remaining passengers and return to depot) if the manufacturer can show that sensor functionality is maintained sufficiently across the TOD. If required, a reduced speed operation may be possible – in other words, the system would enter a 'lower' MEL, and therefore use a more restricted TOD and/ or behavioural competencies definition accordingly. Full loss of headlights or rear lights shall result in an MRM into a safe position with hazard lights activated.</p> <p>Brake lights and reverse lights are to mirror current functionality when the ADS is decelerating the vehicle or controlling its direction instead of the driver.</p> <p>Warning Lights are to be activated in case of any MRM execution or if the LSAV finds itself physically stuck, unable to move or obstructing traffic.</p>	
Requirements on how the ADS and other vehicle systems shall interact with occupants to ensure their safety	11	The ADS shall be designed to protect occupants during operation.	This requirement is intended to link to the occupant safety provisions derived in Work Package 4 of this project, which examines non-ADS aspects of the vehicle - considerations should be given whether the provisions on the vehicle are aligned with the design of the ADS (for example maximum permitted acceleration and deceleration limits during normal operation in case of standing passengers).
	12	The ADS shall not start a journey until it has ensured that occupants have safely boarded the vehicle.	

Proposed Technical Requirement on ADS		Comments and Explanation
13	The ADS controlling an LSAV engaged in passenger operation shall ensure it is stationary at an appropriate stopping location before passengers may board or disembark the vehicle, unless an MRM has been executed or an emergency stop request has been triggered.	<p>The Law Commissions consider securing of load as part of oversight duties, but ensuring that the doors are closed before the vehicle moves off would fall under the responsibility of the ADS.</p> <p>The manufacturer shall describe the implementation that meets the intent of this requirement, and a demonstration of this functionality could be included in potential witnessed testing or during initial supervised trial operation.</p> <p>Whilst it is noted that an MRM may justify stopping in a location that otherwise wouldn't be appropriate, nonetheless this may not be desirable behaviour, and it should be considered whether it would be feasible for the vehicle to reach an MRC in a more appropriate location.</p>
14	The ADS controlling an LSAV shall provide all necessary safety information to the occupants in a clear and unambiguous manner. (e.g., operating status, intention to move off, next stop indication, how to request a stop, what to do in an emergency, that an MRM is being activated)	<p>The manufacturer will be required to describe their functionality and HMI design, and the implementation could be observed during witnessed testing or during initial supervised trial operation. Best practice guidelines for Human Factors should be considered – see the guidance on Human Factors in Section 5.7</p> <p>The intent of requirement 16 is to link to the passenger safety requirements produced by Work Package 4 of this project, which require the provision of emergency stop buttons</p>
15	The LSAV shall provide a means for vehicle occupants to interact with the control centre.	
16	The ADS shall bring the vehicle to a standstill if requested by the vehicle occupants via the emergency stop.	
17	The LSAV shall provide a means for the remote oversight to monitor the LSAV's load or vehicle occupants (e.g., via camera), other than in cases where there are operational procedures that allow a human to perform this role.	

Proposed Technical Requirement on ADS			Comments and Explanation
Requirements on the necessary interaction between the LSAV and remote external supervision	18	The LSAV shall provide a means for off-vehicle remote assistants to support operation and passenger safety, other than where the operational procedures include the continuous presence of a member of staff within, or within visual line of sight and close proximity of, the vehicle	Requirements 18-20 facilitate the recommendation from the Law Commissions' consultation that every vehicle without a driver or user-in charge should have a licensed operator able to support the journey. The manufacturer shall describe the implementation that meets the intent of this requirement, and a demonstration of this functionality could be included in potential witnessed testing or during initial supervised trial operation. This requirement is formulated as an objective to allow for different solutions to be possible. The assessment needs to carefully consider the trade-offs between higher dependability on off-board functionality and required safety and security properties. The design of the interface should follow best practise guidance for Human Best practice should be followed with regards to Human Factors (Section 5.7), Cybersecurity (Section 5.3) and External Inputs (Section 5.6)
	19	The ADS shall provide the necessary information to the control centre operator to carry out their duties.	
	20	If there is a means for a remote assistant to influence the DDT, the manufacturer shall provide evidence for its safe implementation and communicate any conditions or restrictions clearly to the operator.	
	21	If a remote or onboard assistant is required to intervene for any reason, then the appropriate information or system status shall be provided to them.	
Requirements on the ADS and non-ADS vehicle systems	22	The ADS shall monitor the health and status of all vehicle systems that are involved in the execution of the DDT.	

Proposed Technical Requirement on ADS		Comments and Explanation
23	The ADS shall react safely to faults that affect the LSAV's ability to drive safely by executing an appropriate MRM to reach a suitable MRC or by transitioning to a degraded mode of operation (i.e. to a different MEL, potentially resulting in a more limited TOD or set of behavioural competencies).	<p>Without a driver in charge the ADS has to be able to maintain safe operation throughout the lifetime of the vehicle (meaning it must accommodate ageing effects of non-ADS vehicle systems while they remain within defined tolerances. When these tolerances are exceeded, the ADS must detect this as a failure (which would have previously been detected by the driver), unless there are operational procedures in place to ensure that all components are in the appropriate condition to operate safely before each journey and suitable assurance can be given that rapid degradation will not occur mid-journey. The decision on which systems and components need to be monitored during run-time and what is to be addressed with operational procedures should be determined by the manufacturer based on safety and failure mode analysis, reliability calculations and technical expertise, and documented as part of a safety and security/operation manual. The activities carried out to support these requirements should be documented in a manufacturer's Safety Management System and might be based on guidance from ISO 26262, ISO/PAS 21448 or other appropriate standards and guidance.</p> <p>One possible additional mandatory requirement could be that no single point of failure shall result in the loss of braking functionality unless there are additional measures to ensure sufficiently low probability of occurrence. This is to ensure the MRC of "stationary vehicle" can be achieved in all cases. This requirement needs to also be assessed jointly with the MRM(s) the manufacturers sets out, to ensure that malfunctioning behaviour is taken into account in their design, resulting in appropriate design of backup systems where required to provide fail-operational performance. If this is the case, the ADS needs to be shown to adapt its behaviour to the presence of faults such that continued safe operation can be achieved.</p> <p>Security measures triggering should also feed into the monitoring concept, and care must be taken to ensure that safety and security mechanisms are compatible.</p>
24	The ADS shall be free of unreasonable risk from hazards associated with the intended functionality and its implementation, including both hazards due to failures and due to insufficiencies of specification or performance insufficiencies.	
25	The ADS shall be protected from unauthorized access.	This requirement is taken over from the EU draft proposal.

Table 36: Proposed Technical Requirements for GB Approval Scheme.

In order to for the regulator to assess the performance requirements a number of supporting requirements need to be established that set out the process and evidence required to achieve the performance requirements in more detail. These are further discussed in their individual sections but are outlined in principle within Table 37.

Placeholder manufacturers and operators	Requirements for	Evidence required and Acceptance Criteria	
Requirements on manufacturer	26	<p>The manufacturer shall declare the competencies of the ADS.</p> <p>(see the content on competencies described in Sections 3.3 and 4.3)</p>	<p>A minimum set of competencies must be achieved, which is proposed to include all of the operational/control level competencies:</p> <ul style="list-style-type: none"> <li>- maintain lateral/longitudinal position in lane</li> <li>- follow another vehicle</li> <li>- collision avoidance</li> </ul>
	27	<p>The manufacturer shall define the ODD and TOD (as per domain specification requirements - see Section 4.1).</p>	<p>See further information in Section 4.1. The purpose is that the specification of the ODD must be detailed enough to facilitate the assessment of the safe behaviour of the ADS and also that compatibility of the conceptual ODD with any later physical target operating domain(s) (TOD).</p>
	28	<p>The manufacturer shall document how the capabilities they have declared are achieved by describing the</p> <ul style="list-style-type: none"> <li>- Static objects or elements and</li> <li>- Dynamic objects and actors used to determine the inputs required for each behavioural competency, and the</li> <li>- environmental conditions in which the capabilities can be executed</li> </ul>	<p>More information is provided in Section 4.1 – here the aim is that for the minimum required capabilities, the ODD and TOD information that is considered relevant as evidence at type approval is described in more detail.</p> <p>It is proposed that at type approval, the declared list of objects and events is reviewed and a minimum set of static and dynamic elements is used as a checklist by the approval authority, but not specified in the requirements.</p> <p>Exclusions of any dynamic element types must be justified and be appropriate for the declared DDT or addressed with operational procedures. This might include constraints on in-use operation (e.g. by technical oversight) or require in-use monitoring to be in place to ensure the continued validity of any exclusion.</p>
	29	<p>The manufacturer shall declare the strategies taken in case TOD boundaries are reached or exceeded, or fault conditions affecting monitoring of TOD conditions occur.</p>	<p>There must be at least one MRC and MRM of bringing the LSAV to a safe stop. Additional strategies might be implemented and must be shown to be appropriately safe.</p>

	30	The manufacturer of the vehicle shall specify a maintenance schedule for the vehicle, including requirements on regular checks, system tests and component replacements.	This is required to support the operator in maintaining the vehicle in an appropriate condition for safe operation.
	31	The manufacturer shall prepare a safety and security case	This topic is further discussed in Sections 3.2 and 5.
	32	The manufacturer shall implement a Safety and Security Management System	This topic is elaborated upon in Section 7.1.
Requirements on operator	33	The operator shall implement a Safety and Security Management System	
Safety and Security Manual/ Operating Manual / Safe and Secure Operating Manual	34	The manufacturer shall prepare a safety and security manual to support operation.	In order to ensure that the operator is aware of any constraints or assumptions made regarding the type of use for the vehicle, the target operating domain, or the operating organisation processes, there should be a means for this information to be communicated, and also reviewed at the point of approval.
	35	The manual shall include any instructions required for the ADSE, operators, occupants, service personnel, regulators and public authorities.	
	36	The manual shall document any external systems or communication links (including their safety and security properties) that are required to be available, constraints on the target operating domain that the low-speed automated vehicle is able to operate in, and operational	This topic is further described in Section 6.

		processes that are necessary to ensure safety during operation	
	37	The operational manual shall describe the expected response of vehicle occupants, transport service operator, on board operator and remote intervention operator and public authorities in case of failures and ADS request.	This topic is extended in Section 6.

*Table 37: Proposed requirements on manufacturers or operators.*

The proposed performance requirements were put out for comment in the second round of stakeholder consultation. The input sought from consultees was on the proposed approach that requires the burden of evidence to be carried by the manufacturer, who would need to show that the design and implementation of the ADS in their low-speed automated vehicle (LSAV) is safe for their specified ODD and TOD by explaining the design intent and demonstrating the achievement by validation. A minimum number of competencies would need to be achieved and demonstrated, but without definite targets for performance characteristics.

8 out of the 10 consultees provided feedback, with 5 responses showing a broad agreement and understanding of the current style of requirements. Responses also acknowledged that this would require good and early collaboration between manufacturers and approval organisation. This is covered by the approval process outlined in the report, which calls for early engagement.

From those responses that indicated a different preference, the sentiment that was expressed was not completely in disagreement with the proposal, as one response suggested a comparative approach rather than a quantitative approach (which had not been proposed and might have been misread by the stakeholder), while the other indicated a preference for actual specifications. It is acknowledged that future regulation may be able to set more specific requirements, but the report lays out a justification as to why this is currently not possible.



## 6 Operational Safety of Deployment

### 6.1 Considerations for Deployment

#### 6.1.1 Background

##### 6.1.1.1 Definition of Problem Addressed

In addition to safety assurance of the vehicle itself, safety of the deployment environment must also be considered. This includes aspects such as ensuring that hazards within the deployment environment have been identified and adequately mitigated, ensuring that staff with safety-related roles are aware of and capable of performing their duties, ensuring that the public are able to interact with the system in a safe and appropriate way, and ensuring that suitable processes are in place to react to emergencies should one occur. Pre-deployment approval of the operational safety will therefore be a key component of the regulatory process for confirming the acceptability of the operator's safety case. The deployment safety case should include consideration of what the vehicle is and is not able to do safely, as identified in the vehicle safety case, in order to inform the identification and mitigation of hazards.

It is recognised that there can be flexibility in the interpretation of what falls under the scope of a safety management system (SMS); this report draws a distinction between the SMS and other non-SMS elements of the safety cases in terms of the evidence available to audit, analyse or test at the time of approval. If tangible mitigations are able to be assessed by the regulator at the time of approval, these will be assessed directly as components of the VSCR and/ or DSCR. Processes to ensure that mitigations remain effective, and are updated and documented as necessary over time, would fall within the SMS, a sub-component of the safety case where the mitigations themselves are, naturally, not available to be assessed at the time of approval; therefore, SMS approval should instead be done on the basis of reviewing processes rather than reviewing tangible products or mitigations.

This section focusses primarily upon mitigations within the DSCR that are available to assess at the time of approval, but also covers some aspects of the operator's SMS. In contrast, Section 7.1 focusses upon the manufacturer's SMS, particularly with regards to establishing a strong safety culture and feedback process, in order for this approval of 'process' to complement the approval of the 'product' as set out in Section 5. Naturally, there is significant overlap between the SMS processes applied by the manufacturer and operator, and hence a certain amount of overlap between this section and Section 7.1, the differentiation being in the organisation and activity the SMS proposals are aimed at.

##### 6.1.1.2 Current State of the Art

BSI PAS 1881 (2020) sets out methods for assuring the safety of AV tests and trials, and therefore focusses upon operational measures that don't directly relate to the system safety, but instead provide an additional layer of protection that is able to prevent or mitigate harm through methods external to the system. It defines 'operational safety' as: "identification and management of all risks associated with completing any activities within the defined operating environment", and adds the note: "the measures put in place to ensure appropriate operational safety and security are influenced by the capabilities and safety of the system.... in addition to, for example, consideration of human factors or hazards in proximity to the vehicle".

The focus of PAS 1881 is upon what elements need to be included within an operational safety case in order to provide an acceptable level of safety assurance within trials. Although explicitly limited to trials, the scope includes trials where there is not a traditional safety driver in the vehicle – such solutions could range from a safety operator in the vehicle with non-traditional controls, to continuous remote supervision, to the absence of any safety operator whatsoever. As such, the scope extends to trials that would look very similar to full commercial deployments, and therefore the guidance can be seen as having some relevance. Caution should, however, be used, as the exposure to hazards, and therefore

the level of analysis and mitigation that is proportionate, will be significantly different if a commercial deployment is on a far larger scale than a research and development trial.

The PAS identifies particular areas needing focus such as:

- Ensuring the ODD is sufficiently defined such that it can be used as an input to the operational safety case (note: the PAS pre-dates the concept of the TOD, presented in Section 4.1. However, in practice, the requirements specified for the ODD relate directly to the characteristics of the specific trial route(s) or area(s), and could therefore be seen as more analogous to TOD than ODD within the terminology of this report)
- Performing an operational risk assessment, for which the safety case should capture what methodology has been used to identify hazards and assess the resulting risks that are foreseeable for each scenario, including for hazards resulting from system errors and faults, external dependencies (e.g. wireless communications), the route, and other road users. It proposes that the risks should be made 'ALARP' (as low as reasonably practicable), although note that whilst this is widely regarded as an effective method of assessing risk acceptability, other methodologies such as the 'globally, at least equivalent' (GALE) method of comparing relative risk have also been successfully used (Hillman, 2021).
- Operational guidance such as method statements to document safe working practices, roles and responsibilities etc.
- Route selection and assessment to ensure that the route is compatible with safe operation of the system and that stakeholders such as landowners and local authorities are in agreement.
- Safe operation and control (in practice, this primarily assumes a safety driver physically present in the vehicle, and will be of limited relevance to deployments utilising remote assistants.
- Change control – much like for the systems safety case, this will be important for operational safety.
- Monitoring, reporting and continuous improvement - to ensure that there is a process in place such that incidents are identified and learnt from.

It should be noted that, at the time of writing, there is a revision of this document being undertaken by BSI, including public consultation and a steering group of relevant experts. This is expected to expand the scope of the document by introducing topics such as safety management systems, but is not expected to alter the fundamental principles that are relevant to this report.

While BSI PAS 1881 sets high level requirements for relating to safety drivers and safety operators, BSI PAS 1884 (2021) takes this further by focusing exclusively on this aspect of safety assurance. Again, the scope focusses upon trials rather than commercial deployments, but nonetheless, in the absence of references that are specific to the challenges of remote assistants within commercial deployments, PAS 1884 at least provides an approximate benchmark.

Clause 4.3 of PAS 1884 includes an absolute maximum duration for trialling without a break (2.5 hours) and an absolute minimum length of break (a 15-minute break after a duration of not more than 2 hours; otherwise, a 30-minute break). However, these requirements only provide for a basic level, and are effectively caveated by the requirement for each trial to consider the balance of risk and select appropriate policies for break periods on a case-by-case basis. No data or references are provided to support the trialling organisation in making this judgement – understandably so, as this is an aspect that has had minimal research with regards to safety drivers within research trials. This is as opposed to the very different challenge of members of the public being responsible for SAE level 2 or 3 systems, which has received more attention (TRL, 2021a), but arguably holds less relevance for trained professional operators working within a safety management system.

As a point of reference on break periods, the Office of Rail and Road (ORR, 2013) recommends 10 to 15 min breaks every 2 hours (or every 1 hour during the night) where tasks "require continuous sustained attention, with no natural breaks in the task and where a lapse in attention can lead to safety implications". This is well aligned with BSI PAS 1884, and provides a valuable benchmark as it would cover a wide range of safety-critical roles within the rail industry. The report notes the significant risk that fatigue brings, with at least 74 railway accident and incident reports between 2001 and 2009

considering fatigue to be a possible causal or contributory factor. Compliance with the Working Time Regulations 1988 is not in itself sufficient to adequately control risks from fatigue.

Whilst automation may be expected to eliminate many opportunities for fatigue to result in incidents, there will still be a need for humans to perform safety-critical roles, and the ORR report highlights how significant a factor the management of fatigue will be in such cases. The report sets out how there should be a 'fatigue risk management system' in place to cover any staff required to undertake shift work or overtime, and where such work is safety critical, this should be extended to cover management of fatigue within safety critical work.

Whilst aimed at trials upon CAM Testbed UK, the Zenzic safety case guidance (Zenzic, 2021) scope includes 'advanced trials' where no safety driver is present (and thus a high level of assurance of the system safety is required), and is intended to be of use not just to trials upon CAM Testbed UK but also to trials conducted elsewhere. As such, much of the guidance included is relevant to commercial deployments, particularly with regards to operational safety measures. The operational safety guidance builds upon that within BSI PAS 1883 such that the two are compatible, but with the Zenzic guidance adding informative guidance, examples and templates, in order to provide support for those creating or reviewing safety cases. This includes:

- Performing an operational safety risk assessment to identify hazards, prioritise risks and assign mitigations relating to operational hazards. In addition to guidance within the framework document itself, there are also downloadable templates available from the Zenzic website which cover three separate approaches to risk assessment:
  - A two-factor HARA (hazard analysis and risk assessment) – this is scored using ratings for likelihood and severity of incidents, a widespread approach to risk assessments within health and safety.
  - A three-factor HARA, which considers controllability by a safety operator as well as likelihood and exposure – although appropriate for use with safety operator, this has limited relevance to full commercial deployments, and will therefore not be considered further here.
  - A relative risk HARA, which supports an assessment of how the risk will change relative to a pre-existing benchmark. This is valuable where it is difficult to identify an absolute value for risk but where changes relative to an existing system can be readily assessed, with existing manually driven road traffic forming an appropriate benchmark (HumanDrive, 2019).
- Monitoring, reporting and continuous improvement
- Operational guidance – this section adds significant information over and above BSI PAS 1881, particularly with regard to method statements and how they can be used to ensure all personnel with a safety-related role have access to a clear and concise document that captures the key information that they need and serves as a 'single source of the truth'
- Emergency response and crisis communication plans – these are important to ensure appropriate steps are taken should something go wrong
- Route selection and assessment – broadly similar scope to BSI PAS 1883
- Incident reporting procedure – adds significant detail in the form of examples and templates, with flow diagrams to suggest an appropriate methodology
- Safe operation and control – provides some consideration of how remote safety operators could be evidenced to be an acceptable solution. However, the guidance is kept relatively high level, recognising that flexibility is needed given the wide range of potential implementations possible.

The final report by the Law Commissions (2022) envisions that licensing of NUIC (no user-in-charge, i.e. vehicles able to operate without a human available to take over control) operators, together with their plans and procedures, would take place separately to vehicle approval; this will allow consideration of local issues. This report utilises the same subdivision, and therefore the operational safety approval would consider the operator and their safety case. Under our model, the operator would be responsible for providing any oversight required (e.g. a remote operations system where assistants can help when

a which identifies a situation that it is unable to proceed within) and would also be responsible for other duties such as insurance, maintenance (including installation of safety-critical updates), paying of tolls and ensuring loads are secure.

It is worth noting that the Law Commissions have identified a difference to the Uniform Law Commission in the USA, whose approach requires a single 'automated driving provider' to cover all aspects of complying with the legal and technical requirements of AVs, further reinforcing the link between the design and deployment, whereas the GB recommendation would make it optional for them to be the same entity but possible for them to be separate. Whilst there are certain advantages in mandating a single entity for the entire process, such as reducing risks due resulting from inadequate communication or division of responsibility between entities or inadequate coupling between the test programme and the TOD, sections 4.1 and 5.9, covering the specification of the design and deployment domains and the test programme respectively, are intended to mitigate such risks, and in the absence of any concrete evidence to the contrary, it seems reasonable to follow the approach proposed by the Law Commissions in Great Britain such that flexibility is maintained.

The Law Commissions propose that, should the ASDE ('automated self-driving entity', with a meaning broadly equivalent to 'manufacturer' within this report) and operator be the same entity, a single safety case covering both roles would be submitted. Whilst at odds with the approval mechanism described within section 3.1, in practice this wouldn't be problematic as safety cases typically contain multiple documents (Zenzic, 2021) rather than being a single monolithic entity, in order to make reviewing and updating more practicable. Therefore, a single safety case could be submitted that covers the system and operational safety aspects, and these aspects could be reviewed separately by relevant stakeholders or assessors as required.

The operator would be required to obtain a licence, in a manner not dissimilar to existing licensing of public service vehicle operators, albeit with additional requirements relating to the complexities associated with operating an ADS. The Law Commissions included a proposal for a new procedure to grant interim passenger permits, the idea being that this would allow initial services to be approved on a limited scale in order to collect further information before a full licence is granted, and potentially even before type approval of the system has been granted; this has potential to enhance the operational safety by limiting exposure to hazards until they are better understood. Such an approach also received widespread support within the stakeholder consultations that helped inform the drafting of this report.

The findings from the Law Commissions include consideration of working conditions within other industries, noting that an air traffic controller is required to have a half hour break during or after every two-hour period. However, it also notes that little information is available about the required ratio of staff to vehicles in a remote operations centre, bearing in mind that external circumstances such as flash floods may cause many vehicles to require assistance at the same time, and that staffing must be sufficient to cover peak, rather than average, demand. The report observes that there will also be a need to be able to communicate with passengers; there was widespread agreement within the consultation upon the need for passengers to be able to make contact with a human operator when needed, especially following an incident.

The key summary of what the Law Commissions would expect an operator's safety case to contain is set out in recommendation 54. This states that *"To obtain a NUIC operator licence, the applicant should submit a safety case, showing how safety will be assured. Among other things, the applicant's safety case should set out:*

- 1. how oversight will be provided to vehicles, including suitable connectivity, equipment, staff training and rest breaks;*
- 2. incident management, including communication with passengers, road users and the emergency services, together with measures to remove vehicles causing an obstruction;*
- 3. systems, expertise and equipment to maintain vehicles, install updates and ensure cybersecurity;*
- 4. data management;*
- 5. whether safety relies on any element of remote driving, and (if so) how this will be done safely; and*

6. *ways to learn from mistakes, including links with local authorities, highway authorities and the police.*

*Where an ASDE and the NUIC operator are the same entity, the entity may submit a joint safety case covering both roles, to be assessed by the authorisation authority. In other cases, the safety case should address the ASDE's written specifications for what must be done to ensure safe operation."*

Waymo (2020c) have conducted extensive testing of their vehicles upon public roads, including pioneering mobility services that are available to the public, both with and without a safety driver in the vehicle. To support such direct interaction with the public, they have developed in-car features and user interfaces to help 'riders' to understand what the vehicle is doing and to allow them to communicate – for example, to set a direction, ask the vehicle to pull over, or speak to a member of the Rider Support team. They have also considered emergency responses, including not just communication with riders, but also interactions with law enforcement and first responders.

Waymo have developed a smartphone app that enhances the ability of riders to communicate with them and the vehicle. They also provide audio and visual information within the vehicles to keep them informed and to remind them of safety features such as seatbelts. They need to click a 'start ride' button either on the app or in the vehicle for the ride to commence, and have access to a 'pull over' button within the vehicle – when pressed, the vehicle will identify the nearest location to pull over safely such that the rider can exit the vehicle prior to reaching their destination.

In line with the ORR report summarised earlier, Waymo recognise the importance of managing staff fatigue within safety critical operations, and therefore have implemented a Fatigue Risk Management Programme. This is primarily focussed upon safety drivers within vehicles, but may form a reasonable benchmark for equivalent processes for remote assistants.

They have also incorporated various features to aid accessibility for users with impairments, including:

- Accessibility of mobile app
- Audio cues and tools
- Braille labels
- Visual display
- Accessible rider support

This is in line with the recommendations of the law commissions regarding the importance of ensuring accessibility to those with impairments.

It should be noted that whilst operational considerations for trials have been considered by many documents within the public domain, and experience has been gained through application, no such knowledge base exists for full commercial deployment of AVs. For this reason, operational safety guidance relating to trials will be used as a benchmark within this report. However, this highlights consideration of the practicalities of commercial operation safety as an important area for further work; it is therefore hoped that as such systems move towards commercialisation, exploratory work to support the development and regulation of operational procedures will be given equivalent attention to the assurance of system safety.

## Stakeholder Feedback

Within the first round of stakeholder consultations, a representative of a local authority observed that an analysis of the safety of the route and the needs of the vehicle may result in the need for modifications to the infrastructure such that it meets a minimum requirement for safe operation of the AV. However, another stakeholder pointed out that any infrastructure changes to accommodate the vehicle should not be to the detriment of other road user types; for example, extra use of guardrails to prevent pedestrians crossing roads would act as a deterrent to active modes of travel. They also stated that requirements should not be placed upon other road users to take mitigating steps beyond what is required for existing road traffic, such as wearing a hi-vis vest.

Within the same consultation, another interviewee referred to what they include within an operational safety case for existing trials, which "...covers everything from how we choose our safety drivers, what characteristics they have, to what would happen if the vehicle had a flat tyre, or what would happen if somebody ran out in the road in front of the vehicle and it couldn't stop". They stated that it should cover

the whole spectrum of hazards, and be subject to an independent safety case review. It must cover the physical characteristics of the trial route, but also other issues such as weather, lighting, fog etc. Multiple stakeholders highlighted the need for the risk assessment to consider a wide range of road user types, including emerging categories such as e-scooters.

The second round of stakeholder feedback presented initial proposals to the stakeholders and solicited feedback within the form of a survey. This indicated widespread support for the proposed requirements.

### 6.1.1.3 Conclusions Drawn

Due to the wide range of different use cases possible for LSAVs, and the limited guidance currently in the public domain as to how the operational safety of commercial deployments (as opposed to trials) should be assured, the requirements need to remain flexible, and be limited to a relatively high-level. As such, in the absence of detailed requirements, it will be necessary to provide guidance, and to signpost guidance that is already available, to inform decisions.

## 6.1.2 Recommendations

### 6.1.2.1 Proposed Requirements

The safety case report shall evidence the achievement of acceptable operational safety, i.e. safety with regards to hazards in the surrounding environment such as those presented by other road users, passengers or roadside infrastructure. At a minimum, this operational safety case shall include:

- Assessment of the safety and suitability of the route(s) defined within the TOD, including identification of any features or locations that pose a particular hazard (e.g. steep drop, limited line of sight).
- Identification of what hazards may be posed by other road users that the TOD defines as in scope.
- Identification of environmental hazards (e.g. fog, snow) that the TOD defines as in scope.
- Identification of hazards relating to customer interactions with the vehicle.
- An operational risk assessment to prioritise the identified hazards for mitigation, log any resulting mitigations, and (where applicable) assess the acceptability of the post mitigation risk. This shall include identification and justification of the methodology to determine the acceptability of risks, such as confirming that the risks are ALARP (as low as reasonably practicable) or GALE (globally at least equivalent).
- Operational guidance (e.g. a 'method statement') to define safe systems of work, hazard mitigations, roles and responsibilities etc. such that all staff carrying out safety critical work relating to the deployment, including those employed by a third-party organisation, have a clear understanding of how to perform their role safely.
- Defined maintenance procedures setting out how to work upon the vehicles in a manner that is safe and that results in safe functioning of the vehicle. This shall include consideration of selection and training of maintenance staff and consideration of appropriate workshop processes to ensure quality, consistency and safety is maintained.
- Safety management systems such as a process for capturing information on incidents and near misses in service, a change management system, and processes for staff selection, training, assessment and performance management. These are further examined in Section 7.1 (Safety Management Systems), and will therefore not be explored in detail here.
- A fatigue risk management system (FRMS) detailing what steps have been taken to mitigate the risk of safety-related incidents due to fatigue, how the effectiveness of the FRMS will be monitored, and how updates will be made to the FRMS in response to new data.

- Evidence of how it has been ensured that members of the public, including passengers and other road users, understand how to interact with the vehicle. This includes consideration of the clarity of how to accomplish tasks that may have safety implications (e.g. ensuring there is a clear means to request the vehicle to pull over at the next safe opportunity), consideration of accessibility such that persons with impairments can interact appropriately with the vehicle, and consideration of what communications are required in the event of an incident or emergency (e.g. liaison with emergency services, providing evacuation instructions to passengers).
- A description of how appropriate stakeholders were identified to ensure that each stage of the operational safety case had access to appropriate knowledge, experience and sign-off authority.
- An incident response plan that sets out how the scene will be made safe, if such action is necessary, after any foreseeable incident (including all incidents resulting from hazards that were identified as plausible within the risk assessments for the vehicle and the deployment), who should be contacted (including senior staff and emergency services) and what should be done to record and preserve evidence (such as not making changes to the scene other than those necessary to ensure the immediate safety). If it is necessary for staff to take action to ensure that digital data is preserved (e.g. to prevent a data recorded recording over the data before it can be accessed), the procedure for this shall also be documented.

Additionally, if the vehicle utilises remote, on-site or in-vehicle assistants to support the operation of the vehicle where it is unable to proceed safely without human intervention, the operational safety case shall include:

- Consideration of how staff are able to gain situational awareness and to provide inputs in a safe manner (see also Section 5.7 on human factors).
- Consideration of the robustness and security of the mechanism for human assistance, such as the wireless communications link to a remote operations centre (see also Section 5.3 on Cybersecurity and Section 5.6 on External Inputs).
- Consideration of the level of staffing required to ensure that the number of assistants available at any given time is at least equal to the reasonably foreseeable peak demand, bearing in mind the possibility of events that result in multiple vehicles requiring assistance simultaneously.

In all the above steps, consideration should be given to any system limitations identified within the VSCR such that appropriate mitigations can be put in place. For example, if the system is unable to pull into an oncoming lane to overtake parked vehicles, an appropriate mitigation may be for parking to be prohibited in the relevant area. However, mitigations shall not require changes to road infrastructure or road user behaviour that would have a significant detriment upon forms of travel other than LSAVs; for example, excessive use of barriers to separate pavements from roads may sever rights of way and have a detrimental impact upon those using active forms of travel such as walking or cycling.

### 6.1.2.2 Supporting Information

It should be noted that there is potential for lines to be blurred between what is classed as operational safety, and what is classed as part of the safety management system. As set out previously, this report considers the safety management system of the operator to be an essential component of the DSCR, much as the safety management system of the manufacturer will be within the VSCR (see Section 7.1). This section highlights some topics that may in particular be suited to being covered within the SMS produced by the operator, with other topics being more suited to inclusion within the non-SMS evidence such that tangible mitigations are reviewed during the approval process. It is advised that safety cases employ a similar distinction in order to support clarity, but this is not essential provided that the overall requirements and principles of both sections are adhered to.

#### Operational Risk Assessment

The operational risk assessment will be a key component of the operational safety case, and therefore regulators should look to see a thorough and robust process for identifying hazards, prioritising risks and implementing mitigations. This should cover the full scope of the operational lifecycle of the system, and should be proportionate to the complexity and risk associated with the deployment, bearing in mind

both the complexity of the environment (e.g. a shared space area open to many road user types will be inherently more complex than a designated lane for LSAVs) and the complexity of the operations to be undertaken within that environment. It must consider not just risks associated with the driving task that can be mitigated with operational measures external to the system, but also risks associated with other aspects of operation such as ingress/ egress of passengers, falls while the vehicle is in motion, failure to adequately secure cargo, emergency scenarios such as fires or civil unrest, personal safety against crimes such as assault or theft, and safety management within maintenance or storage facilities.

It is recommended that the operational risk assessment should conform to good practice as defined within guidance documents aimed at AV trials, such as BSI PAS 1881 (2020) and the Zenzic safety case framework (Zenzic, 2021); the latter includes examples and templates that could be applied to commercial deployments. Furthermore, it is recommended that the risk assessment should be aligned with best practice for health and safety, such as the BS ISO 31000 (2018) risk management guidelines and the Health and Safety Executive's guidance on reducing risks and protecting people (HSE, 2021).

In addition to using brainstorming methods to generate a list of hazards, the safety case should also utilise statistical sources of hazards; this will ensure lessons are learned from past incidents and reduce the likelihood of significant hazards being missed. Examples of suitable approaches include:

- Using a review of documents such as the Highway Code (2022) or the National Standard for Driving Cars and Light Vans (DVSA, 2010) as a prompt for what scenarios and hazards may be encountered; this may typically be found to result in consideration of hazardous scenarios that are not otherwise obvious (HumanDrive, 2019).
- Obtaining incident data from highway authorities such as the National Highways log of 'Top Level Hazards' or, accident studies such as STATS19 (2021) and RAIDS (2013).
- Hazards identified via an incident reporting process used during trials to develop the system.
- Hazards contained within any hazard log that may be developed in the future for collating and sharing AV safety data across the industry.

However, although such sources are valuable in creating a hazard list, it should be recognised that no two deployments are the same, and therefore the assessment of risks should consider the particulars of the deployment and not merely rely upon generic information. Where statistical data for manually driven traffic is used as a source, consideration should be given for how the nature and distribution of accidents may change for an LSAV deployment, bearing in mind the nature of automation and also the limited speeds and road types. The operational safety case should also provide evidence that appropriate stakeholders have been involved in the creation of the hazard analysis and risk assessment. The hazard log should, in particular, include significant focus upon the vehicle's interactions with vulnerable road users such as cyclists and pedestrians, although other road users such as cars, HGVs and emergency vehicles should also be considered.

The operational safety risk assessment could use a qualitative or quantitative approach, although in practice it may be difficult to acquire sufficient data to make the latter practicable, particularly for early deployments. A suitable qualitative approach would be to use a 'risk matrix' where likelihood and severity are given subjective scores (e.g. from one to five), which are then used to identify an overall score via a matrix. An example of a suitable risk matrix, from the Zenzic (2021) templates, is shown in Figure 42, although there is no universal scoring method and therefore other approaches can be adopted.

Note in particular that while the categories of unacceptable risk, tolerable risk and acceptable risk are in line with the principles set out by the Health and Safety Executive (HSE, 2021), the tolerable risk category (where it must be shown that risks have been reduced to be ALARP, or 'as low as reasonably practicable') has been split into an upper and a lower category in order to aid prioritisation of proportionate mitigation measures. 'Reasonably practicable' measures are those where the cost of implementation is not grossly disproportionate to the benefit in risk reduction.



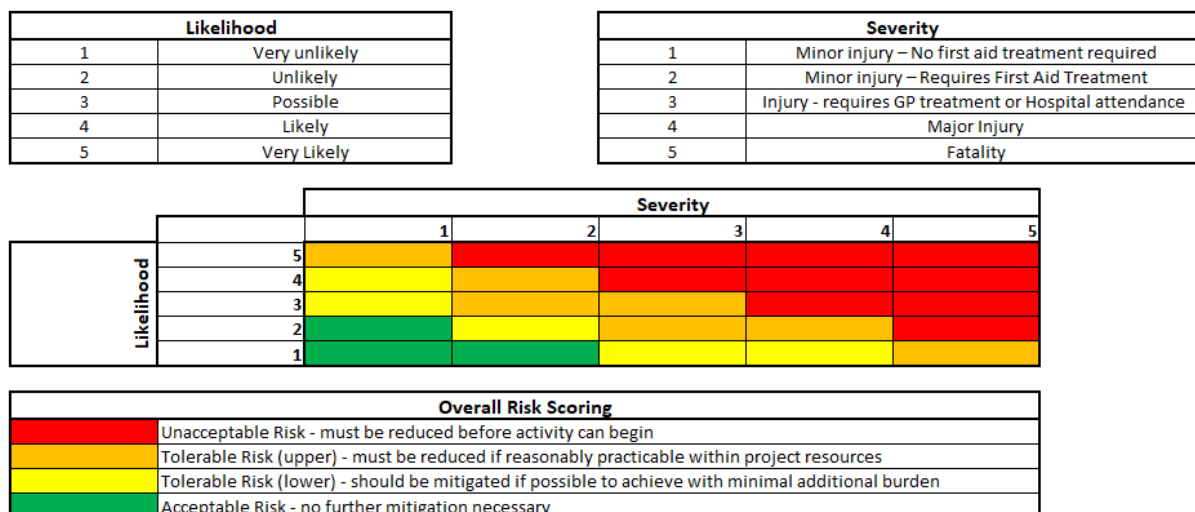


Figure 42: Example of a risk matrix to identify an overall risk based on ratings for likelihood and severity. Source: Zenzic (2021)

At a minimum, the operational safety risk assessment should contain the following headings:

- Description of Hazard
- Stakeholders Affected by Hazard (e.g. passengers, maintenance staff, pedestrians)
- Mitigations in place to control the hazard
- Likelihood score for the hazard being realised
- Severity score should the hazard be realised
- Overall risk level (identified via risk matrix)
- Justification for the scoring

Typically, a risk assessment would also include additional columns to capture further mitigations applied where the initial assessment identified a need for them, followed by a new assessment of the likelihood, severity and risk ratings. Examples of such mitigations include providing human oversight (e.g. having a member of staff present at a busy dispatch point), restricting the available route(s), limiting speed within certain areas, providing clear warning signs or making infrastructure changes (e.g. installing barriers).

It should be noted that other methodologies may be used as an alternative, provided that they can be appropriately justified; for example, GALE (globally at least equivalent, sometimes referred to using the French acronym of GAMAB) allows assessment of the risk relative to a pre-existing benchmark, with the risk being deemed acceptable if the overall level of risk from the new system is not greater than the benchmark one (Hillman, 2021). Care must be taken, however, where some individual risks rise significantly despite an overall reduction, as this could raise ethical concerns if some demographics are disproportionately exposed to increased risk ('risk intensity transfer'). The Zenzic templates include a 'Relative HARA' to support comparison of risks against a benchmark such as the National Highways log of top-level hazards or safety data relating to existing public service vehicles. Note that the UK legal system uses the principle of ALARP, or the similar SFAIRP (so far as is reasonably practicable); therefore, even if a relative risk approach such as GALE is used, it will still be necessary to provide evidence that risks have been made ALARP/ SFAIRP.

Although a risk assessment can be sufficiently complete for a service to commence, it should never be regarded as finished; the operational environment will change over time, and unforeseen hazards will present themselves. As such, it should remain a live document for as long as the vehicles remain in service, and therefore the safety case should include a plan for how monitoring will be used to inform updates. The ongoing monitoring and mitigation of risks should therefore be supported by a safety management system that meets the requirements set out in Section 7.1 of this report.

## Operational Guidance

The operational risk assessment will set out mitigations for risks that are external to the vehicle, but there will also be other safety-critical information that is essential to the procedures employed by the operator, including guidance provided by the manufacturer. For a complex system such as a deployment of multiple vehicles operating within a transport network, there will be many different staff in many different roles who each have a part to play in ensuring acceptably safe operation, and it is not reasonable to expect every one of them to have read, understood and retained all the information contained within the safety case. It is therefore necessary to provide operation guidance such that there is a concise, targeted and accessible 'single source of the truth' regarding staff responsibilities for safe operation. Types of operational guidance include:

- Method statement
- Safe operation of the ADS on the given route(s)
- Remote assistant policies
- Vehicle storage and security
- Vehicle maintenance, inspection and cleaning procedures
- Vehicle fuelling and charging
- Vehicle recovery plan
- Incident reporting policy or procedure
- Emergency response plan

A method statement is a form of document used extensively within health and safety management. It sets out the scope and workflow of the intended activities, the roles and responsibilities of the team, and an overview of the key safety processes, including mitigations logged in the operational risk assessment, that must be adhered to. Method statements are valuable as they collate the key information into one document, in a format that is accessible for all staff to engage with.

For a complex deployment featuring many staff in diverse roles, it may be deemed more appropriate to have separate procedure documents defining each role such that staff only need to be familiar with the aspects that affect them. Should this be the case, it should be confirmed that the separate procedures are compatible, and that there is a process in place to maintain their alignment within future updates.

Operational guidance such as a method statement could legitimately be interpreted as falling under the operational safety assurance or SMS scope; operators should consider either option when constructing their safety case, and select whichever provided the most clarity within the structure of their safety argument and evidence.

A key component of the SMS aspect of the operator's safety evidence should be a procedure for how to manage incident reporting, such that incidents are recorded and are also learnt from. This would typically include a means of categorising incidents according to their severity, of reviewing incidents in a manner that is proportionate to their categorisation, and of triggering suspension of services, updates to the technology or updates to operational procedures where appropriate. An example of such a procedure is shown in Figure 43, which categorises incidents into three levels: 'moderate' (e.g. mechanical breakdown, fault code triggered, MRM performed), 'substantial' (e.g. near miss that had the potential to cause harm) and 'severe' (e.g. collision or severe breach of road traffic laws). Any occurrences that violate the safety case, e.g. a remote assistant exhibiting signs of drowsiness while performing their role, should also be assigned an incident level (e.g. level 2 if there was the potential for a collision, but no collision occurred) and reported.

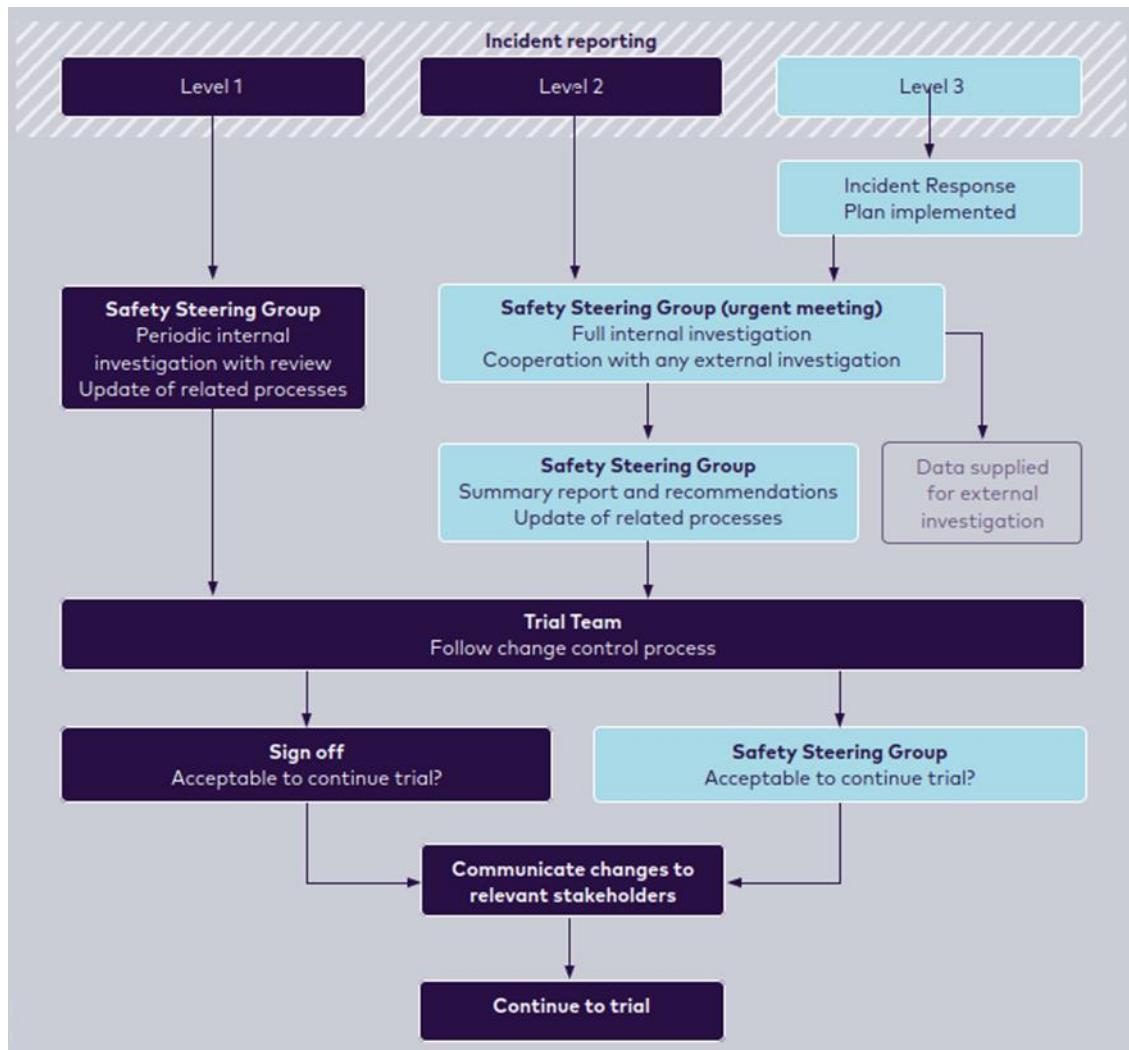


Figure 43: Example of an incident reporting process for safety trialling, which could be adapted for commercial deployments. Source: Zenzic (2021)

The operational safety case should also include plans for how to respond to emergencies in order to protect against further harm to individuals, financial loss or reputational damage. This should include defining the actions required and the roles responsible for performing them, and should be documented within an emergency response plan. It is likely that the operator will also possess a crisis communications plan to address media announcements following an incident, although this has no direct bearing upon safety and should therefore be seen as an internal process for the benefit of the operator rather than a matter for the regulator to consider.

### Route Assessment

As described in Section 4.2, it should be confirmed that the TOD definition accurately represents the true nature of the deployment route(s), to ensure that the safety case is not invalidated in service. However, there is also a need to examine the operational implications of the route such that its suitability can be assured. At a minimum, the following should be considered:

- Space available (length, width and height) relative to the vehicle dimensions and the ADS capabilities. This should include consideration of turning circles and manoeuvrability.
- The presence of any collision ‘hot spots’.
- Whether there are any challenging traffic flows, e.g. areas of high congestion.
- Damage or wear to the road or infrastructure.

- Locations where line of sight may be compromised, e.g. where a pedestrian may be able to enter the path of the vehicle without being visible beforehand.
- Features that could pose a collision hazard such as bridge abutments or buildings in close proximity to the vehicle path.
- Use of the route by other vehicles which may pose a hazard or result in conflicting traffic flows, e.g. bus routes.
- Road layouts that may make interaction between automated and non-automated vehicles challenging.
- Typical distributions of other road users for each location (e.g. some areas may typically feature a high density of pedestrians).
- Locations adjacent to the road where the use of the land may result in an elevated hazard (e.g. schools, hospitals, petrol stations).

The route assessment may consist of multiple stages, e.g. a desk-based analysis using an online map service initially, followed by on-site assessments such as walk-throughs, drive-throughs, measurement (by hand or by laser scanning) or video recording. Hazards identified within the route assessment should be used to inform the operational risk assessment; note that this may be an iterative exercise, as the resulting mitigations may involve changes to the route.

### Fatigue Risk Management System (FRMS)

Where humans have a safety-critical role to play within the deployment, management of fatigue becomes important; this is particularly so when humans are called upon to make decisions and inputs that directly affect the system operation, such as in the case of remote assistants. It is therefore important that a FRMS is put in place. This should include identification of hazards associated with fatigue, prioritisation of the resulting risks, and identification of mitigation measures. Furthermore, it should include a process to identify incidents related to fatigue in order to support continuous improvement. The FRMS may be a standalone document, or it may be incorporated into other documents (e.g. fatigue assessed within the operational risk assessment, and resulting procedures captured within a method statement).

In the absence of detailed guidance on this topic within the AV industry, consideration should be given to the guidance from the ORR (2013), which makes use of many years of practical experience and continuous improvement within the rail sector.

### Staff Training

The SMS element of the operational safety evidence should identify all roles that are safety critical, and for each, set out:

- Selection criteria for staff;
- Training courses and materials provided;
- Assessment methods;
- Refresher training to maintain knowledge and skills and be made aware of key changes;
- Performance management processes to ensure safety practices are followed;
- Reporting processes to ensure that incidents related to training insufficiencies result in updates to the staff training plan.

Particular focus should be placed upon remote assistants if they are used to support the deployment. This should include ensuring that they are adequately knowable and skilled with regard to road traffic laws, the TOD of the system, the information provided to them by the system (e.g. video feeds, system status dashboard), the controls available to them, the operational procedures for the deployment, and the emergency procedures should an incident occur.

### 6.1.3 Future Considerations

The above recommendations have had to make extensive use of materials aimed at trials (rather than deployments) or at other industries due to the lack of experience of LSAVs in full commercial deployments to date. This makes it difficult to foresee all use cases, and it may therefore be supposed that there are many 'unknown unknowns' waiting to be uncovered.

As such, rather than addressing limitations through further research projects, which would only serve to provide a greater abundance of trial-centric data, the emphasis should be on collecting data and updating requirements accordingly, as and when AVs start to be deployed in a commercial manner within GB.

Over time, it may be expected that the technology will become more sophisticated such that there will be less reliance upon humans to support operations, and indeed less reliance upon operational safety mitigations as a whole. This, together with new processes being developed as LSAVs are deployed for real, may make some aspects of the supporting guidance less relevant. However, as the requirements have been kept high level, it is hoped that they will remain relevant on a longer-term basis.

## 6.2 Post-Deployment

### 6.2.1 Background

#### 6.2.1.1 Definition of the Problem Addressed

Once commercial operation is underway, there remains a need to ensure that the deployment remains compatible with the approved safety case and with the regulatory requirements. It is therefore necessary to collect data throughout the operational lifecycle in order to identify discrepancies resulting from flaws within the safety case or from changes to the TOD for the deployment that may occur over time.

In-service monitoring is being examined in detail by Work Package 5 of this project, and will therefore be considered within a separate report. This section does not seek to duplicate this work, but instead to examine how a process would work for in-service data to trigger a change-management process such that safety assurance is maintained. This includes consideration of system updates, safety case updates and situations where it would be acceptable for no changes to result, and also looks at how the operator would be expected to interact with the authorities responsible for the roads used by the automated vehicles.

Section 7.1 considers safety management systems (SMSs) in terms of organisational processes required of the manufacturer to ensure that a safety culture is cultivated. This section instead focusses on the aspects of ongoing safety management that are particular to the operator, and particular to the monitoring of the vehicles and TOD during service; it is acknowledged that there is significant overlap between these objectives, which is reflected within the content of the sections.

### 6.2.2 Current State of the Art

UL4600 (2020) contains significant guidance relating to how in-service data should be used, with a particular emphasis placed upon how it can validate, or lead to updates of, the safety case.

For example, section 5.4.2 requires that lifecycle monitoring is “performed upon any evidence fully or partially based upon any of:

- Unsupported expert or subjective opinion
- Existing practices that are not supported by data and are not supported by written public standards documents, public guidance documents, or similar cited sources
- Assumptions”

This recognises that safety cases inevitably rely upon some arguments that include approximations, estimates or principles that are difficult to prove to a high level of confidence, and therefore it is important that in-service data is used to check the original arguments.

Similarly, section 5.5.2 of UL4600 requires that ‘accepted risks’ documented within the safety case (i.e. risks that have been determined to be sufficiently low such that they can be tolerated as residual risk) be ‘tracked through the item lifecycle via field engineering feedback’ to ensure that the accepted risks in practice are less than or equal to the level of risk estimated within the safety case. This is considered on two levels: it is required that, where residual risks have been accepted, the overall risk that materialises should be compared to that expected within the safety case, but it is ‘highly recommended’ that this analysis should be done on a per-risk basis. The clause also requires that the comparison should be on the basis of confirming that the risk in practice is “less than or equal to” the level of risk expected.

It is questionable whether requiring equal or lower risk (i.e. a one-sided tolerance) is the right approach, as it would mean even the slightest increase in the risk that materialises would trigger an update to the safety case; if the risks are predicted honestly (rather than being artificially inflated to be more conservative), this would be expected to result in around 50% of risks turning out to be higher. As such, this requirement could be argued to either result in an excessive conservatism in over-reacting to minor risk increases, or excessive conservatism in the original estimates used. The former would result in an

excessive administrative burden to update the safety case for insignificant changes, and the latter would be poor practice as it tends to result in inaccurate prioritisation of mitigation measures and gross inflation of costs where multiple conservative estimates compound (Dearden, 2016).

As such, it could be argued that a better approach would be to require the realised risks to not be significantly greater than the predicted risks, with a decision upon what constitutes ‘significantly greater’ requiring a level of engineering judgement, there being no objective way to determine a threshold. Such an approach would require an argument to be made within the safety case for what tolerance band should be deemed acceptable. Nonetheless, the key point from clause 5.5.2 of UL4600 is an important one; field data should be used to validate or correct any risk assumptions made within the safety case.

UL4600 also requires that the hazard log is updated in response to newly identified hazards, and recommends as a hazard identification technique “experience with the item under consideration or similar items” (section 6.3.1). This again requires in-service data to be collected and used to provide feedback.

Section 8.2.4 of UL4600 requires that the ODD “shall be detected and tracked to resolution”, requiring a strategy to be documented for how safety-related changes to the ODD will be detected via different monitoring sources to trigger upversioning of the ODD. Note that UL4600 doesn’t identify a distinction between design and deployment domains, and hence the use of the term ‘ODD’ within section 8.2.4 could be taken to be synonymous with the term ‘TOD’ in this report, as both refer to the reality that the vehicle experiences in service.

Section 9.3 requires that data on defects “shall be collected, analysed and used to improve products and processes.” Similarly, 10.5.1 (“the item shall be acceptably robust”) includes a requirement for the ability to detect and report a range of failures, errors and surprise events. This includes:

- Unexpected operational data (including distributional shifts and surprise events)
- Violations of assumptions made in the safety case
- Incorrect confidence values (such as classification confidence)
- Incorrect prediction values
- Adverse events for which risk was previously ‘unknown’
- Adverse events for which risk was previously ‘accepted’
- Negative consequences of changes (such as bug fixes, retraining)
- Robustness deficiencies
- Ambiguous or inconsistent data or commands
- Faults experienced (to help improve fault model)

10.6.1 of UL4600 requires the detection and reporting of loss events, although with the recognition that this should be done “to the degree that detection is practicable”, recognising that not all loss events may be plausible to detect (the need for acceptable detection is further elaborated upon in 10.6.2). 10.6.8 goes on to define how “the item shall report item status, operational parameters, faults, incidents, and loss event data with acceptable forensic accuracy”, for which it is mandatory that the approach to incident and loss event data recording and reporting is defined. 10.6.9 requires a post-incident analysis approach to be defined, and executed wherever applicable.

Section 11.3 refers to the need for robust data storage, with “data logs of faults, failures, incidents, mishaps” being included amongst the types of data required to be stored.

Section 12.5.1 requires the ability to detect safety related operational faults and design assumption violations, and requires logging of such data. 12.5.2 builds upon this to require an acceptable analysis of the results of the run-time monitoring.

12.6 describes requirements for safety case updates, with one of the triggers for this being “the occurrence of any safety-related incident regardless of whether the item has been changed in response or not”. This includes an impact analysis to identify whether safety case updates are required.

UL4600 also considers aspects of in-service safety such as maintenance procedures or risk assessment of hazards within the deployment location, although this analysis should be conducted and

documented prior to deployment approval; as such, it was considered in Section 6.1 of *this report*, and will not be addressed further here.

Section 16 of UL4600 examines what metrics should be used to assess performance; this will not be examined further here, as the selection of such metrics falls under the scope of Work Package 5 within this project. Section 17.5 requires prompt element lists (i.e. lists used as prompts during safety case creation) to be updated in response to in-service data.

Overall, therefore, UL4600's requirements relating to monitoring in-service could be summarised as encompassing:

- Validation of evidence used in the safety case
- Confirmation that hazards were identified sufficiently, and the resulting risks assessed suitably accurately
- Changes in the operating environment
- Behaviour that is unsafe in service, or that violates the safety case
- Faults identified within the system.

A process should be in place to update the safety case in response to such feedback sources where an impact assessment shows this to be appropriate.

BSI PAS 1881 (2020) contains a section on change control, which states that "the safety case shall remain a live document throughout the trial or testing period" and that "systems or operational changes that could impact safety shall be classified, managed and included in the safety case to ensure it remains up to date" It requires an audit trail for the changes and a classification of the change, based upon the safety impact.

Regarding change control, the safety case is required to include:

- the process for monitoring and capturing changes made;
- the process of assessing the level of risk posed by the change to safety and security;
- the process for documenting, classifying, and testing, as appropriate;
- the process for validating system performance before continuing trials or testing in the public domain;
- how changes to the safety case are communicated and implemented; and
- the method for monitoring the subsequent effects of any changes made

It should be noted that, at the time of writing, there is an updated version of BSI PAS 1881 being developed; this is due to include additional information on safety management systems. However, detailed discussion of safety management systems is beyond the scope of this section, and is examined further within Section 6.1.

The Safety Case Guidance produced by Zenzic (2021) builds upon BSI PAS 1881 to include requirements for what should be included within incident reporting and change management processes. Figure 44 summarises the proposed change management system, and it is advised that key roles might include:

- The change proposer – this could be anyone involved in the trials
- The change owner – a team member tasked with taking the change through the process
- Approver – responsible for checking the process is followed and signing off updates.



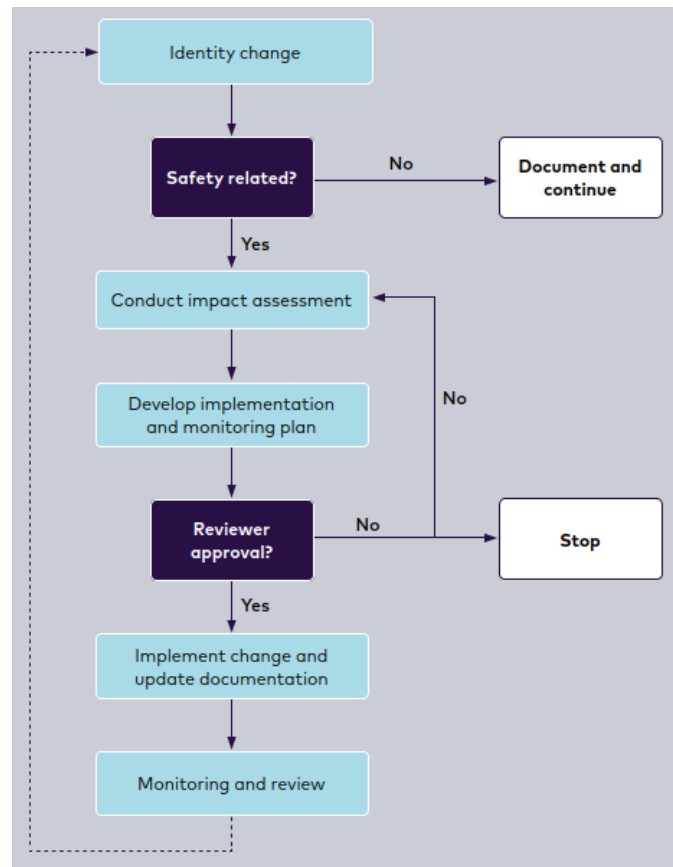


Figure 44: Proposed change management system within the Zenic safety case guidance. Source: Zenic (2021)

The change control should be integrated into a wider process that supports continuous improvement, including the monitoring of operations and the reporting of any identified incidents. This is summarised in Figure 45: Continuous improvement process. Source: Zenic (2021).

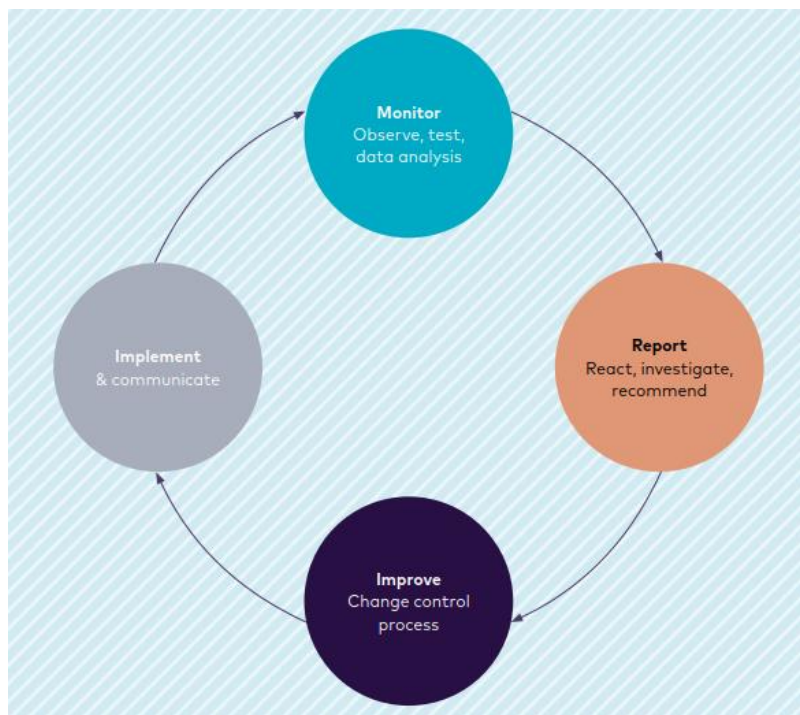


Figure 45: Continuous improvement process. Source: Zenic (2021)

The Zencic guidance also contains an example of an incident reporting process; this was examined in Section 6.1, so will not be re-examined here. For all the processes shown within the Zencic guidance, however, it should be noted that the document makes clear that all examples are optional to follow, and that trialling organisations should be free to use their pre-existing processes or to adapt the examples and templates in order to suit the needs of the application.

The Safety Management International Collaboration Group (SMICG) has published guidance on safety management systems (SMS) for small organizations (SMICG, 2015); this will not be reviewed in full here as the scope is more directly relevant to Section 7.1, but there are some sections that relate to how the in-service monitoring specified by Work Package 5 of the project should feed in to updates to the safety case. In particular, it is required that organisations:

- Decide how they will measure safety performance
- Keep monitoring to ensure mitigations are working as planned
- Take action if things are not improving

It requires the setting of safety performance indicators (SPIs), a term also used by UL4600, as a set of metrics used to monitor how well safety goals, targets and objectives are being met. This includes 'generic SPIs' that all organisations should meet, and 'specific SPIs' that the organisation or the regulator may decide need to be met. Examples of generic SPIs include:

- Number of major risk incidents (as defined in your Safety Management Manual);
- Number of mandatory reports;
- Number of voluntary reports;
- Number of overdue safety report closures;
- Number of safety meetings;
- Number of safety briefings; and
- Number of safety audits.

The guidance also notes that organisations should “be careful when reviewing SPIs, unless you have a reasonably large number of events. A change from one to two incidents per year is a 100% rate increase, but is not nearly as useful an indicator as a 10% change from 50 to 55”. This is an important consideration for automated vehicles; it should be expected that serious incidents will be rare, especially where numbers of vehicles may be low in the early stages of technology rollout, meaning it is difficult to get reliable statistics.

The guidance on management of change is well aligned with that in BSI PAS 1881 and in the Zencic guidance, including the need to initially assess the overall risk of the change itself, to identify all the factors that may be affected (e.g. new risks introduced, and how they can be mitigated) and to document agreed changes, all taking place as part of a continuous improvement cycle.

It is also important to note that regulations and test protocols are developed using extensive data from monitoring of vehicles in service, with this monitoring covering the entire vehicle fleet rather than individual vehicle types. Examples of this can be found within the technical papers available on the Euro NCAP website (Euro NCAP, 2021) or work done to use road collision data to assess opportunities to make vehicles safer and perform a cost-benefit analysis on regulatory options (TRL, 2020). As such, the in-service feedback should not just collect and utilise data for individual vehicle types, but should also look at the wider pool of vehicles that have been approved under the regulatory system to allow assumptions to be validated and thresholds to be calibrated such that regulatory requirements can improve over time.

## Stakeholder Feedback

An ADS developer who participated within the first round of stakeholder consultation expressed the need for greater flexibility than existing approvals when it comes to monitoring systems in service; rather than just being approved as an end goal, the system should be monitored in service over time, with mechanisms in place to address deviations. Change on a regular basis should be expected and accommodated. They also stated that the level of monitoring may depend upon the complexity of the operating environment and the capabilities of the system, with the need for monitoring, and particularly

human monitoring, potentially reducing over time as a record of safe performance is built up. A representative from the police expressed the need for any monitoring data that relates to collisions to be in an accessible format.

## 6.2.3 Recommendations

### 6.2.3.1 Proposed Requirements

The manufacturer and Operator shall reach and document an agreement upon what safety performance indicators (SPIs) will be monitored while the vehicle type is in service, how they will be monitored, and which organisation will be responsible for the monitoring. Data to monitor each SPI may be collected by the manufacturer, the operator or a third party; in the latter case, the 3<sup>rd</sup> party shall have a contractual arrangement with either the manufacturer or the operator, and evidence shall be provided that a quality assurance process is in place for the data collection activity.

The safety case shall document what SPIs will be monitored, shall provide an argument to justify why these are sufficient, and shall record which organisation is going to collect the data and by what means. The safety case shall also document and justify acceptance criteria for each metric such that there is a pre-defined threshold beyond which performance is deemed unacceptable and remedial action taken.

It is proposed that the SPIs shall be sufficient to provide a reasonable level of assurance that:

- Assumptions, models and subjective judgements within the safety case were not significantly inaccurate.
- The frequency of incidents involving harm ('lagging measures') is not higher than the acceptable level.
- The frequency of 'leading measures' (incidents that don't cause harm but nonetheless are a safety concern) is not above a level that implies, according to a model documented in the safety case, that the frequency of incidents involving harm is likely to be higher than acceptable. The documented model for how leading measures relate to a corresponding frequency of lagging measures shall itself be validated once sufficient lagging measure data becomes available to draw statistically justifiable conclusions.
- The log of hazards is sufficiently complete, and the estimate of the risk presented by each hazard is reasonable.
- The COD experienced by the system at any given instant in service is in accordance with the TOD that was defined for the deployment. This shall include consideration of events or object types encountered and also parameter ranges and frequency distributions. Any instances where the COD lies outside the TOD, resulting in remedial action such as an MRM or inhibition of the system, shall be logged. This requirement shall encompass both discrepancies due to limitations in the original definition of the TOD and discrepancies due to a change to the operating environment.
- Faults or erroneous data identified within the system are detected and logged.
- Incidents that required emergency intervention (such as lateral or longitudinal accelerations above a threshold value defined and justified in the safety case) shall be logged, with sufficient data available to support an investigation and potential remedial action.
- Collisions shall be logged, with sufficient data available to support an investigation and potential remedial action.
- Near misses shall be logged where practicable, with sufficient data available to support an investigation and potential remedial action.
- Safety concerns raised by staff, whether related to an incident or a general deficiency, shall be logged. This shall include the definition of an incident reporting process.

In order to make use of such SPIs, the safety case shall define and justify a process for how the significance of any unexpected or out of tolerance data will be assessed, and how this will lead to changes under the safety management system (such as an update to the safety case or a modification to the system). This process shall include a definition of roles and responsibilities, and of key decision points, and shall ensure that any changes to the system and any new data that has safety implications results in an update to the safety case. It shall also include a definition and justification of appropriate reactions to incidents or out-of-tolerance data, such as whether it justifies an immediate review or can be dealt with at a periodic review, whether it requires cessation of the service, whether temporary safety measures need to be put in place while investigation and/ or remedial action is underway, whether the regulator needs to be notified, and whether the safety case needs to be updated.

In order to support the regulator in assessing the overall safety statistics for automated vehicles, the regulator may require the manufacturer or operator to collect and share data relating to metrics that are of interest. Where this is the case, manufacturer and operator shall comply with the regulator's requests.

In order to ensure modifications to the road infrastructure are identified in advance and fed into the SMS, evidence shall be provided to demonstrate that the manufacturer and/ or the Operator have a process in place to liaise with the authorities responsible for any roads and road infrastructure within the deployment route(s). This may involve a contractual arrangement to provide updates; in the absence of any contractual arrangement, it must be justified that there is an alternative means to ensure that updates will be provided by the relevant authorities, including mitigation for the risk that organisational restructuring or staff changing roles may result in lines of communication being broken.

## 6.2.4 Future Considerations

The topic of safety management systems is relatively mature, having been developed in other industries, and therefore no further work is needed in terms of how in-service data feeds into this. However, the data that needs to be collected is not well understood, and is the subject of multiple industry working groups. It is therefore advised that the outputs of such working groups are monitored.

It may be expected that this is an area that will evolve significantly as experience is gained of full commercial deployment of AVs, and a greater understanding is gained as to how incidents occur. The requirements are at a sufficiently high level to provide flexibility such that they will still be relevant as the state of the art evolves, although there may be opportunities to make the requirements more prescriptive such that it is more difficult for gaps to go undetected. Furthermore, it is likely that the supporting guidance will need to be made more detailed to reflect evolving practice, and it may become possible to provide some concrete examples of approaches that have been used successfully.

# 7 Other Evidence to be Supplied to the Regulator

## 7.1 Manufacturer Safety Management Systems

### 7.1.1 Problem Summary

It is vital that an ADS manufacturer demonstrates to an approval body that they have robust, proactive safety management systems in place, and a strong safety culture across their organisation, that supports the safety case for deployment. This section will seek to specify the evidence that the manufacturer should supply across a breadth of topics, such as risk management, operational procedures and policies, incident reporting and analysis, continuous improvement and others.

This section aims to consider the recommended actions a LSAV manufacturer should take in order to maintain an effective Safety Management System (SMS) and strong safety culture. Although the deployment phase is the key focus of the overall approval scheme, the SMS and strong safety culture should also apply throughout LSAV development, testing and trialling of an LSAV. This section also details how manufacturers should evidence an effective SMS and safety culture, for example when submitting documentation to an approval body.

The recommendations within this section draw on a desk-based review of SMS and safety culture good practice and guidelines from the automated vehicle (AV) industry and other transport industries, including aviation and rail, and will inform the wider WP1 work regarding how to structure, manage, and implement an effective SMS for LSAV deployment.

The structure of this section centres on the recommendations (Section 7.1.3) for manufacturers, with the supporting evidence and good practice used to inform them provided in Appendix 6 - A, Appendix 6 - B, Appendix 6 - C and Appendix 6 - D (Section 10.6).

The SMS should address requirements for managing risk in general (including functional safety, cybersecurity, SOTIF, operational safety). Guidance from standards relevant for these disciplines contain requirements on management systems, with many being underpinned by the same processes. These are proposed to be merged into an overarching requirement here:

- Quality Management System
- Change Management
- Configuration Management
- Requirements Management
- Safety Culture

The SMS for the operator, together with other non-SMS safety evidence to support the deployment, is covered separately within Section 6; naturally, there is significant overlap between needs of an SMS for the manufacturer and the operator, and consequently there is overlap between the two sections, the differentiation being in the organisation and application that the guidance is targeted at.

### 7.1.2 Introduction to SMS, Safety Culture and Structure

An SMS is a framework or form of documentation that details an organisation's safety procedures and practices, based on defined safety principles, that is used to inform and improve an organisation's safety operation (AVSC, 2021). These procedures and documents are specific to each organisation and should be fit for the industry in which the organisation operates. It is essential that the SMS is embedded within an organisation and forms an integral part of the safety culture.

Safety culture is the overall attitude, approach, and knowledge of safety and the SMS within an organisation which influences how safety is prioritised, managed, and continually improved. It

determines how effectively an SMS is implemented across an organisation, how focussed employees, management, stakeholders, and key safety personnel are on safety, and how effectively incidents and improvements, or maintenance to the SMS itself, are recorded. Safety culture also comprises a shared set of values and beliefs and demonstrates a proactive approach to managing risks and eliminating incidents (where possible). It also ensures that any safety-related incidents are learned from and fed back into the SMS to ensure continuous improvement.

The relationship between safety culture and SMS is a synergic one. The Automated Vehicle Safety Consortium (AVSC, 2021) defines 4 key elements to an effective SMS:

- **“Safety Policy and Objectives (SPO):** *Establish or enhance safety practices with a clear safety policy, safety roles and responsibilities, and organizational safety objectives.*
- **Safety Risk Management (SRM):** *Proactively manage risk using safety risk assessments.*
- **Safety Assurance (SA):** *Monitor, analyse, and measure overall safety performance, including effectiveness of its safety risk controls, safety management, and associated processes.*
- **Safety Promotion (SP):** *Regularly conduct activities that inform, educate, and heighten the safety awareness of employees”.*

These key elements of an SMS are explained in greater detail in Appendix 6 - A. Each of these elements influences how effective an SMS is, both individually and collectively, and they all sit within the safety culture and the safety values/beliefs of the organisation.

Safety culture is the overall motivation and approach towards safety, which can be evidenced by the structure and effectiveness of an SMS, but the SMS will not be effective if the overall culture of the organisation is not focussed on the important aspects of safety. Leadership, management, consultation, and involvement are critical, so that all employees feel ownership of safety and it's made integral to everything they do. The Rail Safety and Standards Board (RSSB) define the relationship thus: *“the true ‘health’ of the safety of any organisation is primarily defined by the frequency of key day-to-day behaviours (frontline and management) and the extent to which these are encouraged and supported by an effective and flexible safety management system”.*

### 7.1.3 Requirements and Recommendations

This section identifies the requirements and recommendations to evidence and maintain an effective SMS and positive safety culture for LSAV deployments. These are based on a review of SMS and safety culture good practices and guidelines from the AV industry and other transport industries, including aviation and rail. By following these requirements and recommendations, organisations will be able to maintain an SMS that is effective in identifying and resolving safety issues within a safety culture that shares values of proper conduct and oversight.

Recommendations were identified from the research that can be found in Appendices 6 - A, 6 - B, 6 - C and 6 - D.

The requirements and recommendations are grouped by priority on a scale of one to three.

- **Priority 1** – This requirement must be followed for an effective SMS to be formulated or used. It is a key factor within the structure of an SMS or safety culture. It could also be required by law or regulation.
- **Priority 2** – This recommendation is very strongly encouraged. Failure to comply may not result in an ineffective SMS overall but will almost certainly limit application and effectiveness. It is not a required feature and may also be context dependent.
- **Priority 3** – This is a suggestion on improving or facilitating use of an SMS in a safety culture. It is not an essential need within effective SMS and safety culture but is likely to make an SMS more effective or improve clarity.

The following sub-sections summarise the recommendations from this report in short form, sorted by priority. Priority 1 recommendations are also separated into their specific areas of focus which are:

- Formulation and documentation

- Safety objectives and safety performance indicators
- Key safety personnel
- Specifics to AVs
- Safety risk assessments and safety reporting
- Employee consultation.

### 7.1.3.1 Priority 1 Recommendations

#### **Formulation and Documentation**

- **An SMS must use documentation to demonstrate commitment to SMS process and upkeep**
  - Include documentation of formulation, maintenance, changes after safety incidents and employee suggestions
- **Define a structure for the SMS and how it is broken down, including into safety objectives (SOs) and safety risk management**
  - The structure must be appropriate to the organisation and documentation must feed into appropriate sections
- **Follow the “Plan, Do, Check, Act” process from the BS ISO 45001 (2018) to formulate/improve the SMS**
  - The process documents and facilitates overall safety oversight, from identifying hazards or improvements, to implementing and auditing them
- **An SMS must be bespoke and unique to the organisation currently using it**
  - A generic SMS will not address the safety needs of specific organisations with unique structures and deployments
- **Map the SMS onto existing safety processes, where suitable processes already exist, and the safety management hierarchy during creation of an SMS**
  - This creates the most appropriate structure and content for the SMS. It will identify safety processes that are currently not considered or undertaken in the existing SMS or structure and should be incorporated. Mapping onto safety processes that are irrelevant for the organisation will result in unapplicable processes.
- **Use all possible relevant documentation to formulate an SMS, including vehicle certification, responses to safety incidents, etc. This must also account for regulations, standards and best practice guidance and requirements.**
  - Having a large evidence/knowledge base for the SMS is crucial to it being formulated properly
- **Conduct regular overall safety audits of the entire organisation and SMS process**
  - An overall picture of how an SMS works, alongside assessment of specific features, ensures an organisation is clear on how well an SMS is performing
- **Use gap analysis to examine the current safety culture before formulating new and more appropriate SMS processes to ensure issues are solved when it is being updated or created**
  - A gap analysis method is crucial to identifying issues to resolve in the SMS and should be used when auditing and updating the SMS
- **A safety policy (SP) must be established to capture the organisation’s values and commitment to safety and SMS implementation. This should define the key safety personnel and leadership’s commitment and responsibility in monitoring safety appropriately and helping other employees to do the same**

- This reflects the beliefs of the organisation and is key to establishing a strong safety culture. The culture will be based on the safety policy and therefore the values and actions of the culture must be captured in the safety policy

### **Safety objectives and safety performance indicators**

- **SOs must be specific to organisations and in-line with organisational focus**
  - Irrelevant SOs do not promote good safety values or culture and cannot be assessed by an organisation
- **SOs must be regularly reviewed and updated to reflect current practices relating to AVs**
  - This allows SOs to be useful for the organisation. Past SOs may have been met or are no longer applicable. Promotes ongoing safety oversight
- **SOs should be high-level aims that are defined by deeper research and knowledge of the organisation and in-line with organisational focus**
  - SOs should be goals to aim for, with safety performance indicators (SPIs) defining specific measurable elements. If they are not informed by research, they will likely not be effective. The SPIs should also follow a SMART (Specific, Measurable, Achievable, Relevant and Timed) structure of targets in order to be effectively applied.
- **SOs must be made in collaboration between senior leadership and key safety personnel**
  - These parties will be most informed about the safety culture and SMS of the organisation so both of their approval is required
- **SPIs must be defined as measures used to assess whether an organisation is achieving SOs**
  - A required part of the SMS. SPIs can be assessed by a variety of means, but must be specific to organisational work programmes and domains
- **SPIs must be specific measures against SOs and be updated in-line with these SOs so they remain relevant**
  - If SPIs are not relevant to SOs, then they are not usable. When SOs are updated to reflect organisational change or progress, SPIs should follow suit
- **SPIs must include a measurable element to evidence progress towards SOs**
  - If SPIs are not measurable, then they are not usable. They could be measured by number of reported incidents, timeline of changes or comments by employees
- **SPIs and SOs must be updated when new methods of data analysis are introduced, new work programmes are launched and when other trigger events occur**
  - Changes to safety processes and needs require new safety goals. Can be in response to trigger events and safety incidents
- **Relevance and performance of SPIs must be regularly assessed in reference to safety incidents and reporting, to ensure they are accurate measures**
  - SPIs are a measure of safety performance, but performance of the SPIs must also be reviewed so the organisation continues to assess its relevant safety needs.

### **Key Safety Personnel**

- **Nominate key safety personnel to take responsibility for application of and adherence to the SMS**
  - Having key personnel allows for focussed and dedicated oversight of the SMS rather than diluting the focus over a larger number of employees
- **Appoint employees with technical or AV safety expertise for key safety roles**



- Those editing and overseeing the SMS should be experts who are knowledgeable of safety needs and policy. Uninformed personnel will lead to a less effective and comprehensive SMS
- **Key safety personnel must collaborate with senior leadership to establish and promote safety culture**
  - Input is needed from both parties to fully account for organisational aims and for including the appropriate values in SMS integration with safety culture
- **A plan for promoting awareness and education of an organisation's SMS should be produced by key safety personnel**
  - The plan must be specific, and organised to give appropriate information to employees. Must not be simply handing out of generic SMS documentation
- **Key safety personnel must assess whether employees have the appropriate knowledge, skills, and motivation to work in accordance with the organisation's SMS**
  - If employees are not aware of safety processes or cannot perform these correctly, then the SMS will not be able to be actioned effectively
- **The SMS must be treated as a dynamic and regularly updated system that needs to be adapted considering new safety information, particularly for commercial LSAV applications**
  - An SMS is not a one-time consideration. It will develop and change with new information. Not assessing, auditing or updating this will lead to the SMS becoming obsolete
- **The SMS must be updated due to safety incidents and organisational changes or triggers**
  - This includes change of management, restructuring and inception of new work programmes
- **Nominated safety personnel should familiarise themselves with, and act upon, any applicable legislation relating to a duty of candour – for example, the Law Commissions (2022) proposed within their report on automated vehicles that there should be a duty of candour, although it should be noted that this is currently only a proposal, and not legislation.**
  - Following the Duty of Candour could help foster a no-blame culture and ensure oversight of appropriate safety systems including software and in-use monitoring. This ensures the primary focus is on safety and not punitive action on employees post-incident. The Duty of Candour is in itself a recommendation to Government only, and may or may not be taken forward.

### Specifics to AVs

- **Organisations working with AVs must define which data analysis or monitoring methods are used to determine whether a safety incident has occurred, and this should be updated as methods change**
  - Referencing irrelevant or outdated data analysis techniques means the SMS will not be usable
- **SMS for LSAV organisations should have dedicated sections for in-use monitoring and safety reporting, because in-use is where most safety incidents may be expected to occur**
  - LSAV organisations must be aware of which situations present the most hazards and address these accordingly within the SMS
- **Operational design domains (ODDs) and Target Operating Domains (TODs) should be assessed and updated alongside the SMS for safety reporting and processes upon ODD or TOD exit, to contain the appropriate actions to mitigate further incidents**

- ODDs and TODs should be well-documented and understood by organisations, and there should be dedicated processes for responding to these incidents
- **Software must be monitored alongside equipment and vehicles when safety monitoring is taking place**
  - This integrates the recommendations by the Law Commissions (2022). Not addressing software or data issues or needs during safety incidents will lead to important safety issues being ignored.

#### **Safety Risk Assessments and Reporting**

- **Safety risk assessments (SRAs) must regularly take place, particularly when new work programmes are introduced, to ensure that safety processes are able to mitigate significant hazards. These can be used to update the SMS**
  - SRAs are a legal requirement. They are effective learning opportunities to use for assessing current safety processes
- **Safety reporting and methods to report safety incidents must be available to all employees and there should be training in place to educate the organisation in how and when to report incidents**
  - Safety incidents must be reported accurately and promptly. Not reporting safety incidents is a serious indictment on a bad safety culture. The ability to report safety incidents involves employees in the safety culture and ensures continued safety oversight within an organisation
- **The SMS must include a safety report form that is used to report incidents. It should document information such as time/location, context and mitigating actions that took place**
  - Safety reports need to be accurate, otherwise they cannot be used to inform future safety decisions
- **Safety report forms must be documented and compiled to allow key safety personnel to use the learning to update the SMS with new information**
  - These are effective learning opportunities and evidence for where improvement to the SMS is needed. Not retaining report forms shows little focus on safety values.

#### **Employee Consultation**

- **Employees must be consulted on their experiences and issues with SMS operation, to gain information on real-world application of SMS processes and any qualitative changes needing to be made**
  - Employees need knowledge of their responsibility within the SMS to be consulted. They will offer insight into SMS operations during deployment where key safety personnel may not be able to assess this directly
- **Employee consultation should involve all levels of the organisation to gain an overall view of employees' opinions on and usage of the SMS**
  - The SMS must refer to all levels, so therefore all levels should be consulted. This ensures that the SMS is applicable and usable for all employees
- **Issues reported from the employee consultation cannot be used for reprisals against the reporter. This is to foster the “no-blame” safety culture**
  - Safety culture must include protection for all employees. Safety issues should be addressed to protect the organisation as a whole, without pointing fingers or punishing employees for complaining
- **Employees must be knowledgeable of what constitutes a safety incident and when and how these occur**

- Employees may ignore safety issues if not appropriately trained. This can lead to more severe or significant safety incidents in future if gaps in employee knowledge or education are not addressed.

### 7.1.3.2 Priority 2 Recommendations

- **The SMS should be formulated for LSAVs and their resulting complexity. This complexity will differ depending on the technology used and the scope of the system operation (including the complexity inherent in the ODD/TOD and behavioural competencies), and should be considered when formulating the SMS**
  - The complexity of the ADS should be addressed through an appropriately designed SMS for said vehicle
- **Investigate technical and software errors after safety incidents to eliminate reasonable doubt of human error**
  - This must be done, but in conjunction with addressing and investigating human error as well. Investigating software errors helps to prevent a punitive safety culture and follows the Law Commission's proposed Duty of Candour (Law Commissions, 2022).
- **Foster a “no-blame culture” around safety incidents unless there is clear human error and criminal liability**
  - Organisations should focus on addressing the cause of safety incidents and not immediately assume an element of human error as a direct cause.
- **Utilise the company lexicon when formulating the SMS, to ensure relevance and applicability to the organisation who will use it**
  - This should be done when referring to safety processes in-organisation but is not useful when collaborating with other organisations, as it can lead to confusion
- **Statements made in the SMS and the processes defined should relate to specific roles and members within the organisation and not be generic, but also not be overly specific**
  - The documentation of the SMS should use specific role names, reference relevant processes and follow the company lexicon, but need not include names or personal details of employees.
- **Collaborative partners should agree on shared safety frameworks when collaborating, and share relevant safety reporting information and training or education**
  - A shared safety framework is not always required, depending on the specific collaboration, but needs for sharing safety information should be addressed. This will ensure safety of all employees but should limit how much safety processes cross over between organisations so that each organisational SMS remains relevant to the organisation it was designed for
- **The SMS of individual collaborative partners should remain differentiated from each other so internal safety monitoring is appropriate**
  - Organisations must have bespoke SMSs, but some crossover will be required during collaborative projects with close partners
- **An SMS should reflect the organisational focus; for example commercial passenger pods vs off-highway material transport**
  - The focus needn't be explicitly stated everywhere throughout the SMS, but rather the processes changed to reflect the safety needs within the organisational focus. Focus on content, not name
- **If using aggregated datasets with collaborating organisations, safety incidents and responses should be discussed between partners**

- Organisations can individually report these incidents, but collaboration allows for more accuracy and insights into the context of safety incidents
- **Employee consultation should involve regular sessions to discuss issues, and dedicated channels where employees can report on any issues they encounter at any time. This is to be decided by key safety personnel**
  - Employees must be consulted, but the method of this is up to an organisation to design. This is the suggested method, as regular update sessions promote discussion of safety values and employees having 24/7 opportunity to report incidents allows for constant safety oversight.

### 7.1.3.3 Priority 3 Recommendations

- **Define the SAE Level of the AV in the SMS, which could help identify that the SMS was designed specifically for deployment**
  - Defining the SAE Level directly may help give more clarity, but it is most important to include the needs from the SAE Level and complexity of the LSAV requiring safety oversight rather than simply the naming the Level itself
- **Review the SMS of other organisations**
  - The SMS must be specific to an organisation, but review of other organisations SMSs can give insight into gaps in documentation
- **Organisations should be aware of the innovative nature of commercial LSAV organisations, and therefore be prepared to regularly update SOs and SPLs to maintain an up-to-date SMS**
  - Organisations will likely be aware by nature of being part of the industry. Disseminating this awareness in the safety culture could help promote a focus on this innovation and need for regular updating of the SMS
- **Very small organisations (<5 members) should have ubiquitous SMS knowledge across all employees**
  - Organisations this small will likely require all employees to take part in SMS formation and documentation – however, this will not always be the case, e.g. a member of staff may have a role that is purely administrative or financial, with no involvement in safety, hence this recommendation being only priority 3
- **There should be both proactive and reactive action to mitigate the effects of safety incidents**
  - By following proper SMS documentation methods, this should already be addressed, but reminders of being proactive and reactive are likely to help employees follow processes correctly.

### 7.1.3.4 Evidence for review

The recommendations in Section 7.1.3 are the actions for organisations to take when formulating and maintaining the SMS, in order to make it as effective as possible. They must also be able to evidence this commitment to a strong SMS, and a positive safety culture and values, to external reviewers or auditors. These extra-organisational personnel should be able to review the documentation and integration of the SMS within the organisation. Below are suggestions as to what evidence can be used for the review of an effective SMS. There may be other examples specific to certain organisations, and some organisations may not document all of the below examples; it is an overall suggestion for what to review.

- Overall SMS policy document and write-up
- Where applicable, previous versions of the SMS document to demonstrate change management and updates over time

- Documentation examples that were used to inform the formulation of the SMS, including but not limited to:
  - Vehicle certifications
  - Previously documented safety processes
  - Records of driver training and certification
  - Organisation certifications, e.g., FORS
- Safety risk assessments created through the SMS
  - Also include safety risk assessment guidance documents and templates
- Safety incident reports created through the SMS
  - Also include safety incident report guidance documents and templates
- Examples of specific changes made after safety incidents or reports
  - Documentation of recommended changes, then examples of where the change occurred
- Documented safety objectives and safety performance indicators
  - Both current and past examples should be reviewed to evidence how these have been adapted
  - Evidence of audits
- Documentation and identification of key safety personnel and the hierarchy/ structuring of safety work across the organisation
- Health and safety policies operated under the onus of the SMS
  - Possibly employee opinion on the effectiveness of these policies
- Evidence of regular employee consultation sessions and the suggestions/minutes taken from these meetings
  - Could include interviews of employees about the effectiveness of the consultation methods
- Evidence of channels where employees can report safety issues or suggest changes/areas for improvement of the SMS
- Discussion/ interviews with employees focussing on the safety culture and their perception/ involvement.

### 7.1.4 Summary

If a manufacturer follows the requirements and recommendations above for all 3 priority areas, it will be able to produce a strong SMS that effectively monitors and maintains safety practices within the organisation. It will also promote effective safety reporting and responsibility in safety culture. It is most important to follow the priority 1 requirements, as these are needed for an SMS to function properly. Each organisation must apply these to their own safety processes and deployments uniquely to have proper safety oversight. The most important requirement is an SMSs bespoke nature; it must be designed specifically for the organisation. Were an organisation not to follow the priority level 1 requirements, then there would be gaps and errors within the safety processes they follow, key reporting processes would not be followed, and the safety culture would not be well promoted.

The priority 1 requirements identify how to maintain a strong safety culture through these specific areas:

- Formulation and documentation
- Safety objectives and safety performance indicators
- Key safety personnel

- Specifics to AVs
- Safety risk assessments and safety reporting
- Employee consultation.

The priority 2 recommendations give guidance that should be followed to ensure a much more effective SMS. The priority 3 recommendations give examples of how to improve clarity in an SMS and general considerations on how an organisation should think about an SMS. Finally, an overview of the evidence to be reviewed is given, which can be used to establish whether these requirements and recommendations are followed. The rationale behind and further guidance on these requirements, recommendations and evidence can be found in Appendices 6 - A, 6 - B, 6 - C and 6 - D.

## 7.2 Collated Systems-Level Administrative Requirements

Preceding sections have recommended a series of requirements, procedures and processes as part of the LSAV approval scheme. Many stages of the approval framework require formal documented outputs to be submitted to regulators; in other words, there are administrative requirements placed upon manufacturers and operators. Table 38 collates together all these administrative requirements and states by whom the documentation should be provided, and during what phase within the overall framework.

Information Required:		Further Details:	Provided by:	During Phase:
<b>Manufacturer's information</b>		Include at a minimum the name, registered address, and also the name and address of any parent company.	Manufacturer	Pre-Approval
<b>System Design Capability Definition (to define the design intent)</b>	This shall include at least one ODD (Operational Design Domain), and least one behavioural competency definition, and at least one MEL (Minimum Equipment List)	<p>This is required in order to define what functionality the system can provide, under what conditions, and is required to support the downstream safety analysis and testing. Further detail on what is required can be found within Section 4.1.</p> <p>It is permitted for the manufacturer to provide more than one of each of the constituent definitions such that degraded performance modes can be defined for example, the system may provide a narrower set of behavioural competencies where component faults result in transition to a different MEL or where adverse weather results in transition to a different ODD.</p>	Manufacturer	Pre-Approval
<b>Safety &amp; Security Manual; including:</b>	<ul style="list-style-type: none"> <li>- Operational constraints / instructions;</li> <li>- requirements on external infrastructure / interfaces;</li> <li>- operator instructions;</li> <li>- user instructions;</li> <li>- maintenance guidelines.</li> </ul>	This shall contain the evidence from the SMS activities (see Section 7.1).	Manufacturer	Pre-Approval

<b>Safety &amp; Security Case; containing,</b>	<p>ADS Design description, including:</p> <ul style="list-style-type: none"> <li>- Capabilities implemented in ADS;</li> <li>- Sensing Functionality;</li> <li>- Planning Functionality;</li> <li>- Strategies implemented if ODD/TOD exit detected;</li> <li>- MRMs implemented and conditions for their activation.</li> </ul>	<p>Minimum capabilities required are:</p> <ul style="list-style-type: none"> <li>- maintain lateral/longitudinal position in lane;</li> <li>- Follow another vehicle;</li> <li>- Collision avoidance.</li> </ul> <p>The acceptance will be determined through evaluation whether the defined functionality achieves the required behavioural competence in the context of applicable scenarios that are expected to occur in the ODD and ultimately the TOD, considering,</p> <ul style="list-style-type: none"> <li>(a) Have sufficient capabilities been declared, e.g., are lane changes required, are there crossings or intersections?</li> <li>(b) Is the design appropriate to perform the required functionality?</li> </ul> <p>(see Section 5.4 on Proposed Technical requirements).</p>	Manufacturer	Pre-Approval
	<p>Safety Concept for hazards caused by malfunctioning behaviour &amp; evidence of appropriate strategies implemented</p>	<p>This shall contain the evidence from the Functional Safety activities (see Section 5.1).</p>	Manufacturer	Pre-Approval
	<p>Security Concept for risks arising from external threats</p>	<p>This shall contain the evidence from the Cybersecurity activities (see Section 5.3).</p>	Manufacturer	Pre-Approval
	<p>Acceptance criteria determined by manufacturers for hazardous behaviours caused by the implementation of the intended ADS functionality &amp; associated validation targets</p>	<p>This shall contain the evidence from the SOTIF activities (see Section 5.2).</p>	Manufacturer	Pre-Approval
<b>Regulatory standards</b>	<p>Identifies the 'standards' with which vehicles must conform.</p>	Regulator	Pre-approval	
<b>Vehicle type SMS</b>	<p>An SMS for the vehicle, perhaps based upon an Organisation SMS which is tailored for the specific vehicle.</p>	Manufacturer	Pre-approval	



<b>Deployment SMS</b>		<p>An SMS for vehicle operation including how necessary inspection and maintenance activities for the vehicle are supported across the deployed fleet.</p> <p>This may potentially be submitted later within the process, if the operator and/ or deployment route(s)/ area(s) haven't been identified at the pre-approval stage.</p>	Operator	Pre-approval
<b>SMS Audit report</b>		A report of the audit of the <i>Vehicle Type SMS</i> and <i>Deployment SMS</i> . It is expected this will be conducted periodically during the operational life of the vehicle/service.	Regulator	Pre-approval
<b>VSCR (Vehicle safety case report)</b>		A report providing a snapshot of the vehicle safety case at a point in time.	Manufacturer	Vehicle Type Approval
<b>Vehicle Type Approval</b>		Certification provided by the regulator to confirm acceptance of the safety and security of the vehicle type	Regulator	Vehicle Type Approval
<b>System Deployment Capability Definition (to define the actual deployment)</b>	<p>This shall include at least one TOD (Target Operating Domain), and least one behavioural competency definition, and at least one MEL (Minimum Equipment List)</p>	<p>This is required in order to define what functionality the system will be required to provide within the deployment, and is required to support the scenario-based testing that is specific to the deployment location (whether defined routes or a geofenced area) and the analysis of operational safety. Further detail on what is required can be found within Section 4.1.</p> <p>It is permitted for the manufacturer to provide more than one of each of the constituent definitions such that degraded performance modes can be defined for example, the system may provide a narrower set of behavioural competencies where component faults result in transition to a different MEL or where adverse weather results in transition to a different ODD</p>	<p>Operator *</p> <p>* Note that there remains an open question as to whether the scenario-based testing upon the actual deployment location should take place within the system or deployment approval phases.</p> <p>If the former option is chosen, it would also be necessary for the manufacturer to provide a System Deployment Capability Definition during the Vehicle Type Approval phase.</p> <p>This way, a TOD – rather than just an ODD – is available to form the basis for assessments of test programme coverage (see Section 5.9.4).</p>	Licensing and Deployment

<p><b>DSCR (Deployment safety case report)</b></p>	<p>This should include identification of all operational safety hazards, analysis of their associated risk, and details of any resulting mitigations put in place. It should also include consideration of the suitability of the route and of any provisions for humans to act as remote assistants, and should evidence that the roles and responsibilities of all staff with a safety-critical role are defined in a manner that it is reasonable to expect those staff to understand. See Section 6 for further detail.</p> <p>Detail of how the trial will be monitored post-deployment, including the division of responsibility between manufacturer, operator and any third parties, should also be provided (see Section 6.2).</p>	<p>Operator</p>	<p>Licencing &amp; Deployment</p>
<p><b>Vehicle operating licence</b></p>	<p>Certification provided by the regulator to confirm acceptance of the safety and security of the deployment, together with acceptance of the deployment from non-technical perspectives such as urban planning, traffic flows, or competition with other service operators</p>	<p>Regulator</p>	<p>Licencing &amp; Deployment</p>
<p><b>Regulatory Notice</b></p>	<p>A notice limiting operation issued in response to a notifiable event.</p>	<p>Regulator</p>	<p>Monitoring</p>
<p><b>Incident Reports</b></p>		<p>Operator / Manufacturer</p>	<p>Monitoring</p>
<p><b>Operational data</b></p>		<p>Operator / Manufacturer</p>	<p>Monitoring</p>
<p><b>Change proposal</b></p>	<p>Definition of a set of coordinated changes in vehicle design and / or in operation of the vehicle.</p>	<p>Operator / Manufacturer</p>	<p>Response</p>
<p><b>Defined SMS structure</b></p>	<p>The organisation's SMS structure should be defined to understand how it is broken down into relevant parts (including into safety objectives and safety &amp; risk management).</p>	<p>Operator</p>	<p>Licensing and Deployment</p>

<b>Defined safety objectives</b>	The safety objectives (SOs) should be defined to provide benchmarks against in-service safety data. Details should also be provided regarding how SOs are regularly reviewed and updated.	Manufacturer and Operator	Licensing and Deployment
<b>Defined safety performance indicators</b>	The safety performance indicators should be defined to provide benchmarks against in-service safety data. They should be specific to the OD.	Manufacturer and Operator	Licensing and Deployment
<b>Defined means of MRM instigation</b>	Detail regarding how passengers and/ or staff can instigate an MRM if required.	Operator	Licensing and Deployment
<b>Defined means of ADS data collection</b>	Detail regarding how, and what, in-service vehicle and ADS data will be captured and how it will be used to trigger improvements to the SMS and associated documentation/processes.	Operator	Licensing and Deployment
<b>Defined means of ADS data access</b>	Detail regarding how in-service vehicle and ADS data interfaces for the emergency services can be accessed and utilised in emergency scenarios.	Operator	Licensing and Deployment

*Table 38: Collated list of minimum administrative requirements due from the parties at the systems-level.*

## 8 Conclusion

Whilst it must be acknowledged that the state of the art in automated vehicle safety and security assurance is continually evolving, and therefore whilst the proposals contained within this report should be regularly reviewed in light of new developments, nonetheless it is hoped that they provide a stronger basis for regulatory approval decisions than was hitherto available. It remains the case, however, that industry regulators must expect to have to apply considerable flexibility in order to accommodate use cases, technical solutions or assurance methodologies that are unforeseen at the time of writing.

This report should, therefore, be viewed as an initial step within an evolutionary process that will continue to refine the requirements and guidance and to expand the scope. In so doing, it is hoped that the requirements will become more detailed and prescriptive, with less ambiguity, such that there is a reduced need for regulators to apply judgement within less defined areas. This will foster an ever-increasing level of understanding from all stakeholders of what is required in order to gain approval, resulting in improved efficiency in the creation and assessment of safety evidence and in greater consistency of decisions. Such clarity will also be important for industry, helping de-risk investments through increased confidence in the time and cost, and the actual pass/ fail outcome, of the assurance programme for an LSAV type under development.

In particular, commercial deployments of automated vehicles within Great Britain could be expected to result in a significant increase in the data available to validate the current proposals and to make updates accordingly. Careful attention should be paid to public and political opinion, especially should collisions or near misses occur within early deployments, to better understand whether the level of safety assurance is regarded as acceptable; this report sets out various approaches that could be adopted, but ultimately it remains a political decision whether systems should be expected to be at least as safe as the 'average' driver, as a 'careful and competent' driver, or some other benchmark. For example, whereas the public accept the relatively high risk presented by road travel, it is generally expected that other modes of transport such as rail or air should achieve far better safety, and it may prove to be the case that such levels of safety are required for LSAVs if they are to be politically acceptable.

It is important to recognise the wide gap that exists between being able to perform a technology demonstration in a research setting (asking the question: "is the system capable of performing this function *correctly*"), and delivering a production ready, safe and secure system (asking the question: "is the system capable of performing this function *incorrectly*"). The difference in technological maturity and in the level of safety assurance required to get from the former to the latter should not be underestimated, and claims that technology solutions are ready should be treated with caution until robust evidence of true production-readiness is presented.

Opportunities to align on an international basis should be sought on an ongoing basis. For example, the term 'TOD' presented within this report appears to be broadly synonymous with other terms such as 'Deployment Domain' (DD) and 'Operating Envelope Specification' (OES), but the opportunity for alignment will become clearer as more precise definitions for the terms evolve, supported by concrete examples of their use. Ultimately, the goal should be to achieve harmonisation of not just the terminology, but also the regulatory approach, on an international basis; this will allow far greater pooling of data and expertise on the regulatory side whilst facilitating economy of scale on the industry side.

## 9 References

- ALKS (2021) *UN Regulation No. 157 Uniform provisions concerning the approval of vehicles with regard to Automated Lane Keeping Systems (ALKS)*, United Nations. Available at: <https://unece.org/transport/vehicle-regulations-wp29/standards/addenda-1958-agreement-regulations-141-160>
- Almeida, L., Menezes, P., & Dias, J. (2020) *Interface transparency issues in teleoperation*. Applied Sciences, 10(18), 6232
- AMLAS (2022) *Guidance on the Assurance of Machine Learning in Autonomous Systems (AMLAS), version 1.1*. Available at: <https://www.york.ac.uk/assuring-autonomy/guidance/amlas/> . Last accessed 31/03/2022
- Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J. and Mané, D. (2016) *Concrete problems in AI safety*. arXiv preprint arXiv:1606.06565
- Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L. and Ridella, S. (2012) *In The 'K' in K-fold Cross Validation*. In ESANN, 2012
- Ashmore, R., Calinescu, R. and Paterson, C. (2021) *Assuring the machine learning lifecycle: Desiderata, methods, and challenges*. ACM Computing Surveys (CSUR), 54(5), pp.1-39
- Aymerich-Franch, L., Petit, D., Ganesh, G., & Kheddar, A. (2017) *Object touch by a humanoid robot avatar induces haptic sensation in the real hand*. Journal of Computer-Mediated Communication, 22(4), 215-230
- AVSC (2020) *AVSC Best Practice for Describing an Operational Design Domain: Conceptual Framework and Lexicon*, Automated Vehicle Safety Consortium, SAE International. Available at: <https://www.sae.org/standards/content/avsc00002202004/>
- AVSC (2021) *Information Report for Adapting a Safety Management System (SMS) for Automated Driving System (ADS) SAE Level 4 and 5 Testing and Evaluation*, Automated Vehicle Safety Consortium, SAE Industry Technologies Consortia
- Barber, C. (2015) *Evaluating human-computer interaction*, J. R. Wilson, & S. Sharples (Eds.), Evaluation of Human Work (4th ed., pp. 359-381). Boca Raton: Taylor & Francis
- Blumenthal, M.S., Fraade-Blanar, L., Best, R. and Irwin, J.L. (2020) *Safe Enough: Approaches to Assessing Acceptable Safety for Automated Vehicles*, RAND Corporation. Available at: [https://www.rand.org/pubs/research\\_reports/RRA569-1.html](https://www.rand.org/pubs/research_reports/RRA569-1.html)
- British Standards (2018) *BS ISO 45001:2018, Occupational health and safety management systems*, BSI
- BS ISO 31000 (2018) *BS ISO 31000:2018 Risk management — Guidelines*. Available at: <https://www.iso.org/standard/65694.html>
- BSI (2021) Digital Commentary Driving: The new safety technique that can help put automated vehicles on our roads, BSI. Available at: <https://www.bsigroup.com/en-GB/about-bsi/media-centre/press-releases/2021/june/digital-commentary-driving-the-new-safety-technique-that-can-help-put-automated-vehicles-on-our-roads/#:~:text=Digital%20Commentary%20Driving%20is%20a,that%20constitute%20automated%20driving%20behaviour>
- BSI Flex 1890 (2022) *Connected and Automated Vehicles – Vocabulary, Version 4*. Available at: <https://www.bsigroup.com/en-GB/CAV/cav-vocabulary/>

- BSI PAS 1883 (2020) *PAS 1883:2020 Operational Design Domain (ODD) – taxonomy for an automated driving system – Specification*. Available at: <https://www.bsigroup.com/en-GB/CAV/pas-1883/>
- BSI PAS 1884 (2021) *PAS 1881:2020 Safety operators in automated vehicle trials and testing – Guide*, available at: <https://www.bsigroup.com/en-GB/CAV/pas-1881/>
- Burton, S., Gauerhof, L., Sethy, B.B., Habli, I. and Hawkins, R. (2019) *Confidence arguments for evidence of performance in machine learning for highly automated driving functions*, International Conference on Computer Safety, Reliability, and Security (pp. 365-377). Springer, Cham
- Burton, S., Habli, I., Lawton, T., McDermid, J., Morgan, P. and Porter, Z. (2020) *Mind the gaps: Assuring the safety of autonomous systems from an engineering, ethical, and legal perspective*, Artificial Intelligence, 279:103201
- Burton, S., Kurzidem, I., Schwaiger, A., Schleiss, P., Unterreiner, M., Graeber, T. and Becker, P. (2021) *Safety Assurance of Machine Learning for Chassis Control Functions*, International Conference on Computer Safety, Reliability, and Security, pp. 149-162. Springer, Cham, 2021
- California Path (2016) *Peer Review of Behavioral Competencies for AVs*, University of California PATH Program. Available at: <https://www.nspe.org/sites/default/files/resources/pdfs/Peer-Review-Report-IntgratedV2.pdf>
- CertiCAV (2021) *CertiCAV Assurance Paper - A framework approach for assuring the behaviour of highly automated vehicles*, Connected Places Catapult. Published June 2021
- Cheng, H., Garrick, D.J. and Fernando, R.L. (2017) *Efficient strategies for leave-one-out cross validation for genomic best linear unbiased prediction*, Journal of animal science and biotechnology, 8(1):38
- Clark, H., McLaughlin, A. C., & Feng, J. (2017) *Situational awareness and time to takeover: exploring an alternative method to measure engagement with high-level automation*, S. CA (Ed.), Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 61, pp. 1452-1456. Los Angeles: SAGE Publications CA. doi:10.1177/1541931213601848
- Cooke, N. J. (2006) *Human factors of remotely operated vehicles*, S. CA (Ed.), Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 50, pp. 166-169. Los Angeles: SAGE Publications
- Cummings, M. L., Mastracchio, C., Thornbury, K. M., & Mkrtchyan, A. (2013) *Boredom and distraction in multiple unmanned vehicle supervisory control*, Interacting With Computers, 25(1), 34-47. doi:10.1093/iwc/iws011
- Czarnecki, K. (2018a) *Operational World Model Ontology for Automated Driving Systems Part 1: Road Structure*, Waterloo Intelligent Systems Engineering (WISE) Lab
- Czarnecki, K. (2018b) *Operational World Model Ontology for Automated Driving Systems Part 2: Road Users, Animals, Other Obstacles, and Environmental Conditions*, Waterloo Intelligent Systems Engineering (WISE) Lab
- Dearden, H. T. (2016) *Functional Safety in Practice*, SIS Suite
- De Vos, J., Waygood, E. O., & Letarte, L. (2020). *Modeling the desire for using public transport*, Travel Behaviour and Society(9), 20-98. doi:10.1016/j.tbs.2019.12.005
- De Winter, J. C., Happee, R., Martens, M. H., & Stanton, N. A. (2014). *Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence*, Transportation Research Part F: Traffic Psychology and Behaviour, 27, 196-217

DfT (2018) *Bus and Coach Security: Recommended Best Practice* (3rd, Ed.), Retrieved November 29, 2021, available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/875547/bus-and-coach-security-recommended-best-practice.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/875547/bus-and-coach-security-recommended-best-practice.pdf)

DfT (2019) *Code of Practice: Automated Vehicle Trialling*, Retrieved November 11, 2021, from <https://www.gov.uk/government/publications/trialling-automated-vehicle-technologies-in-public/code-of-practice-automated-vehicle-trialling>

DfT (2021) *Future of Transport: Deliberative Research*. Available at: <https://www.gov.uk/government/publications/future-of-transport-deliberative-research>

Dichabeng, P., Merat, N., & Markkula, G. (2021) *Factors that influence the acceptance of future shared automated vehicles - A focus group study with United Kingdom drivers*, Transportation Research Part F: Traffic Psychology and Behaviour, 82, 121-140. doi:10.1016/j.trf.2021.08.009

Diels, C., & Bos, J. E. (2016). Self-driving car sickness. *Applied Ergonomics*, 53

DVSA (2010) *National Standard for Driving Cars and Light Vans (Category B)*, Driver and Vehicle Standards Agency. Available at: <https://www.gov.uk/guidance/national-standard-for-driving-cars-and-light-vans-category-b>

DVSA (2013) *National Standard for Driving Buses and Coaches (Category D)*, Retrieved November 25, 2021, available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/377669/national-standard-for-driving-buses-and-coaches.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/377669/national-standard-for-driving-buses-and-coaches.pdf)

Edwards, T., Homola, J., Mercer, J., & Claudatos, L. (2017) *Multifactor interactions and the air traffic controller: the interaction of situation awareness and workload in association with automation*, Cognition, Technology & Work, 19, 687-698. doi:doi.org/10.1007/s10111-017-0445-z

Endsley, M. R. (1995). *Toward a theory of situation awareness in dynamic systems*, Human Factors, 37(1), 32-64. doi:10.1518/001872095779049543

Engström, J., Markkula, G., Victor, T., & Merat, N. (2017) *Effects of cognitive load on driving performance*, Human Factors, 59(5), 734-764

EU (2022) *Draft Regulation for ADS, working group documents reviewed are 6<sup>th</sup> Meeting (Com Impl act AD annexes v4.1\_urban\_shuttles) and 11<sup>th</sup> Meeting (Com Impl act AD annexes v8.2\_TCMV 21-12-2021 clean)*, last accessed May 2022. Available at: [https://circabc.europa.eu/ui/group/4273d650-b8a9-4093-ac03-18854fba4b5/library/9c15a20d-0551-4c89-8d42-b0149456e427?p=2&n=10&sort=modified\\_DESC](https://circabc.europa.eu/ui/group/4273d650-b8a9-4093-ac03-18854fba4b5/library/9c15a20d-0551-4c89-8d42-b0149456e427?p=2&n=10&sort=modified_DESC)

EU 2018/858 (2018) *Regulation (EU) 2018/858 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles*, European Union

EU 2020/683 (2020) *Commission Implementing Regulation (EU) 2020/683 implementing Regulation (EU) 2018/858 with regards to the administrative requirements for the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles*, European Union

Euro NCAP (2021) *EuroNCAP Website*. Last Accessed 04/09/2021. Available at: <https://www.euroncap.com/en>

European Commission (2022) *Automated Cars – Technical Specifications*. Draft last accessed July 2022, available at: [https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12152-Automated-cars-technical-specifications\\_en](https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12152-Automated-cars-technical-specifications_en)

- Favarò, F. M. (2021) *Exploring the Relationship Between "Positive Risk Balance" and "Absence of Unreasonable Risk"*, arXiv preprint arXiv:2110.10566. Available at: <https://arxiv.org/ftp/arxiv/papers/2110/2110.10566.pdf>
- Fawcett, T (2006) *An introduction to ROC analysis*, Pattern recognition letters, 27(8):861–874
- Federal Aviation Administration (2018) *VS 800.367B Aviation Safety (AVS) Safety Management System Requirements*, FAA
- Fong, T., Thorpe, C., & Baur, C. (2001) *Advanced interfaces for vehicle teleoperation: collaborative control, sensor fusion displays, and remote driving tools*, Autonomous Robots, 11, 77-85
- Fraade-Blanar, L., Blumenthal, M.S., Anderson, J.M. and Kalra, N (2018) *Measuring Automated Vehicle Safety: Forging a Framework*, Santa Monica, CA: RAND Corporation. Available at: [https://www.rand.org/pubs/research\\_reports/RR2662.html](https://www.rand.org/pubs/research_reports/RR2662.html)
- Gauerhof, L., Hawkins, R., Picardi, C, Paterson, C., Hagiwara, Y. and Habli, I (2020) *Assuring the safety of machine learning for pedestrian detection at crossings*, SAFECOMP 2020 (39th International Conference on Computer Safety, Reliability and Security, York
- Gnatzig, S., Chucholowski, F., Tang, T., & Lienkamp, M. (2013) *A system design for teleoperated road vehicles*, ICINCO, 2, 231-238.
- Golightly, D. (2015) *Situational awareness*, J. R. Wilson, & S. Sharples (Eds.), Evaluation of Human Work (4th ed., pp. 549-563). Boca Raton: CRC Press.
- Gravells, A. (2017) *Principles and Practices of Teaching and Training: A Guide for Teachers and Trainers in the FE and Skills Sector*, London: SAGE Publications Ltd.
- Groeger, J. A. (2002) *Trafficking in cognition: applying cognitive psychology to driving*, Transportation Research Part F: Traffic Psychology and Behaviour, 5(4), 235-248. doi:10.1016/S1369-8478(03)00006-8
- GSN (2018) *Goal Structuring Notation Community Standard Version 2*. Available at: <https://www.goalstructuringnotation.info/>
- Gyllenhammar M, Johansson R Warg F, Chen D and Heyn H-M (2020) *Towards an Operational Design Domain That Supports the Safety Argumentation of an Automated Driving System*, 10th European Congress on Embedded Real Time Software and Systems (ERTS 2020), Jan 2020, TOULOUSE, France. fihal-02456077
- Highway Code (2022) *The Highway Code*, Department for Transport, last updated 2022. Available at: <https://www.gov.uk/guidance/the-highway-code>
- Hillman, R. (2021) *An approach to managing the operational safety of autonomous vehicle trials*, Journal of Safety and Reliability, Oxford: Taylor and Francis. Available at: <https://doi.org/10.1080/09617353.2021.1920300>
- Hohenberger, C., Spörrle, M., & Welpea, I. M. (2016) *How and why do men and women differ in their willingness to use automated cars? The influence of emotions across different age groups*, Transportation Research Part A: Policy and Practice, 94(December), 374-385. doi:10.1016/j.tra.2016.09.022
- HSA (2015) *Safety Signs and Signals: 3rd edition*, Retrieved October 22, 2021, available at: <https://www.hse.gov.uk/pUbns/priced/l64.pdf>



HSE (2020) *HSE Information Sheet – Organisational Change and Major Accident Hazards*, Health and Safety Executive. Available at: <https://www.hse.gov.uk/pubns/chis7.pdf>. Website last accessed December 2021.

HumanDrive (2019) *HumanDrive: Autonomous Vehicle Project Safety Management*. Available at: <https://humandrive.co.uk/downloads/>

HumanDrive (2020) *Test Methods for Interrogating Autonomous Vehicle Behaviour - Findings from the HumanDrive Project*. Available at: <https://humandrive.co.uk/downloads/>

IEEE (2022) *Assumptions in Safety-Related Models for Automated Driving Systems*, available at: <https://ieeexplore.ieee.org/document/9761121>

Impacars (2021) *Impacars project website*, available at: <https://www.impacars.com/>. Last accessed Dec 2021

ISO 26262 (2018) *Road Vehicles – Functional Safety*. Geneva, Switzerland: International Organization for Standardization

ISO 9241 (1998) *Ergonomics of Office Work with VDTs-guidance on usability*, Geneva, Switzerland: International Organization for Standardization

ISO 9241-210 (2008) *Ergonomics of human-system interaction*, Geneva, Switzerland: International Organization for Standardization

ISO/AWI 5083 (2022) *Safety for automated driving systems — Design, verification and validation*, International Organization for Standardization, available on ISO platform for WG members

ISO/DIS 22737 (2021) *Intelligent transport systems — Low-speed automated driving (LSAD) systems for predefined routes — Performance requirements, system requirements and performance test procedures*, Geneva, Switzerland: International Organization for Standardization. Available at: <https://www.iso.org/standard/73767.html>

ISO/DIS 24089 (2022) *Road Vehicles - Software Update Engineering*, under development, draft last accessed May 2022, Geneva, Switzerland: International Organization for Standardization. Available at: <https://www.iso.org/standard/77796.html>

ISO/FDIS 34502 (2022) *Road vehicles — Test scenarios for automated driving systems - Scenario based safety evaluation framework*, Geneva, Switzerland: International Organization for Standardization, draft last accessed May 2022. Available at: <https://www.iso.org/standard/78951.html>

ISO/PAS 21448 (2019) *Road Vehicles – Safety of the Intended Functionality*, Geneva, Switzerland: International Organization for Standardization. available at: <https://www.iso.org/standard/70939.html>

ISO/FDIS 21448 (2022) *Road Vehicles – Safety of the Intended Functionality*, Geneva, Switzerland: International Organization for Standardization. available to committee members

ISO/SAE 21434 (2021) *Road vehicles – Cybersecurity Engineering*, Geneva, Switzerland: International Organization for Standardization. Available at: <https://www.iso.org/standard/70918.html>

ISO/TS 4804 (2020) *Road vehicles — Safety and cybersecurity for automated driving systems — Design, verification and validation*, Geneva, Switzerland: International Organization for Standardization. Available at: <https://www.iso.org/standard/80363.html>

Jansen, A. M., Giebels, E., van Rompay, T. J., & Junger, M. (2018) *The Influence of the Presentation of Camera Surveillance on Cheating and Pro-Social Behavior*, *Frontiers in Psychology*(October). doi:10.2289/fpsyg.2018.01937

Kempapidis, T., Castle, C. L., Fairchild, R. G., Hussain, S. F., Cash, A. T., & Gomes, R. S. (2020) *A scientific evaluation of autonomous vehicle user experience on sighted and visually impaired passengers based on FACS (Facial Analysis Coding System) and a user experience questionnaire*, *Journal of Transport & Health*, 19, 100906. doi:10.1016/j.jth.2020.100906

KI-Absicherung (2022) *KI-Absicherung - Safe AI for Automated Driving*, Project website. Last accessed June 2022, available at: <https://www.ki-absicherung-projekt.de/en/>

- Koppel, S., Lee, Y., Mirman, J. H., Peiris, S., & Tremoulet, P. (2021) *Key factors associated with Australian parents' willingness to use an automated vehicle to transport their unaccompanied children*, Transportation Research Part F: Traffic Psychology and Behaviour, 78(April), 137-152. doi:10.1016/j.trf.2021.02.010
- Krajewski, J. (2014) *Situational Awareness—The Next Leap in Industrial Human Machine Interface Design*, White Paper. Houston: Invensys Systems
- Law Commissions (2022) *Automated Vehicles: Joint Report*, Law Commission and Scottish Law Commission. Available at: <https://www.lawcom.gov.uk/project/automated-vehicles/>
- Leveson, N. G., & Thomas, J. P. (2021, 11 30) *STPA*, Retrieved from MIT: [https://psas.scripts.mit.edu/home/get\\_file.php?name=STPA\\_handbook.pdf](https://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf)
- Linkov, V., & Vanžura, M. (2021) *Situation awareness measurement in remotely controlled cars*, Frontiers in Psychology (April). doi:10.3389/fpsyg.2021.592930
- Lu, Z., Coster, X., & De Winter, J. (2017) *How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving*, Applied Ergonomics, 60, 293-304.
- Luz, R., Corujeira, J., Grisoni, L., Giraud, F., Silva, J. L., & Ventura, R. (2019) *On the use of haptic tablets for UGV teleoperation in unstructured environments: system design and evaluation*, IEEE Access, 7, 95443-95454.
- McDermid J, Hawkins R, Parsons M (2021) *Vehicle Regulation Objectives*, University of York, Assuring Autonomy International Programme
- Mizukoshi, Y., Sato, R., Eto, T., M, K., Matsuzaka, A., & L, Y. (2020) *A low cognitive load and reduced motion sickness inducing Zoom Method based on typical gaze movement for master-slave teleoperation systems with HMD*, 2020 IEEE/SICE International Symposium on System Integration (SII) (pp. 28-33). IEEE. doi:10.1109/SII46433.2020.9026260
- Mohapatra, J., Chen, P.Y., Liu, S., and Daniel, L. (2020) *Towards verifying robustness of neural networks against semantic perturbations*, arXiv preprint arXiv:1912.09533
- Mohebbi, R., Gray, R., & Tan, H. Z. (2009) *Driver reaction time to tactile and auditory rear-end collision warnings while talking on a cell phone*, Human Factors, 51(1), 102-110.
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002) *Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study*, Cognitive Brain Research, 14(1), 115-128. doi:10.1016/S0926-6410(02)00066-6
- Mouratis, K., & Serrano, V. (2021) *Autonomous buses: intentions to use, passenger experiences, and suggestions for improvement*, Transport Research Part F(76), 321-332. doi:10.1016/j.trf.2020.12.007
- MuCCA (2019) *MuCCA: Cars That Talk and Take Decisive Action Will Save Lives on our Motorways*. Available at: <https://mucca-project.co.uk/cars-that-talk/>
- Mutzenich, C., Durant, S., Helman, S., & Dalton, P. (2021) *Situational awareness in remote operators of autonomous vehicles: developing a taxonomy of situational awareness in video-relays of driving scenes*, Frontiers in Psychology(November). doi:10.3389/fpsyg.2021.727500
- NHTSA (2018) *A Framework for Automated Driving System Testable Cases and Scenarios (DOT HS 812 623)*, National Highway Traffic Safety Administration. Available at: [https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/13882-automateddrivingsystems\\_092618\\_v1a\\_tag.pdf](https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/13882-automateddrivingsystems_092618_v1a_tag.pdf)
- NIST (2021) *Automated Driving System Safety Measurement Part I: Operating Envelope Specification*, National Institute of Standards and Technology. Available at: <https://www.nist.gov/publications/automated-driving-system-safety-measurement-part-1-operating-envelope-specification>
- Nostadt, N., Abbink, D. A., Christ, O., & Beckerle, P. (2020) *Embodiment, presence, and their intersections: teleoperation and beyond*, ACM Transactions on Human-Robot Interaction (THRI), 9(4), 1-19.

O'Regan, S., Faul, S., & Marnane, W. (2013) *Automatic detection of EEG artefacts arising from head movements using EEG and gyroscope signals*, *Medical Engineering and Physics*, 35(7), 867-874

ORR (2013) *Managing Rail Staff Fatigue*, Office of Rail and Road. Available at: <https://www.orr.gov.uk/media/10934>

Pattinson, J., Chen, H., & Basu, S. (2020) *Legal issues in automated vehicles: critically considering the potential role of consent and interacted digital interfaces*, *Humanities and Social Sciences Communications* Volume, 7(153). doi:10.1057/s41599-020-00644-2

PEGASUS (2019) *PEGASUS Method – An Overview*. Available at: <https://www.pegasusprojekt.de/files/tmp/Pegasus-Abschlussveranstaltung/PEGASUS-Gesamtmethe.pdf>

Petty, G. (2009) *Evidence-Based Teaching: A Practical Guide (2nd ed.)*, Cheltenham: Nelson Thornes Ltd

Pezzementi, Z., Tabor, T., Yim, S., Chang, J.K., Drozd, B., Guttendorf, D., Wagner, M. and Koopman, P. (2018) *Putting image manipulations in context: robustness testing for safe perception*, In 2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), page 1–8. IEEE

Prechelt, L. (2012) *Early Stopping — But When?*, Montavon, G., Orr, G.B., Müller, KR. (eds) *Neural Networks: Tricks of the Trade*. Lecture Notes in Computer Science, vol 7700. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-35289-8\\_5](https://doi.org/10.1007/978-3-642-35289-8_5)

Proctor, R. W., & Van Zandt, T. (2008) *Human Factors in Simple and Complex Systems (2nd ed.)*, Boca Raton: CRC Press

Rae, A. (2007) *Acceptable Residual Risk - Principles, Philosophies and Practicalities*, 2nd Institution of Engineering and Technology International Conference on System Safety. Available at: <https://ieeexplore.ieee.org/document/4399904?arnumber=4399904>

Rahimi, M., Guo, J.L., Kokaly, S. and Chechik, M (2019) *Toward requirements specification for machine learned components*, 2019 IEEE 27th International Requirements Engineering Conference Workshops (REW), page 241–244. IEEE

RAIDS (2013) *Road Accident In-Depth Studies (RAIDS)*. Available at: <https://www.gov.uk/government/publications/road-accident-investigation-road-accident-in-depth-studies/road-accident-in-depth-studies-raids>

RAND (2016) *Driving to Safety – How Many Miles Would it Take to Demonstrate Autonomous Vehicle Reliability*, RAND Corporation. Available at: [https://www.rand.org/content/dam/rand/pubs/research\\_reports/RR1400/RR1478/RAND\\_RR1478.pdf](https://www.rand.org/content/dam/rand/pubs/research_reports/RR1400/RR1478/RAND_RR1478.pdf)

RSSB (2022) *Route Knowledge*, last accessed 27/01/2022. Available at: <https://www.rssb.co.uk/en/safety-and-health/guidance-and-good-practice/route-knowledge>

SAE J3016 (2021) *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, J3016\_202104*, SAE International. Available at: [https://www.sae.org/standards/content/j3016\\_201806/](https://www.sae.org/standards/content/j3016_201806/)

SaFAD (2019) *Safety First for Automated Driving*. Authored by Aptiv, Audi, Baidu, BMW, Continental, Daimler, Fiat Chrysler, HERE, Infineon, Intel and Volkswagen. Available at: <https://www.daimler.com/innovation/case/autonomous/safety-first-for-automated-driving-2.html>

Salonen, A. O. (2018) *Passenger's subjective traffic safety, in-vehicle security and emergency management in the driverless shuttlebus in Finland*, *Transport Policy*, 61(January), 106-110. doi:10.1016/j.tranpol.2017.10.011

Sargent, R. (2010) *Verification and validation of simulation models*, *Proceedings of the 2010 winter simulation conference*, pages 166–183. IEEE, 2010

SASWG (2020) *Safety of Autonomous Systems Working Group - Safety assurance objectives for autonomous systems*, Technical Report SCSC-153A 2.0 ed, Safety Critical Systems Club, 2020. URL: <https://scsc.uk/SCSC-153A>

- Shalev-Schwartz, S., Shammah, S., & Shashua, A. (2018) *On a Formal Model of Safe and Scalable Self-driving Cars*, CoRR, abs/1708.06374. Retrieved from <http://arxiv.org/abs/1708.06374>
- Shen, X., Chong, Z. J., Pendleton, S., Fu, G. M., Qin, B., Frazzoli, E., & Ang, M. H. (2016) *Teleoperation of on-road vehicles via immersive telepresence using off-the-shelf components*, *Intelligent Autonomous Systems*, 13, 1419-1433
- SMICG (2015) *SMS for Small Organisations*, Safety Management International Collaboration Group. Available at: <https://skybrary.aero/articles/sms-small-organizations>
- Stahl, P., Donmez, B., & Jamieson, G. A. (2014) *Anticipation in driving: the role of experience in the efficacy of pre-event conflict cues*, *IEEE Transactions on Human-Machine Systems*, 44(3), 603-613. doi:10.1109/THMS.2014.2325558
- Stanton, N. A., Chambers, P. R., & Piggott, J. (2001) *Situational awareness and safety*, *Safety Science*, 39(3), 189-204. doi:10.1016/S0925-7535(01)00010-8
- Stapel, J., Mullakkal-Babu, F. A., & Happee, R. (2019) *Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving*, *Transportation Research Part F: Traffic Psychology and Behaviour*, 60, 590-605
- STATS19 (2021) *STATS19 Forms and Guidance*. Available at: <https://www.gov.uk/government/publications/stats19-forms-and-guidance>
- Steinberger, F., Schroeter, R., & Watling, C. N. (2017) *From Road Distraction to Safe Driving: evaluating the effects of boredom and gamification on driving behaviour, physiological arousal, and subjective experience*, *Computers in Human Behavior*, 75, 714-726.
- TfL (2009) *Pictogram Standards: Issue 4*, Retrieved October 24, 2021, from <https://content.tfl.gov.uk/tfl-pictogram-standard.pdf>
- TfL (2019) *Tram Graphics Standard (CR4000): Issue 4*. Retrieved October 23, 2021, from <https://content.tfl.gov.uk/london-trams-tram-graphic-standard-cr4000-issue-04.pdf>
- The Management of Health and Safety at Work Regulations (1999). Available at: <https://www.legislation.gov.uk/ukxi/1999/3242/contents/made>
- The Public Service Vehicles Accessibility Regulations (2000). Retrieved November 20, 2021, from <https://www.legislation.gov.uk/ukxi/2000/1970/contents/made>
- The Public Service Vehicles Regulation (2020). Retrieved November 20, 2021, from <https://www.legislation.gov.uk/ukdsi/2020/9780111196021>
- Tikanmäki, A., Bedrník, T., Raveendran, R., & Röning, J. (2016) *The remote operation and environment reconstruction of outdoor mobile robots using virtual reality*, 2017 IEEE International Conference on Mechatronics and Automation (ICMA) (pp. 1526-1531). IEEE.
- Transport for London (2020) *Transport for London: Safety, Health and Environmental Report*, Transport for London. London: Transport for London
- Tremoulet, P., Seacrist, T., McIntosh, C. W., DiPierro, A., & Tushak, S. (2019) *Transporting children in autonomous vehicles: an exploratory study*, *Human Factors*, 62(2), 278-287. doi:10.1177/0018720819853993
- TRL (2020) *EU General Safety Regulation - TRL's Experience Developing the EU General Safety Regulation and the Pedestrian Safety Regulation*, available at: <https://www.trl.co.uk/projects/eu-general-safety-regulation>

TRL (2021a) *Driver Availability Monitoring Systems*, Brussels: European Commission. Available at: <https://trl.co.uk/uploads/trl/documents/MIS070-Technical-study-for-General-Safety-Regulation.-driver-assistance-monitoring-systems-DAMS.pdf>

TRL (2021b) *Remote Operation of Connected and Automated Vehicles – Project Endeavour Summary Report*. Available at: <https://trl.co.uk/uploads/trl/documents/PPR1012-Remote-operation-of-CAVs---Project-Endeavour---Summary-Report.pdf>

UK Government (2012) *Anti-Social Behaviour on Public Transport: Safety Measures*, Department for Transport. Retrieved November 23, 2021, available at: <https://www.gov.uk/guidance/anti-social-behaviour-on-public-transport-safety-measures>

UL 4600 (2020) *Standard for Evaluation of Autonomous Products*, Underwriters Laboratories. Available at: <https://www.shopulstandards.com/ProductDetail.aspx?productid=UL4600>

UNECE (2021) *11<sup>th</sup> Session of WP29 GRVA Item 11: German Act Amending the Road Traffic Act and the Compulsory Insurance Act – Act on Autonomous Driving*, United Nations Economic Commission for Europe

UNECE (2022) *Draft UNECE regulation for ADS, reviewed documents taken from 16<sup>th</sup> Session (FRAV-16-10.pdf) and 21<sup>st</sup> Session (FRAV-21-05.pdf)*, United Nations Economic Commission for Europe. Available at: <https://wiki.unece.org/pages/viewpage.action?pageId=87622236>

UNECE Regulation 79 (2017) *Addendum 78: UN Regulation No. 79, Revision 3: Uniform Provisions Concerning the Approval of Vehicles with Regard to Steering Equipment*, United Nations Economic Commission for Europe. Available at: <https://unece.org/DAM/trans/main/wp29/wp29regs/2017/R079r3e.pdf>

United Nations (1969). *Vienna Convention on the Law of Treaties*, 1155, 331, United Nations. Available at: <https://www.refworld.org/docid/3ae6b3a10.htm>

VeriCAV Project (2021) *VeriCAV Website*. Last accessed 04/09/2021. Available at: <https://vericav-project.co.uk/>

VMAD (2020) *VMAD-13-03, 13th VMAD IWG session*, Agenda item 4, October 9, 2020, United Nations Economic Commission for Europe

VMAD (2022) *New Assessment/Test Method for Automated Driving (NATM) Guidelines for Validating Automated Driving System (ADS) – amendments to ECE/TRANS/WP.29/2022/58*, VMAD, United Nations Economic Commission for Europe. Available at: <https://unece.org/sites/default/files/2022-05/GRVA-13-35e.pdf>

Waymo (2020a) *Waymo Public Road Safety Performance Data*, Waymo LLC. Available at: <https://waymo.com/safety/>

Waymo (2020b) *Waymo Safety Methodologies and Readiness Determinations*, Waymo LLC. Available at: <https://waymo.com/safety/>

Waymo (2020c) *Waymo Safety Report*, Waymo LLC. Available at: <https://waymo.com/safety/>

White, H., Large, D. R., Salanitri, D., Burnett, G., Lawson, A., & Box, E. (2019) *Rebuilding Drivers' Situation Awareness During Take-Over Requests in Level 3 Automated Cars*, H. White, D. R. Large, D. Salanitri, G. Burnett, A. Lawson, & E. Box (Ed.), *Proceedings of the Contemporary Ergonomics and Human Factors*. Stratford-upon-Avon

Zenic (2021) *Safety Case Framework: The Guidance Edition*. Available at: <https://zenic.io/projects-and-resources/safety-case-framework/>

Zhang, M., Zhang, Y., Zhang, L, Liu, C. and Khurshid, S. (2018) *Deeproad: Gan-based metamorphic autonomous driving system testing*. arXiv preprint arXiv:1802.02295

Zhang, N., Zhang, L. and Cheng, Z. (2017) *Towards simulating foggy and hazy images and evaluating their authenticity*, International Conference on Neural Information Processing, pages 405–415. Springer

# 10 Appendices

These appendices contain extra details and extended discussion, in order to elaborate upon the topics and themes addressed within the main body of this document.

## 10.1 Appendix 1: Literature review methodology for 4.7 Human Factors

Search terms that were used in the databases to search for relevant literature:

Review of literature where the ADS interacts with	1st Level Search Terms	2nd Level Search Terms
1. Passengers	“Automated-vehicle”, “AV”, “autonomous”, “autonomous driving system”, “ADS”, “shared autonomous vehicle”, “SAV”, “remote”, “operator”, “supervisor”	AND passengers”, “shared”, “bus”, “participants”, “on-board”, “experience”, “attitudes”, “behaviour”, “impact”, “concerns”, “bus”, “train” “standards” guidelines”, “regulations”, “signs”, signage”, “pictograms”
2. Remote Operators	“human factors”, “human machine”, “interface”, “usability”, “user-experience”, “considerations”, “emergency”	“remote”, “operator”, “interaction”, “interface”, “user”, “usability”, “standards” guidelines”, “regulations”
3. Other parties in the event of an incident, such as other road users, emergency services personnel, and recovery personnel		“response”, “planning”, “hazard”, “coordination”, “team”, “standards” guidelines”, “regulations”.

Table 39: Literature review methodology.

## 10.2 Appendix 2: Measurement methods for 4.7 Human Factors

Integration of methods for measuring situational awareness includes (Linkov & Vanžura, 2021; Golightly, 2015):

Method	Measure	Example measure
<b>Performance methods</b>	Measures Operator results. Assumes that the results are related to SA so concerns with validity	Readback errors Eye tracking EEG
<b>Experimental methods</b>	Measures the Operator’s behaviour and SA knowledge during performance	Situation Awareness Global Assessment Technique (SAGAT)-simulator research with pauses  Situation Present Assessment Method (SPAM)-real-time research during activity  Verbal protocol recoding and analysis <sup>15</sup>
<b>Subjective methods</b>	Based on either an observer’s ratings or Operator’s self-assessment	<b>Self-rating</b>  Situational Awareness Rating Technique (SART)  Post-assessment of Situational Awareness Rating  <b>Expert-rating</b>  Situation Awareness Behavioural Rating Scale (SABARS)

Table 40 Measurement methods

---

<sup>15</sup> The Operator describes what they are doing and thinking while undertaking the task



## 10.3 Appendix 3: Summary of requirements from 21<sup>st</sup> FRAV session

<b>UNECE Working Party on Automated/Autonomous and Connected Vehicles                      Functional Requirements for Autonomous Vehicles (FRAV) ADS Safety Elements proposed                      in 21st Session (document titles “FRAV -21-05.pdf” - <i>This document is a working draft</i>)</b>		
<b>ADS Safety Requirements</b>	4.1.1	The ADS should be capable of performing the entire Dynamic Driving Task (DDT)
	4.1.2	The ADS shall recognize the conditions and boundaries of the ODD of its feature(s) pursuant to the manufacturer’s declaration under paragraph 3.2.
	4.1.3	The ADS shall detect and respond to objects and events relevant to the DDT. 36
	4.1.4	The ADS shall comply with safety-relevant traffic laws according to the ODD of the feature in use
	4.1.5	The ADS should interact safely with other road users
	<b>ADS interactions with ADS vehicle users</b>	4.2.1
<b>ADS management of safety-critical situations</b>	4.3.1	The ADS shall execute a fallback response in the event of a failure in the ADS and/or other vehicle system that prevents the ADS from performing the DDT
	4.3.2	The ADS shall signal its intention to place the vehicle in an MRC.
	4.3.3	Pursuant to a traffic accident, the ADS shall stop the vehicle.
<b>ADS management of system failures</b>	4.4.1	The ADS shall detect and respond to system malfunctions and abnormalities relevant to its performance of the DDT
	4.4.2	The ADS shall be protected from unauthorized access
	4.4.3	The ADS shall signal [faults/failures] compromising its capability to perform the entire DDT relevant to the ODD of its feature(s)
	4.4.4	The ADS shall signal [faults/failures] compromising its capability to perform the entire DDT relevant to the ODD of its feature(s)
	4.4.5	The ADS may continue to operate in the presence of [faults/failures] that do not prevent that ADS from fulfilling the applicable safety recommendations.

	4.4.6	The ADS shall signal [faults/failures] compromising its ability to execute the DDT.
--	-------	---

*Table 41 Summary of requirements from 21st FRAV session*

## 10.4 Appendix 4: Summary of requirements from European Commission draft (Dec. 2021)

<b>Draft European Commission Regulation</b> <b>Uniform Procedures and technical specifications</b> <b>for the type-approval of motor vehicles with regard to their automated driving system</b> <i>Draft Dec 2021</i>	
<b>2 Dynamic Driving Task (DDT) under nominal traffic scenarios.</b>	The ADS shall be capable of performing the entire Dynamic Driving Task (DDT).
	The capability of the ADS to perform the entire DDT shall be determined in the context of the ODD of the ADS
	As part of the DDT, the ADS shall be able to: <ul style="list-style-type: none"> <li>-Operate at safe speeds;</li> <li>-Maintain appropriate distances from other road users by controlling the longitudinal and lateral motion of the vehicle;</li> <li>-Adapt its behaviour to the surrounding traffic conditions (e.g., by avoiding disruption to the flow of traffic) in an appropriate safety oriented way.</li> <li>-Adapt its behaviour in line with safety risks (e.g., by giving all road users and vehicle occupants the highest priority)</li> <li>-Activate the relevant other vehicle systems when necessary (Opening doors, activate wipers in case of rain, etc)</li> </ul>
	The vehicle equipped with ADS shall be able to drive in the reverse direction (reverse gear)
	The ADS shall detect and respond appropriately to objects and events relevant for the DDT
	Objects and events might include, but are not limited, to: <ul style="list-style-type: none"> <li>-Vehicles, motorcycles, bicycles, pedestrians, obstacles (e.g., debris, lost cargo, animals)</li> <li>-Road accidents</li> <li>-Road safety agents / enforcement agents.</li> <li>-Emergency vehicles.</li> <li>- traffic signs, road markings and speed limits</li> <li>-environmental conditions (e.g., lower speed due to rain, snow).</li> </ul>
	The ADS shall comply with traffic rules of the country of operation
	The ADS shall interact safely with other road users, such as via: <ul style="list-style-type: none"> <li>-Signalling manoeuvre intentions.</li> <li>-Signalling ADS status active/inactive.</li> <li>-Using the horn where appropriate.</li> </ul>
	Vehicles with ADS intended to carry standing or unrestrained vehicle occupants shall not exceed a combined horizontal acceleration of 2.4 m/s <sup>2</sup> in normal operation. Depending on the factors influencing the risk to occupants and other road users, it might be appropriate to exceed these limits.
	<b>3 DDT under critical traffic scenarios (emergency manoeuvre).</b>
The ADS shall be able to detect the risk of collision with other road users or a suddenly appearing obstacle (debris, lost load) and shall be able to automatically perform appropriate emergency manoeuvres (braking, evasive steering) to minimize risks to safety of the vehicle occupants and other road users.	

	<p>If a crash can be avoided without causing another one, it shall be avoided.</p> <p>After the evasive manoeuvre the vehicle shall aim at resuming a stable motion.</p> <p>If the emergency manoeuvre results in the vehicle with ADS being at standstill, the signal to activate the hazard warning lights shall be generated automatically in accordance with traffic rules. If the vehicle with ADS automatically drives off again, the signal to deactivate the hazard warning lights shall be generated automatically.</p> <p>Pursuant to a traffic accident, the ADS shall stop the vehicle. ADS reactivation shall not be possible until the safe operational state of the ADS has been verified by the on-board operator or the remote intervention operator.</p>
<b>4 DDT at system boundaries</b>	<p>The ADS shall recognize the ODD conditions and boundaries of the ODD of its feature(s).</p> <p>The ADS shall be able to determine when the conditions are met for activation.</p> <p>The ADS shall detect and respond when one or more ODD conditions are not or no longer fulfilled.</p> <p>The ADS shall be able to anticipate planned exits of the ODD.</p> <p>The ODD conditions and boundaries (measurable limits) shall be established by the manufacturer.</p> <p>The ODD conditions to be recognized by the ADS shall include:          -Precipitation (rain, snow)          -Time of day (light intensity, including the case of the use of lighting devices)          -Visibility          -Road and lane markings</p> <p>When the ADS reaches the boundaries of the ODD of the ADS, it shall fall back to a Minimum Risk Condition (MRC).</p>
<b>5 DDT under failure scenarios</b>	<p>The ADS shall detect and respond ADS and vehicle malfunctioning behaviour</p> <p>The ADS shall perform self-diagnosis of faults and failures.</p> <p>The ADS shall detect malfunctioning behaviour and evaluate ADS's ability to fulfil the entire DDT.</p> <p>Provided a failure does not significantly compromise ADS performance, the ADS shall respond safely to the presence of a fault/failure in the ADS</p> <p>The ADS shall execute a safe fallback response directly to a Minimal Risk Condition (MRC) in the event of a failure of the ADS and/or other vehicle system that prevents the ADS from performing the DDT</p> <p>The ADS shall immediately upon detection, signal major failures and resulting operational status to vehicle occupants, the operator (if relevant) or the remote operator (if relevant), as well as to other road users (e.g., activation of the hazard warning lights)</p> <p>If failures are affecting the braking or steering performance of the vehicle, the manoeuvre shall be carried out with consideration for the remaining performance.</p>

<p><b>6 Minimum risk manoeuvre</b></p>	<p>During the minimum risk manoeuvre the vehicle with the ADS shall be slowed down, with an aim of achieving a deceleration demand not greater than 4.0 m/s<sup>2</sup>, to a full standstill in the safest possible place taking into account surrounding traffic/road infrastructure. Higher deceleration demand values are permissible in case of a severe ADS or severe vehicle failure.</p> <p>The ADS shall signal its intention to place the vehicle in an MRC to ADS vehicle occupants as well as to other road users (e.g., by hazard lights)</p> <p>The vehicle can only leave the minimum risk condition only after the confirmation by the on-board operator or remote intervention operator that the cause(s) of the risk manoeuvre is not present anymore.</p>
<p><b>7 Human machine interface for vehicles transporting vehicle occupants</b></p>	<p>Wherever needed for safe operation and with regard to safety of occupant from hazards, adequate information shall be given to the passengers.</p> <p>The ADS shall provide means for vehicle occupants to call a remote intervention operator through an acoustic and a video interface. Unambiguous signs shall be used for the video interface (e.g., ISO7010 E004 )</p> <p>The ADS shall provide means to allow vehicle occupants to request a stop to the ADS. The release of the doors shall be made automatically in case of emergency.</p> <p>The ADS shall provide Camera monitor system to allow the remote intervention operator to understand what is happening inside of the vehicle.</p> <p>It shall be possible for the remote intervention operator to open the power operated service door.</p>
<p><b>8 Functional and operational safety during the ADS lifecycle</b></p>	<p>The manufacturer shall demonstrate that an acceptable consideration of functional and operational safety for the ADS has been done during the design and development processes of the ADS, that the measures put in place by the manufacturer will guarantee that the ADS is free of unreasonable safety risks to vehicle occupants and other road users during the vehicle lifecycle (design, development, production, field operation, decommissioning) when compared with comparable transport services and situations.</p> <p>A safety target for design and development, shall be used. As indicative target, 10<sup>-7</sup> fatalities per hour should be considered as a minimum for applications covered by this regulation. The manufacturer may use other metrics and method provided it can demonstrate that it leads to an equivalent level of safety.</p> <p>The manufacturer shall manage the safety and continued compliance of the vehicles with automated driving function system over lifetime (wear and tear especially for sensors, new traffic scenarios, etc.).</p>
<p><b>9 Specific requirements regarding Cybersecurity and Software-Updates</b></p>	<p>The ADS shall be protected from unauthorized access. The measures ensuring protection from an authorized access shall be provided in alignment with engineering best practices. The effectiveness of the security measures on the ADS shall be demonstrated in compliance with UN Regulation No. 155 during the whole vehicle type-approval.</p>
<p><b>10 Specific requirements regarding data recorder for ADS</b></p>	<p>Vehicles of categories M1 and N1 shall be fitted with an event data recorder system of a vehicle that complies with the technical requirements set out in the 01 Series of Amendments to UN Regulation No 160;</p> <p>In addition, for all vehicle categories each vehicle shall be equipped with a data recorder that at least record an entry for each of the following occurrences upon activation of the ADS:</p>

Request sent to the remote operator
Remote operator request/input
Re (initialisation) of the ADS (if applicable)
Deactivation/over-run of the ADS (if applicable)
Start of Emergency Manoeuvre
End of Emergency Manoeuvre
Involved in a detected collision and crash relevant data
Minimum Risk Manoeuvre engagement by the ADS
ADS failure
Data elements
For each event listed in paragraph 8.2., the data recorder shall at least record the following data elements in a clearly identifiable way:
The recorded occurrence flag
Reason for the occurrence, as appropriate,
Date (Resolution: yyyy/mm/dd);
Position (GPS coordinates)
Timestamp:
Resolution: hh/mm/ss timezone e.g. 12:59:59 UTC
Accuracy: +/- 1.0 s.
For each Recorded occurrence, the RXSWIN, or the software versions, indicating the software that was present at the time when the event occurred, shall be clearly identifiable.
A single timestamp may be allowed for multiple elements recorded simultaneously within the timing resolution of the specific data elements. If more than one element is recorded with the same timestamp, the information from the individual elements shall indicate the chronological order.
Data availability
Once the storage limits of the data recorder are achieved, existing data shall only be overwritten following a first in first out procedure with the principle of respecting the relevant requirements for data availability.
Documented evidence regarding the storage capacity shall be provided by the manufacturer.
For vehicles of Category M <sub>1</sub> and N <sub>1</sub> The data shall be retrievable even after an impact of a severity level set by UN Regulations Nos. 94, 95 or 137.
For vehicles of Categories M <sub>2</sub> , M <sub>3</sub> , N <sub>2</sub> and N <sub>3</sub> , the data elements listed in paragraph 8.3.1 shall be retrievable even after an impact. To demonstrate that capability, the following applies:
Either:

	<p>(a) (a) After a mechanical shock applicable to on-board data storage devices, if any, at a severity level as specified in the component test of Annex 9C of the 03 series of amendment to UN Regulation No. 100, and</p> <p>(b) (b) On-board data storage device(s) shall be mounted in the vehicle cab/passenger compartment or in a position of sufficient structural integrity to protect against physical damage that would prevent the retrieval of data. This shall be demonstrated to the technical service together with appropriate documentation (e.g. calculations or simulations);</p> <p>Or,</p> <p>(c) (c) The manufacturer demonstrates fulfilling the requirements of paragraph 8.4.3.1. (e.g. for M<sub>2</sub> / N<sub>2</sub> vehicles derived from M<sub>1</sub> / N<sub>1</sub>).</p> <p>If the main on-board vehicle power supply is not available, it shall still be possible to retrieve all data recorded on the data recorder.</p> <p>Data stored in the data recorder shall be easily readable in a standardized way via the use of an electronic communication interface, at least through the standard interface (OBD port).</p> <p>Instructions from the manufacturer shall be provided on how to access the data.</p> <p>Protection against manipulation</p> <p>It shall be ensured that there is adequate protection against manipulation (e.g. data erasure) of stored data such as anti-tampering design</p> <p>Availability of the data recorder</p> <p>The data recorder shall be able to communicate with the ADS to inform that the data recorder is operational.</p>
<p><b>11 Manual driving for emergency cases or for the purpose of maintenance or similar cases</b></p>	<p>If manual driving is possible for emergency or for the purpose of maintenance or similar cases is provided in the ADS vehicle, the vehicle shall be provided with means to enable a person driving the vehicle to perform the driving task safely in accordance with the safety concept of the manufacturer.</p> <p>If manual driving control is limited to 6 km/h, it is not necessary for the driver to stay within the vehicle with an autonomous driving function. The control can be performed via a remote control located in the vicinity of the vehicle. The maximum distance over which control is possible by a remote control shall not exceed 6 metres measured in direct, straight connection. Compliance with this maximum distance shall be ensured by appropriate technical means.</p> <p>If, in manual driving, the vehicle is intended to be controlled at speeds higher than 6 km/h, a seating position shall be provided for the person driving the vehicle. It shall be designed in accordance with regulatory acts listed in Annex II to Regulation 858/2018.</p>
<p><b>12 Operation manual</b></p>	<p>The manufacturer shall draw up an operation manual. The purpose of the operations manual is to ensure the safe operation of the vehicle in operation by means of detailed instructions to the owner, vehicle occupants, transport service operator, on board operator, remote intervention operator and public authorities.</p> <p>When manual driving at low speed is possible for emergency cases or for the purpose of maintenance or similar cases, it shall also be covered by the operation manual.</p>

	<p>The operation manual shall include the functional description of the vehicle equipped with the ADS.</p> <p>The operation manual shall include the necessary technical measures (e.g. needed off board infrastructure), operational restrictions (e.g. speed limit, dedicated lane), environmental conditions (e.g. no snow) and operational measures (e.g. on-board operator or remote intervention operator needed) to be met to ensure safety during the ADS vehicle operation.</p> <p>The operational manual shall describe the expected response of vehicle occupants, transport service operator, on board operator and remote intervention operator and public authorities in case of failures and ADS request.</p> <p>The operation manual shall contain rules to ensure proper performance of maintenance, overall tests, further examinations.</p> <p>The Operating Manual shall be submitted to the type approval authority together with the application for a type approval.</p> <p>The Operating Manual shall be made available to the vehicle transport service operator, the owner, needed on-board operator, remote intervention operator, and any relevant public authorities.</p>
<b>13 Provisions for periodic roadworthiness tests</b>	<p>The manufacturer shall ensure the feasibility of periodic roadworthiness testing by taking appropriate measures (e.g.: manual driving, accessibility of brakes). In particular, it shall be able to be tested on brake test benches, it shall have light adjustment positions, etc. for all prescribed tests to be carried out.</p>

*Table 42 Summary of requirements from UNECE draft (Dec. 2021)*



## 10.5 Appendix 5: National Driving Standard – summary of review

The Driver and Vehicle Standards Agency (DVSA) has published a standard that sets out what it takes to be a safe and responsible driver. The standard covers all phases of a journey and sets out what a driver has to be able to do and what they must know and understand in all potential driving situations.

The standard sets out the skills, knowledge and understanding in 6 different roles, with each role setting out

- “performance standards” that set out what a driver must be able to do
- “knowledge and understanding requirements” that describe what a driver must know and understand.

The different roles that are described are

Role 1: Prepare yourself, the vehicle, and its passengers for a journey

Role 2: Guide and control the vehicle

Role 3: Use the road in accordance with The Highway Code

Role 4: Drive safely and responsibly in the traffic system

Role 5: Review and adjust driving behaviour over lifetime

Role 6: Demonstrate developed skills, knowledge and understanding

This standard has been reviewed in order to determine whether the derived requirements could be derived from it for the LSAV Assurance scheme.

Role 1: Prepare yourself, the vehicle, and its passengers for a journey					
Reference		Performance standards You must be able to:	Knowledge and understanding requirements You must know and understand:	Applicable to	Comment
<b>Unit 1.1 Prepare yourself, the vehicle and its passengers for a journey</b>	Element 1.1.1: Choose a suitable mode of transport	<ul style="list-style-type: none"> <li>- assess your own and your passengers' physical, emotional and other needs</li> <li>- assess the environmental impact and cost of other modes of transport</li> <li>- decide whether it's suitable to use a vehicle for the journey</li> </ul>	<ul style="list-style-type: none"> <li>- the pros and cons of different modes of transport, and how each affects the environment</li> <li>- how using a car for very short journeys affects the environment</li> <li>- how vehicle exhaust gases (for example, carbon dioxide, carbon monoxide, sulphur dioxide and lead) affect the environment</li> <li>- the environmental implications of different:                             <ul style="list-style-type: none"> <li>&gt; types of power unit</li> <li>&gt; fuel types</li> <li>&gt; tyres</li> </ul> </li> <li>- how much it costs to own and run different types of vehicles over their life</li> <li>- how vehicle noise can affect the environment</li> </ul>	operator/ passenger not in scope for Type Approval	./.

	<p>Element 1.1.2: Make sure you're fit to drive</p>	<ul style="list-style-type: none"> <li>- assess whether your ability to drive safely and legally is affected or likely to be affected by the use of:             <ul style="list-style-type: none"> <li>&gt; over-the-counter medicines</li> <li>&gt; prescription medicines</li> <li>&gt; illegal or controlled substances</li> <li>&gt; alcohol</li> </ul> </li> <li>assess whether your ability to drive safely and legally is affected by:             <ul style="list-style-type: none"> <li>&gt; your emotional state</li> <li>&gt; a short or long-term physical condition</li> <li>&gt; tiredness</li> </ul> </li> <li>- make other travel arrangements when your ability to drive safely or legally is affected</li> <li>- get help to make any changes needed for you to drive safely and responsibly if you have a long-term physical condition</li> </ul>	<ul style="list-style-type: none"> <li>- what the law says about driving while you have illegal or controlled substances or alcohol in your system</li> <li>- how illegal or controlled substances or alcohol affect your ability to drive safely, and:             <ul style="list-style-type: none"> <li>&gt;that regardless of any legal limits, it's best to have no alcohol in your system</li> <li>&gt;how the strength of alcohol varies in different types of drink</li> <li>&gt;what a 'unit' of alcohol is equivalent to in different types of drink</li> <li>&gt;how the body processes drugs and alcohol and the rate at which they're removed from your system</li> <li>&gt;that any alcohol can make you more likely to fall asleep, even if the levels in your blood are below the legal limit</li> </ul> </li> <li>- how over-the-counter or prescription medicines can affect your ability to drive safely</li> <li>- the risks linked to any combination of:             <ul style="list-style-type: none"> <li>&gt;over-the-counter medicines</li> <li>&gt;prescription medicines</li> <li>&gt;illegal or controlled substances</li> <li>&gt;alcohol</li> </ul> </li> <li>- that any remedy or medicine with instructions that say 'may cause drowsiness' is highly likely to cause drowsiness</li> <li>- the range of possible solutions to help people with long-term physical conditions drive safely and responsibly</li> <li>- how being tired before or during your journey affects your ability to drive, and:             <ul style="list-style-type: none"> <li>&gt;how a poor seating position and bad posture can make you tired</li> <li>&gt;that a poor diet or eating food at the wrong time can make you more likely to fall asleep</li> <li>&gt;that there are times of the day when people are likely to feel more sleepy</li> </ul> </li> <li>- how emotional states (like anger, grief, sadness and joy) can affect your ability to drive safely</li> </ul>	<p>not applicable for Automated Driving</p>	<p>./.</p>
--	---	---	---	---	------------

			<ul style="list-style-type: none"><li>- that being careless, thoughtless and/or reckless are frequent causes of crashes</li><li>- how a short-term injury (like a sprained ankle) can affect your ability to drive safely</li><li>- that eyesight gets worse over time, and that not realising or doing anything about it can affect your ability to drive safely and legally</li><li>- the need to have an eyesight test at least every 2 years</li><li>- that you must wear glasses or contact lenses all the time when driving if you need them to meet the driving eyesight rules</li><li>- how different sorts of tinted and light-sensitive lenses or visors react in different driving conditions</li><li>- that changes to your physical and mental abilities, particularly as you get older, can affect your ability to drive safely (such as slower reaction times or reduced muscle strength)</li><li>- how to make other travel plans when your ability to drive safely or legally is affected</li></ul>		
--	--	--	--	--	--

	<p>Element 1.1.3: Control the risks linked with carrying passengers, loads and animals</p>	<ul style="list-style-type: none"> <li>- manage how passengers affect your ability to drive safely</li> <li>- make sure passengers are seated legally, correctly and securely</li> <li>- make sure loads are secure and distributed according to the manufacturer's guidelines</li> <li>- allow for the effect that extra loads may have on how the vehicle handles</li> <li>- make sure animals are secure and correctly restrained within the vehicle</li> </ul>	<ul style="list-style-type: none"> <li>- the law for fitting and using seatbelts</li> <li>- the law for fitting and using baby seats, child seats, booster seats and booster cushions</li> <li>- the importance of using head-restraints, where fitted, and of adjusting them correctly</li> <li>- the correct use of airbags (such as when using a baby seat)</li> <li>- the law on the carriage of loads on the outside of the vehicle</li> <li>- how to use the vehicle handbook to identify how best to safely load the vehicle</li> <li>- what types of load-carrying and securing equipment you can use with the vehicle and how to fit and use them</li> <li>- how to restrain animals safely</li> <li>- how to make sure that you can still see clearly if windows or mirrors are blocked by passengers or by a load</li> <li>- how to adjust the vehicle to allow for extra weight and changed weight distribution</li> <li>- how to adjust your driving behaviour to allow for extra weight or changed weight distribution</li> <li>- how to deal with social pressure and distractions that passengers cause</li> </ul>	<p>partially applicable: manufacturer / operator</p>	<p>Type Approval regulation to consider requirements for passenger safety and permissible load weight (WP4) as well as for ADS (WP1). For goods transport the operator is responsible for safe loading of the vehicle in line with its permitted pay load.</p> <p>Addressed by Technical Requirements 11, 12, 17 and to be further addressed in operator SMS (WP1)</p>
--	--	--	--	--	--

<p><b>Unit 1.2 Make sure the vehicle is safe to drive</b></p>	<p>Element 1.2.1: Make routine checks that your vehicle's safe to drive</p>	<ul style="list-style-type: none"> <li>- check all fluid levels, including windscreen washer reservoir(s)</li> <li>- check that the horn is working correctly</li> <li>- check that all lights and reflectors are:                             <ul style="list-style-type: none"> <li>&gt; legal</li> <li>&gt; clean</li> <li>&gt; in good working order</li> </ul> </li> <li>- check electrical equipment is in good working order</li> <li>- check there is no damage that would:                             <ul style="list-style-type: none"> <li>&gt; affect your ability to drive the vehicle safely</li> <li>&gt; make the vehicle illegal</li> <li>&gt; have an adverse environmental impact</li> </ul> </li> <li>- check all tyres, including any spare, are:                             <ul style="list-style-type: none"> <li>&gt; legal</li> <li>&gt; correctly inflated</li> </ul> </li> <li>- check any equipment, such as the car jack, is in good working order</li> <li>- check all controls are in good working order</li> <li>- check windscreen, mirrors and other viewing devices are clear and adjusted to give the best view</li> <li>- check registration plates are:                             <ul style="list-style-type: none"> <li>&gt; fitted</li> <li>&gt; visible</li> <li>&gt; legal</li> </ul> </li> <li>- check that any ancillary equipment (like aftermarket sat nav systems or 'head-up' displays) is legal to use in the vehicle and securely fitted in a position that minimises distraction to you</li> <li>- make sure checks are carried out by a competent person where you are unable or unwilling to carry them out yourself</li> </ul>	<ul style="list-style-type: none"> <li>- that different vehicles may permit different levels of access to check and maintain fluid levels, check electric systems etc, and some checks or maintenance on some vehicles should only be carried out by qualified mechanics</li> <li>- that the vehicle handbook identifies which checks can be carried out by the owner or user and explains how and when to carry them out, either directly or using the vehicle's instrumentation</li> <li>- that overfilling with engine oil can:                             <ul style="list-style-type: none"> <li>&gt;damage your engine</li> <li>&gt;increase the amount of environmental pollution the vehicle creates</li> </ul> </li> <li>- that using oil that isn't to the manufacturer's specification:                             <ul style="list-style-type: none"> <li>&gt;can increase fuel consumption</li> <li>&gt;may cause damage</li> <li>&gt;could affect the vehicle warranty</li> </ul> </li> <li>- what fluids to add to the vehicle coolant system and the need to maintain the level of coolant additive</li> <li>- how to check that tyres:                             <ul style="list-style-type: none"> <li>&gt;are correctly fitted and inflated</li> <li>&gt;meet legal requirements for tread depth</li> <li>&gt;are free from defects that would make them unsafe or illegal to use</li> </ul> </li> <li>- the rules that apply to the fitting of different types of tyres</li> <li>- that tyres specially adapted for different weather conditions are available (such as winter tyres or all-season tyres)</li> <li>- that the operation of any equipment could result in the driver taking their eyes off the road</li> <li>- how to spot signs of abnormal tyre wear and the need to have the vehicle checked if abnormal wear is found</li> <li>- that the windscreen and other windows should be clean and free from obstructions and that there are legal limits to the amount</li> </ul>	<p>partially applicable: operator</p>	<p>In scope for licensing process and to be addressed as part of the requirements to be covered in an operator's SMS (WP1) as well as partially addressed by Technical Requirements 8 &amp; 22.</p>
---	---	--	---	---------------------------------------	---

			<p>and location of damage to windscreens, beyond which they must be replaced</p> <ul style="list-style-type: none"><li>- that lights, indicators, reflectors and number plates must be clean at all times</li><li>- any rules that apply to the fitting and use of ancillary equipment and how to make sure it can be used safely and with the minimum of distraction</li><li>- what electrical equipment to check</li><li>- what controls to check</li><li>- the legal need to dispose of or recycle oil, batteries and tyres correctly</li></ul>		
--	--	--	--	--	--

	<p>Element 1.2.2: Check the vehicle is fit for the journey</p>	<ul style="list-style-type: none"> <li>- familiarise yourself with the vehicle if it is the first time you have driven it</li> <li>- conduct pre-journey checks and configure the vehicle correctly</li> <li>- make changes to your driving position so that you:                             <ul style="list-style-type: none"> <li>&gt; are safely and comfortably seated</li> <li>&gt; have good all-round visibility</li> <li>&gt; have control of the vehicle</li> <li>&gt; minimise tiredness</li> </ul> </li> <li>- check there is enough fuel of the right type</li> </ul>	<ul style="list-style-type: none"> <li>- what pre-journey checks are needed and what adjustments to make</li> <li>- the effect of filling a vehicle with the wrong sort of fuel</li> <li>- how to check what sort of fuel your vehicle uses</li> <li>- the operation of low-fuel, mpg or range indicators and how much fuel is left in the tank when low-fuel indicators operate</li> </ul>	<p>partially applicable: operator</p>	<p>In scope for licensing process and to be addressed as part of the requirements to be covered in an operator's SMS (WP1)</p>
--	--	--	---	---	--



	<p>Element 1.2.3: Make sure the vehicle's documents meet the legal requirements</p>	<ul style="list-style-type: none"> <li>- make sure your driving licence is valid for the category of vehicle being driven</li> <li>- make sure the vehicle is registered and taxed</li> <li>- make sure you have valid insurance for the use you intend to make of the vehicle</li> <li>- make sure that the vehicle has a current MOT certificate (where applicable)</li> <li>- display red L plates (or if you wish, red D plates in Wales) if you are a provisional licence holder</li> <li>- make sure that the correct documents are in place even if you don't own the vehicle</li> <li>- where your journey will take you into an area where different rules apply, make sure that you follow those rules</li> </ul>	<ul style="list-style-type: none"> <li>- that you must:             <ul style="list-style-type: none"> <li>&gt; have a valid driving licence for the vehicle you drive</li> <li>&gt; meet any restrictions on your licence</li> </ul> </li> <li>- that learner drivers, holding a provisional licence, must be supervised by somebody who:             <ul style="list-style-type: none"> <li>&gt; is at least 21 years old, and</li> <li>&gt; has held a licence to drive the category of vehicle for at least 3 years</li> </ul> </li> <li>- that any vehicle driven by a learner must clearly display legal, red L plates (or in Wales either red L or red D plates, or both)</li> <li>- that L (D) plates should be removed when a vehicle is not being driven by a learner</li> <li>- that the vehicle must be registered with the Driver and Vehicle Licensing Agency (DVLA)</li> <li>- the law on the taxation of vehicles and the need to make a statutory declaration (SORN) if you take the vehicle off the road and stop taxing it for any period of time</li> <li>- that you must notify the DVLA if you:             <ul style="list-style-type: none"> <li>&gt; change your name or address</li> <li>&gt; have or develop a medical condition that will affect your ability to drive</li> <li>&gt; buy or sell a vehicle</li> <li>&gt; make any substantive changes to your vehicle</li> </ul> </li> <li>- that you must have a minimum of third party insurance covering you for the intended use of the vehicle, and what insurance companies require you to do to meet your obligations under that insurance</li> <li>- that you must hold a valid MOT test certificate for the vehicle if it is more than 3 years old</li> <li>- that, if required by an authorised person, you must be able to produce:             <ul style="list-style-type: none"> <li>&gt; your driving licence</li> <li>&gt; a valid insurance certificate</li> <li>&gt; a current MOT certificate either immediately or within seven days to a police</li> </ul> </li> </ul>	<p>partially applicable: manufacturer / operator</p>	<p>In scope for licensing process and to be addressed as part of the requirements to be covered in an operator's SMS (WP1)</p>
--	---	---	--	--	--

			<p>station</p> <ul style="list-style-type: none"><li>- that if you borrow or rent a vehicle you still must make sure that you have the correct documents</li><li>- that if you lend somebody your vehicle you still must make sure that they have the correct documents</li><li>- that if you drive outside Great Britain there may be different document rules, like a need to have your documents with you at all times</li></ul>		
--	--	--	---	--	--

<p><b>Unit 1.3 Plan a journey</b></p>	<p>Element 1.3.1 Plan a journey</p>	<ul style="list-style-type: none"> <li>- plan a suitable route taking into account:                             <ul style="list-style-type: none"> <li>&gt; road conditions</li> <li>&gt; weather conditions</li> <li>&gt; traffic</li> <li>&gt; driving experience</li> <li>&gt; the vehicle you are using</li> </ul> </li> <li>- work out the time needed to complete your journey safely and legally, including rest breaks and refuelling stops</li> <li>- decide whether it is safe to make a journey in poor weather conditions</li> <li>- consider other routes if your planned route is blocked, or if weather conditions make it unsafe to continue</li> <li>- program any sat nav systems before you start your journey so that you're not distracted while driving</li> <li>- be prepared for the possibility that your journey may be delayed or affected by poor weather conditions, by taking:                             <ul style="list-style-type: none"> <li>&gt; suitable clothing</li> <li>&gt; equipment</li> <li>&gt; food and drink</li> </ul> </li> <li>- plan where you intend to park at the end of your journey</li> </ul>	<ul style="list-style-type: none"> <li>- the principles of mapping, the technologies available for route planning and for monitoring road traffic conditions, and the limitations of these technologies</li> <li>- the need to build in extra time to allow for unforeseen delays</li> <li>- how congestion charges and road and bridge tolls may affect your choice of route</li> <li>- how the risks involved in travelling on some routes can change at different times, such as:                             <ul style="list-style-type: none"> <li>&gt; heavier traffic at rush hour or in the holiday season</li> <li>&gt; lower stability on exposed routes in windy conditions</li> </ul> </li> <li>- the link between your level of skill and experience and whether you should choose a particular route</li> <li>- how to get information on likely weather conditions and how they might affect your journey</li> <li>- when using sat nav systems:                             <ul style="list-style-type: none"> <li>&gt; how to program them</li> <li>&gt; the information they can provide</li> <li>&gt; that they can sometimes fail, and how to prepare for that happening</li> </ul> </li> <li>- the importance of minimising distractions while driving</li> <li>- how to find safe, secure, legal and convenient places to park</li> </ul>	<p>fully applicable: manufacturer / ADS / operator</p>	<p>The safety goals that the type approval regulation is based on are setting out requirements that ensure the ADS will be capable of performing the DDT safely in the conditions and for the environment that it is designed for (ODD), including strategies (MRMs and MRCs) in case of unplanned events (ODD exits) . This will be assessed at type approval. The compliance of the actual target operating domain to the ODD and any required in-service monitoring of the conditions and the environment will be part of the operator's responsibilities, which will be set out as requirements for the operator's SMS. (WP1)</p>
---------------------------------------	---	--	---	--	---

Table 43 National Driving Standard Role 1 Summary

**Role 2: Guide and control the vehicle**

Reference		Performance standards You must be able to:	Knowledge and understanding requirements You must know and understand:	Applicable to	Comment
<p><b>Unit 2.1</b> <b>Start, move off, stop and leave the vehicle safely and responsibly</b></p>	<p>Element 2.1.1: Start the vehicle</p>	<ul style="list-style-type: none"> <li>- carry out pre-start checks on:                             <ul style="list-style-type: none"> <li>&gt; doors</li> <li>&gt; parking brake</li> <li>&gt; seat</li> <li>&gt; steering</li> <li>&gt; seatbelt</li> <li>&gt; mirrors</li> </ul> </li> <li>- disengage anti-theft devices</li> <li>- make sure the gear lever is in neutral (or 'P' or 'N' if driving an automatic vehicle)</li> <li>- consider the effect of starting the engine on other road users, particularly vulnerable road users such as passing cyclists, pedestrians or horse riders</li> <li>- monitor vehicle instruments and gauges during engine start up</li> <li>- respond correctly to information given by instruments and gauges during engine start up</li> <li>- start the engine correctly</li> </ul>	<ul style="list-style-type: none"> <li>- how to read and respond correctly to instruments, like:                             <ul style="list-style-type: none"> <li>&gt; gauges</li> <li>&gt; indicators</li> <li>&gt; warning lights</li> </ul> </li> <li>- on-board diagnostic systems and other aids fitted to the vehicle to allow you to monitor its operation and performance</li> <li>- that different vehicles may have different starting mechanisms, types of instrument, parking brakes and other aids, and that it is vital to use the vehicle handbook to find out how they work</li> <li>- how to start the engine when it is cold</li> <li>- the benefits of using anti-theft devices, and how turn them on and off</li> </ul>	<p>fully applicable: manufacturer / ADS / operator</p>	<p>WP1: As the responsibilities of a human driver will be shared between the ADS and the operator there will be requirements on the ADS to ensure that all vehicle systems required for the DDT are in functioning order prior to a journey. This task may be shared with in-service checks that an operator must perform prior to the commencement of each service operation. This requirement will be established as part of the operator's SMS.</p>

	<p>Element 2.1.2: Move off safely and smoothly</p>	<ul style="list-style-type: none"> <li>- carry out all-round visual checks, including blind spots, to make sure that it is safe to move-off</li> <li>- signal your intention to move off to other road users, where needed</li> <li>- move off straight-ahead, on the level and on slopes, safely and smoothly keeping control of the vehicle at all times</li> <li>- move off at an angle from behind a parked vehicle or obstruction, safely and smoothly, keeping control of the vehicle at all times</li> <li>- check that controls are operating correctly</li> <li>- restart quickly and safely if the vehicle stalls</li> </ul>	<ul style="list-style-type: none"> <li>- the importance of carrying out all-round, effective observation of the road and other road users before moving off</li> <li>- the importance and location of blind spots and how to carry out blind spot checks before moving away</li> <li>- the importance of using a safe, systematic routine to help you to move off safely and smoothly</li> <li>- the importance of applying the footbrake before selecting drive on an automatic vehicle</li> <li>- where applicable, the relevance of the 'biting point', that is the point at which the clutch plate and the flywheel come into firm contact and start to transmit drive</li> <li>- the operation of the parking brake release mechanism</li> <li>- the limitations of hill assist systems, where fitted</li> <li>- the effects of 'dry steering', that is turning the wheels when the vehicle is not moving</li> <li>- how to check controls, such as steering and brakes, are operating correctly</li> </ul>	<p>fully applicable: manufacturer / ADS</p>	<p>WP1: The ADS will be responsible for the DDT within its ODD and type approval will set out the technical requirements for safe driving, including moving off safely. (see operational/ control level behavioural competence "Pull Away from Standstill") (Section 3.3)</p>
--	--	--	--	---	---

	<p>Element 2.1.3: Decelerate and bring the vehicle to a stop safely</p>	<p>- use the accelerator and brakes correctly to regulate speed and bring the vehicle to a stop safely- stop the vehicle safely and under control in an emergency- use the parking brake when stationary, where needed</p>	<p>- how to apply a safe, systematic approach when stopping- the distance a vehicle requires to stop from different speeds and in different road and weather conditions- that a vehicle's overall stopping distance consists of 2 parts: &gt; thinking distance - which is the distance travelled from the point where you decide to brake to the point where you start braking &gt; braking distance - which is the distance travelled from the point where you start to brake to the point where you stop- the importance of anticipation and judgement to allow for progressive use of the brakes- how aids such as an Anti-lock Braking System (ABS) can help in safe and effective braking</p>	<p>fully applicable: manufacturer / ADS</p>	<p>WP1: The ADS will be responsible for the DDT within its ODD and type approval will set out the technical requirements for safe driving, including longitudinal deceleration control. (see operational/ control level behavioural competence "Pull Away from Standstill") (Section 3.3)</p>
--	---	--	---	---	---

	<p>Element 2.1.4: Park the vehicle safely and responsibly</p>	<ul style="list-style-type: none"> <li>- select a safe, legal and convenient place to stop and park and, once stationary, secure the vehicle on slopes, facing both up and down, as well as on the level</li> <li>- make sure the parking brake is applied effectively</li> <li>- select a gear to hold the vehicle safely when parked</li> <li>- switch the engine off</li> <li>- make sure that vehicles fitted with automatic transmission are left with the lever in the Park position</li> <li>- make sure lights are left on where required</li> <li>- check for oncoming cyclists, pedestrians and other traffic before opening your door</li> </ul>	<ul style="list-style-type: none"> <li>- what factors to take into consideration when looking for a safe, legal and convenient place to stop or park</li> <li>- the pros and cons of reversing or 'pulling through' into a parking space rather than reversing out</li> <li>- that you must switch off the headlights, fog lights if fitted and engine when parked</li> <li>- the rules in The Highway Code that apply when leaving your vehicle on different roads and in different lighting and weather conditions</li> <li>- how and when to set the position of the steering wheels of the vehicle to prevent it rolling away</li> <li>- how to make sure that the parking brake is applied effectively</li> <li>- that, when parking a vehicle with manual gears, selecting a gear will help to hold the vehicle if the parking brake should fail</li> <li>- the possible outcomes of opening a door, particularly on the offside of the vehicle, when not safe to do so</li> </ul>	<p>fully applicable: manufacturer / ADS / operator</p>	<p>WP1: The ADS will be responsible for the DDT within its ODD and type approval will set out the technical requirements for safe driving, including parking the vehicle (see tactical/ manoeuvre level behavioural competence "Park Vehicle") (Section 3.3) Operator requirements (to be covered in operator SMS) might include setting out permitted parking areas if applicable.</p>
--	---	---	--	--	---

<p><b>Unit 2.2 Drive the vehicle safely and responsibly</b></p>	<p>Element 2.2.1: Monitor and respond to information from instrumentation, driving aids and the environment</p>	<p>- monitor and respond correctly to gauges, warning lights and other aids when driving- monitor and respond appropriately to instructions provided by sat nav systems without being distracted from the driving task- respond to the actual situation on the road ahead- make effective use of driving aids such as adaptive cruise control, daytime running lights, automatic headlights and lane warning systems and override or disable them if it is safer to do so- make effective use of mirrors and other aids to vision to identify and monitor other road users and hazards- judge speed and distance correctly and effectively- signal your intentions correctly to other road users in a safe and systematic way- use the vehicle's lights, indicators and horn correctly- use the windows, wipers, demisters and climate and ventilation controls so that you can see clearly</p>	<p>- the purpose and meaning of dashboard warning lights- the location of switches and controls and how to use them without being distracted or losing control of the vehicle while on the move- that you must always act on the basis of what is in front of you and not just rely on the information provided by sat nav systems or other aids- when it is safer to override or disable driving aids- when and how to use dipped headlights- the rules that apply to the use of fog lights- how different types of mirror can make other road users appear to be nearer or further away than they actually are- how to identify and respond to changes in road surfaces and weather conditions</p>	<p>partially applicable: manufacturer / ADS / operator</p>	
---	---	---	--	--	--



	<p>Element 2.2.2: Control the acceleration of the vehicle effectively</p>	<p>- use the accelerator smoothly to achieve and maintain a suitable speed</p>	<p>- that correct use of the accelerator will help:          &gt; vehicle performance          &gt; safety          &gt; the environment          - the disadvantages of over-revving when moving away and while stationary          - how to operate cruise control systems safely, if fitted          - the importance of using a driving position that allows you to use the accelerator smoothly</p>	<p>fully applicable:          manufacturer / ADS / operator</p>	
	<p>Element 2.2.3: Use gears correctly</p>	<p>- change gear smoothly and in good time- select the most suitable gear for the speed of the vehicle, given road and traffic conditions- combine the use of gears with braking and acceleration- use an automatic or automated gear box effectively, when fitted</p>	<p>- that different vehicles may have different numbers of gears and those gears may be set up differently- the effect that unsuitable gear selection can have on:          &gt; the performance of the vehicle          &gt; the driver's ability to drive safely and responsibly          &gt; the environment- the use of selective gear changing (sometimes known as block changing)- the benefits of timely gear selection when going up and down slopes, particularly when loaded- the use of 'kick down' and 'lock up' when using an automatic vehicle- how to use gears to assist safe parking- the difference between automatic and automated gearboxes</p>	<p>partially applicable:          manufacturer / ADS / operator</p>	<p><i>As only EV applications are planned, the gear control is limited to selection of forward or reverse movement.</i></p> <p>WP1: The ADS will be responsible for the DDT within its ODD and type approval will set out the technical requirements for safe driving, including Longitudinal Acceleration and Deceleration Control (see operational/ control level behavioural competence "Perform Long accel or decel control (FW or REV)") (Section 3.3)</p> <p>The behavioural competence to reverse the LSAV is included in Technical Requirement [2].</p>

	<p>Element 2.2.4: Steer the vehicle safely</p>	<ul style="list-style-type: none"> <li>- steer the vehicle safely and responsibly in all road and traffic conditions</li> <li>- hold and control the steering wheel to steer the vehicle accurately and safely</li> <li>- continue to steer the vehicle safely and responsibly while operating other controls</li> </ul>	<ul style="list-style-type: none"> <li>- how to keep safe control of the steering wheel</li> <li>- the effect that the vehicle's turning circle has on steering the vehicle</li> </ul>	<p>fully applicable: manufacturer / ADS / operator</p>	<p>WP1: The ADS will be responsible for the DDT within its ODD and type approval will set out the technical requirements for safe driving, including Lateral Steering Control (see operational/ control level behavioural competence "Perform Lateral steering control (FW or REV") (Section 3.3)</p> <p>The behavioural competence to reverse the LSAV is included in Technical Requirement [2].</p>
--	--	--	--	--	---

	<p>Element 2.2.5: Manoeuvre the vehicle</p>	<p>- make proper use of all controls to manoeuvre the vehicle safely and responsibly in: &gt; all road and weather conditions &gt; forward and reverse gear- continue to make effective observations, including checks of blind spots, - while manoeuvring- position the vehicle correctly to carry out manoeuvres safely- use a safe and systematic way to keep yourself and other road users safe, such as 'mirrors, signal, manoeuvre, position, speed, look'- use reversing camera systems or proximity sensors effectively, where fitted</p>	<p>- how the use of safe, systematic routines will contribute to safe and responsible manoeuvring- the blind spots for the vehicle and how to check them- the correct procedure: &gt; for reversing into a side road on the left or right &gt; to carry out a turn-in-the-road or U-turn manoeuvre &gt; for carrying out any reverse parking exercise on and off road- the rules about when and where you cannot make U-turns- the effects of sudden or harsh use of the accelerator, brakes or steering whilst manoeuvring- that different vehicles will react differently in a possible skid situation depending on their configuration (such as front-wheel or rear-wheel drive) and on the technologies fitted (such as ABS or electronic stability program (ESP))- why a skid may occur, how to avoid skids and how to correct them if they do occur- how to allow for vulnerable road users when carrying out a manoeuvre- the benefits of engine braking and when to use it- the risks linked to reversing a vehicle further than necessary- the risks linked to 'coasting'- how to identify a suitable place for manoeuvring- that use of reversing aids, such as camera systems and proximity sensors, does not replace the need to practise good, all-round, effective observation</p>	<p>fully applicable: manufacturer / ADS / operator</p>	<p>WP1: The ADS will be responsible for the DDT within its ODD and type approval will set out the technical requirements for safe driving, including parking the vehicle</p> <p>(see tactical/ manoeuvre level behavioural competence "Park Vehicle") (Section 3.3)</p> <p>Operator requirements (to be covered in operator SMS) might include setting out permitted parking areas if applicable.</p> <p>This is covered in the Technical Requirements 1-5)</p>
--	---	---	--	--	---

<p><b>Unit 2.3</b>  <b>Drive the vehicle while towing a trailer or caravan</b></p>	<p>Element 2.3.1 Drive the vehicle while towing a trailer or caravan</p>	<p>- make sure you have the correct licence to drive the combination of vehicle and trailer or caravan- make sure that the trailer or caravan is suitable and legal for use on the road- make sure that you are insured to drive the combination of vehicle and trailer or caravan- make sure that your vehicle is capable of towing the trailer or caravan- make sure that the trailer or caravan is safely and correctly coupled to the vehicle- carry out correct safety checks- make sure that any load is evenly distributed and secure- allow more time and brake earlier when slowing down or stopping- allow more distance and time to overtake safely- make allowances for the extra length of the vehicle with the trailer or caravan, particularly when turning or emerging at junctions- safely and correctly uncouple the trailer or caravan from the vehicle when it is no longer needed- reverse the vehicle with the trailer or caravan attached</p>	<p>- the driving licence regulations on towing trailers or caravans- that not all insurance policies cover towing a trailer or caravan- that most manufacturers make recommendation for the maximum size of trailer or caravan that can be safely towed by each type of vehicle, and for how they should be attached, and that these recommendations must be followed- how to find the trailer or caravan's 'nose weight' and how to check that this does not exceed the limits of the vehicle's tow bar- how to couple and uncouple a trailer or caravan safely- that towing a trailer or caravan may increase the number of blind spots- how and when to use aids to observation, such as extra mirrors- what safety checks should be made on a trailer or caravan- the speed limits when towing a trailer or caravan- that vehicles towing trailers on motorways are not allowed in the outside lane where there are 3 or more lanes- that towing a trailer or caravan will change the way a vehicle handles, and how to deal with those changes</p>	<p>not in scope of the LSAV Assurance Scheme</p>	<p>./.</p>
--	--	--	---	--	------------

			<ul style="list-style-type: none"> <li>- that it may be necessary to take up a different position on the road when dealing with junctions or roundabouts</li> <li>- what 'snaking' is and how to correct it</li> <li>- that strong winds pose a particular hazard for caravans or high-sided trailers</li> <li>- how to steer correctly when reversing a vehicle with a trailer or caravan attached</li> <li>- the effect that towing a trailer or caravan may have on braking, the concept of brake fade and what to do when descending slopes to make sure you keep in control</li> <li>- that you may have to check height or width restrictions on your route when you tow a trailer or caravan</li> <li>- that rescue services may not include recovery of a trailer or caravan</li> <li>- the benefits of carrying a spare wheel and any other equipment for the trailer or caravan</li> </ul>		
--	--	--	--	--	--

*Table 44 National Driving Standard Role 2 Summary*

**Role 3: Use the road in accordance with The Highway Code**

Reference		Performance standards You must be able to:	Knowledge and understanding requirements You must know and understand:	Applicable to	Comment
<b>Unit 3.1 Negotiate the road correctly</b>	Element 3.1.1: Maintain a suitable position on the road	<ul style="list-style-type: none"> <li>- select and maintain a suitable position on the road</li> <li>- change lanes safely and responsibly</li> <li>- overtake other road users legally, safely and responsibly</li> </ul>	<ul style="list-style-type: none"> <li>- how to select a suitable position on the road</li> <li>- where you may not drive, for example on the pavement, hard shoulder or in cycle lanes</li> <li>- what lane discipline is and why it is important</li> <li>- that your position on the road may be affected by a range of factors including weather, road and traffic conditions</li> <li>- the importance of:                             <ul style="list-style-type: none"> <li>&gt; scanning the road ahead for reasons to change your position, such as roadworks</li> <li>&gt; taking timely action to reposition yourself</li> </ul> </li> <li>- how to use a safe and systematic way to change position safely and responsibly in time</li> <li>- how the performance and handling of your vehicle will affect your ability to overtake safely and responsibly</li> <li>- where you may and may not overtake</li> </ul>	fully applicable: manufacturer / ADS	WP1: The ADS will be responsible for the DDT within its ODD and type approval will set out the technical requirements for safe driving, including lateral steering control and maintaining lateral position in lane.  (see operational/ control level)

	<p>Element 3.1.2: Negotiate bends</p>	<ul style="list-style-type: none"> <li>- assess bends correctly on approach</li> <li>- select a safe position and speed to enter a bend</li> <li>- maintain safe speed and control throughout a bend</li> <li>- exit bends safely</li> </ul>	<ul style="list-style-type: none"> <li>- how to use various methods such as 'limit point analysis' to judge the severity of a bend</li> <li>- that when deciding on the line to take and the speed at which it is possible to negotiate a bend safely you should take into account factors such as:                             <ul style="list-style-type: none"> <li>&gt; adverse camber</li> <li>&gt; banking</li> <li>&gt; uneven or slippery surfaces</li> <li>&gt; weather conditions</li> <li>&gt; visibility</li> <li>&gt; road junctions</li> <li>&gt; other road users</li> <li>&gt; that different vehicles will perform and handle differently through bends</li> </ul> </li> <li>- the importance of coordinating the use of gears, accelerator, brakes and steering to negotiate a bend safely and responsibly</li> <li>- how the use of a safe and systematic way to negotiate bends safely</li> <li>- the effect that loads and passengers may have on the handling of the vehicle through bends</li> </ul>	<p>fully applicable: manufacturer / ADS</p>	<p>behavioural competence "Perform lateral steering control" and tactical/manoeuvre level behavioural competence ("Maintain lateral position in lane") (Section 3.3)  This is covered in the Technical Requirements 1-5)</p>
--	---	--	---	---	--

	<p>Element 3.1.3: Negotiate all types of junctions, including roundabouts, and all types of crossings</p>	<ul style="list-style-type: none"> <li>- apply a safe and systematic way to negotiate all types of junctions, roundabouts and crossings safely and responsibly</li> <li>- actively scan for more vulnerable road users at junctions, roundabouts and crossings ‘ for example cyclists and motorcyclists</li> <li>- turn left and right and go ahead safely and responsibly</li> <li>- emerge safely and responsibly into streams of traffic</li> <li>- cross the path of traffic safely when turning right</li> </ul>	<ul style="list-style-type: none"> <li>- the rules that apply to particular junctions and roundabouts, such as priority rules</li> <li>- how to turn left and right safely and responsibly</li> <li>- the issues that apply to turning right at crossroads</li> <li>- the rules that apply to:                             <ul style="list-style-type: none"> <li>&gt; merging into a stream of traffic</li> <li>&gt; crossing the path of an approaching stream of traffic</li> <li>&gt; all types of pedestrian crossing</li> <li>&gt; train and tram crossings</li> </ul> </li> <li>- the meaning of warning lights used at pedestrian and train and tram crossings and how to respond correctly</li> <li>- how the use of a safe, systematic routine, including effective observations, will support the safe negotiation of junctions, roundabouts and crossings</li> <li>- the rules that apply to other road users, particularly drivers of large vehicles or vulnerable road users such as cyclists and motorcyclists, and the position that they may select on the road as a result</li> </ul>	<p>fully applicable: manufacturer / ADS</p>	<p>WP1: The ADS will be responsible for the DDT within its ODD and type approval will set out the technical requirements for safe driving including negotiating crossings and junctions</p> <p>(see Secondary Manoeuvre Capabilities ("Negotiate intersections" and Negotiate crossings") (Section 3.3))</p> <p>This is covered in the Technical Requirements 1-5))</p>
--	---	---	---	---	---



	<p>Element 3.1.4: Drive on motorways and dual carriageways</p>	<ul style="list-style-type: none"> <li>- join a motorway or dual carriageway safely and responsibly from the left or the right</li> <li>- leave a motorway or dual carriageway safely and responsibly to the left or the right</li> <li>- drive in the most suitable lane</li> <li>- allow for other road users joining or leaving the motorway or dual carriageway</li> <li>- change lanes safely and responsibly</li> </ul>	<ul style="list-style-type: none"> <li>- how to join a motorway or dual carriageway, safely and responsibly, from traffic light controlled or uncontrolled slip roads</li> <li>- how to leave a motorway or dual carriageway safely and responsibly, including the need to position yourself well in advance to allow other road users enough time to react</li> <li>- how to join or leave a motorway or dual carriageway safely in a safe way systematic way</li> <li>- that you may not stop on a motorway except in an emergency</li> <li>- when and for what purposes you are allowed to use the hard-shoulder</li> <li>- that you mustn't pick up or set down anybody, or walk on a motorway, except in an emergency</li> <li>- that you mustn't cross the central reservation, or drive against the traffic flow on a motorway or dual carriageway, unless directed to do so by an authorised person or traffic signs</li> <li>- the rules that apply when using a motorway or dual carriageway</li> <li>- that some stretches of motorway may have local, active traffic management (also known as smart motorways or managed motorways) control systems installed, which will change speed limits or the direction of flow in particular lanes, and that it is vital to obey the instructions given by such systems</li> <li>- the need to scan well ahead on the approach to junctions to make sure you are aware of:             <ul style="list-style-type: none"> <li>&gt; other road users joining or leaving</li> <li>&gt; queuing traffic</li> </ul> </li> <li>- the correct use of hazard warning lights</li> <li>- the risks posed by drivers of left-hand-drive vehicles, in particular large goods vehicles</li> </ul>	<p>not in scope of the LSAV Assurance Scheme</p>	<p>./.</p>
--	--	---	--	--	------------

<p><b>Unit 3. 2 Comply with signals, signs and road markings Make sure the vehicle is safe to drive</b></p>	<p>Element 3.2.1: Comply with signals, signs and road markings</p>	<ul style="list-style-type: none"> <li>- respond correctly to all permanent and temporary traffic signals, signs and road markings</li> <li>- respond correctly to signals given by authorised persons</li> <li>- respond safely and responsibly to signals given by other road users</li> </ul>	<ul style="list-style-type: none"> <li>- the meaning of, and how to respond to:                             <ul style="list-style-type: none"> <li>&gt; mandatory traffic signs</li> <li>&gt; warning signs</li> <li>&gt; road markings</li> </ul> </li> <li>- how to work out the speed limit when you can't see speed limit signs</li> <li>- the meaning of, and how to respond correctly to, signals given by:                             <ul style="list-style-type: none"> <li>&gt; police officers</li> <li>&gt; crossing patrols</li> <li>&gt; others authorised to control traffic</li> </ul> </li> <li>- who is authorised to control traffic</li> <li>- signals that other road users are likely to use and how to respond safely and responsibly to them</li> </ul>	<p>fully applicable: manufacturer / ADS</p>	<p>WP1: The ADS will be responsible for the DDT within its ODD and type approval will set out the technical requirements for. This includes compliance with road rules and observing all necessary infrastructure elements to perform the DDT (see all Capabilities (Section 3.3)</p> <p>This is covered in the Technical Requirements 1-5)</p>
---	--	--	---	---	---

Table 45 National Driving Standard Role 3 Summary

**Role 4: Drive safely and responsibly in the traffic system**

Reference		Performance standards You must be able to:	Knowledge and understanding requirements You must know and understand:	Applicable to	Comment
<b>Unit 4.1 Interact correctly with other road users</b>	Element 4.1.1: Communicate intentions to other road users	<ul style="list-style-type: none"> <li>- use indicators and arm signals to signal intentions correctly</li> <li>- support the use of any signals given by positioning the vehicle correctly and safely</li> <li>- use horn and lights to communicate with other road users where necessary</li> </ul>	<ul style="list-style-type: none"> <li>- the arm signals shown in The Highway Code and when they may need to be given</li> <li>- when and how to use indicators</li> <li>- why you should make sure signals are given in good time and cancelled as soon as possible</li> <li>- how to employ a safe and systematic way to make the best use of signals</li> <li>- when signals must be given and when it is acceptable not to use them</li> <li>- the law on the use of the horn</li> <li>- when the flashing of headlights may be used as a warning of approach or instead of the horn</li> <li>- the risks linked to incorrect use of headlights or the horn as a signal</li> <li>- how and when to use hazard warning lights</li> <li>- how and when to use road positioning to confirm your intentions</li> </ul>	fully applicable: manufacturer / ADS	<p>WP1: The ADS will be responsible for the DDT within its ODD and type approval will set out the technical requirements for. This includes compliance with road rules and observing all necessary infrastructure elements to perform the DDT (see all Capabilities (Section 3.3</p> <p>This is covered in the Technical</p>

	<p>Element 4.1.2: Co-operate with other road users</p>	<ul style="list-style-type: none"> <li>- be aware of and predict the likely actions of other road users</li> <li>- give other road users enough time and space to perform manoeuvres</li> <li>- monitor and manage your own reaction to other road users</li> <li>- respond to emergency vehicles correctly</li> <li>- make progress in the traffic stream and overtake with consideration for other road users</li> </ul>	<ul style="list-style-type: none"> <li>- how to scan the road ahead to gather useful information</li> <li>- the rules that apply to other road users, particularly drivers of large vehicles or vulnerable road users such as cyclists and motorcyclists, and the position that they may select on the road as a result</li> <li>- the importance of predicting the likely actions of other road users, especially vulnerable road users such as cyclists, motorcyclists, children and the elderly</li> <li>- the importance of always keeping a safe stopping distance between the vehicle and other road users</li> <li>- how traffic and weather conditions may affect other road users, such as by reducing visibility, and how to allow for this</li> <li>- how to act safely and responsibly when emergency vehicles are responding to incidents</li> <li>- how to make safe progress in the traffic stream</li> <li>- the rules that apply to overtaking on the left</li> <li>- that driving without due care and attention and reasonable consideration for other road users is an offence</li> </ul>	<p>fully applicable: manufacturer / ADS</p>	<p>Requirements 1, 2 and 10)</p>
--	--	--	---	---	----------------------------------

<p><b>Unit 4.2 Minimise Risk when driving</b></p>	<p>Element 4.2.1: Identify and respond to hazards</p>	<ul style="list-style-type: none"> <li>- continually scan the driving space close to the vehicle and into the distance</li> <li>- use visual clues to predict possible hazards and prepare for situations that may arise</li> <li>- judge the significance of possible hazards and prioritise your responses</li> <li>- respond to hazards safely</li> <li>- keep focused when faced with distractions</li> </ul>	<ul style="list-style-type: none"> <li>- methods you can use to scan your driving space effectively, both close to and into the distance</li> <li>- what can affect your field of vision, such as parked vehicles, and how to allow for this</li> <li>- how the construction of your vehicle may affect your field of vision, and how to overcome this</li> <li>- what aquaplaning is and when it might happen</li> <li>- factors that might cause you to skid, such as oil or gravel on the road</li> <li>- how to read the road ahead and prepare for the unexpected</li> <li>- which kinds of hazard to particularly look for in different environments, such as tractors on rural roads, deer on forest roads or flooding in heavy rain</li> <li>- that many tunnels are equipped with radio transmitters so that drivers can tune in to be warned of any incidents, congestion or roadworks</li> <li>- that if you come across congestion in a tunnel that causes you to stop you should leave at least a 5 metre gap between you and the vehicle in front</li> <li>- when other road users are vulnerable and how to allow for them</li> <li>- factors that can distract the driver (such as talking to passengers or using a sat nav system) and how to manage them so that you are aware of the driving space and possible hazards</li> <li>- the law on the use of mobile phones whilst driving</li> </ul>	<p>partially applicable: manufacturer / ADS</p>	<p>WP1: The ADS will be responsible for managing "critical scenarios, e.g., by executing Obstacle Avoidance Manoeuvres (see Technical Requirements Section 5.12 and "Avoid Obstacle behavioural competence" - Section 3.3).</p> <p>This is covered in the Technical Requirement 3 (and implied in 1 &amp; 2)</p>
---	---	---	---	---	--

	<p>Element 4.2.2: Drive defensively</p>	<ul style="list-style-type: none"> <li>- create and maintain a safe driving space</li> <li>- scan and check your surroundings, especially blind spots</li> <li>- position your vehicle to maximise visibility to other road users</li> <li>- use dipped headlights when necessary during daylight hours</li> <li>- manage your own physical and emotional state to make sure you can manage risks to your safety</li> <li>- drive at such a speed that you can always stop safely in the distance you can see to be clear</li> <li>- assess your own driving behaviour and identify areas needing work</li> </ul>	<ul style="list-style-type: none"> <li>- the importance of using a safe and systematic way to make sure you are always in control of your vehicle and travelling at the right speed, in the right gear and in the correct position on the road for the conditions</li> <li>- the importance of keeping a safe separation distance in all weather and traffic conditions</li> <li>- how to assess your own ability to drive safely and responsibly against best practice</li> </ul>	<p>partially applicable: manufacturer / ADS</p>	<p>WP1: The ADS will be responsible for managing system failures safely, e.g. by executing a MRM to achieve a MRC. (see Technical Requirements Section 5.12).</p> <p>There will be requirements on the operator to ensure processes are in place to handle emergency situations as part of their SMS.</p>
--	---	---	--	---	---

	<p>Element 4.2.3: Drive in an ecologically responsible (eco-safe) way</p>	<ul style="list-style-type: none"> <li>- accelerate and decelerate smoothly and progressively</li> <li>- foresee the need to stop, and use timely and smooth deceleration to reduce fuel - consumption and general vehicle wear and tear</li> <li>- drive in the highest responsive gear to keep full control and avoid labouring the engine</li> <li>- remove extra load from the vehicle when not needed</li> <li>- turn off the engine when you are likely to be stationary for some time</li> </ul>	<ul style="list-style-type: none"> <li>- what affects a vehicle's fuel consumption</li> <li>- how effective scanning and planning can help you to use smooth acceleration or deceleration to keep momentum</li> <li>- how fuel consumption is increased by:                             <ul style="list-style-type: none"> <li>&gt; extra load</li> <li>&gt; incorrectly inflated tyres</li> <li>&gt; wind resistance, for example from carrying luggage on roof racks</li> </ul> </li> <li>- that selecting the most suitable gear will avoid engine labour and maximise the effects of engine braking</li> <li>- the use of technologies to reduce exhaust pollution</li> <li>- under which circumstances it is appropriate to turn off the engine when stationary, rather than leave it idling</li> <li>- that you should never reduce safety to improve economy</li> </ul>	<p>partially applicable: manufacturer / ADS</p>	<p><i>This requirement is mainly formulated for ICE vehicles while the scope of the scheme is EV only. Some of the principles can be applied to driving to conserve EV range, but this is mainly an economical consideration for the operator.</i></p>
--	---	---	--	---	--

<p><b>Unit 4.3 Manage incidents effectively</b></p>	<p>Element 4.3.1: Take suitable action if your vehicle breaks down</p>	<ul style="list-style-type: none"> <li>- stop, in a safe place if possible, and switch off the engine</li> <li>- make sure passengers, animals and loads are managed safely</li> <li>- where suitable, give warning to other road users</li> <li>- seek appropriate help</li> </ul>	<ul style="list-style-type: none"> <li>- where possible, how to keep control of the vehicle if it breaks down</li> <li>- the law on using the hard-shoulder on motorways and the guidance on waiting for breakdown services</li> <li>- how to identify your precise location on motorways, to allow breakdown services to reach you quickly</li> <li>- that it is better to use an emergency roadside telephone than a mobile phone because it allows the operator to find your exact position</li> <li>- how and when to use a warning triangle</li> <li>- how and when to use hazard warning lights</li> </ul>	<p>applicable: manufacturer / ADS</p>	<p>WP1: The ADS will be responsible for managing system failures safely, e.g. by executing a MRM to achieve a MRC. (see Technical Requirements Section 5.12).</p> <p>There will be requirements on the operator to ensure processes are in place to handle emergency situations as part of their SMS.</p>
---	--	---	--	---------------------------------------	---



	<p>Element 4.3.2: Take suitable action if you're involved in or witness a collision</p>	<ul style="list-style-type: none"> <li>- where suitable, stop and park your vehicle in a safe place</li> <li>- make sure passengers, animals and loads are managed safely</li> <li>- make sure warning is given to other road users</li> <li>- assess the incident scene and your safety</li> <li>- note the condition of any casualties</li> <li>- give clear and accurate information to emergency services</li> <li>- give suitable help to others at the scene</li> <li>- where possible, record information about what you saw or the scene as you found it, including taking photographs and drawing sketch plans</li> <li>- comply with legal requirements accurately and in good time, if required</li> </ul>	<ul style="list-style-type: none"> <li>- the importance of making sure further injury and damage is not caused by:                             <ul style="list-style-type: none"> <li>&gt; managing uninjured passengers, animals and passers-by</li> <li>&gt; giving warning to other road users as quickly as possible</li> </ul> </li> <li>- how to contact the emergency services and the vital importance of giving them accurate information</li> <li>- the importance of being able to give information about the condition of casualties to the ambulance service</li> <li>- the benefits of gathering and recording information as soon as possible after the event</li> <li>- if you're involved in an incident that causes damage or injury to another person, vehicle, animal or property, you must know the laws that apply to:                             <ul style="list-style-type: none"> <li>&gt; stopping</li> <li>&gt; providing your details</li> <li>&gt; giving statements</li> <li>&gt; producing documents</li> </ul> </li> <li>- the principles of first aid and the limits of your own first aid skills</li> </ul>	<p>applicable: manufacturer / ADS</p>	<p>WP1: The ADS will be responsible for detecting and stopping in case a collision occurs. (see Technical Requirements Section 5.12 - requirement 5). There will be requirements on the operator to ensure processes are in place to handle emergency situations as part of their SMS.</p>
--	---	---	--	---	--

Table 46 National Driving Standard Role 4 Summary

**Role 5: Review and adjust driving behaviour over lifetime**

Reference		Performance standards You must be able to:	Knowledge and understanding requirements You must know and understand:	Applicable to	Comment
<b>Unit 5.1: Learn from experience</b>	Element 5.1.1 Learn from experience	<ul style="list-style-type: none"> <li>- demonstrate that you have continued to develop and update your driving skills since you took your driving test</li> <li>- recognise when your ability to drive safely and responsibly is affected by factors such as:                             <ul style="list-style-type: none"> <li>&gt; changes in your personal circumstances, such as changes in working patterns</li> <li>&gt; changes in your state of health and your physical abilities, through illness or age-related deterioration</li> <li>&gt; a break from driving</li> <li>&gt; changing to an unfamiliar vehicle</li> </ul> </li> <li>- assess the seriousness of the factors identified and:                             <ul style="list-style-type: none"> <li>&gt; change your driving behaviour to reduce the risks</li> <li>&gt; make plans for recovering or improving your driving ability</li> </ul> </li> <li>- seek professional help where needed</li> <li>- tell DVLA if you have a health or medical condition</li> </ul>	<ul style="list-style-type: none"> <li>- that you can learn from experience and continue to improve your ability to drive safely and responsibly all through your driving career</li> <li>- how to assess your own ability to drive safely and responsibly against best practice</li> <li>- how to assess and learn from others' driving behaviour</li> <li>- how to use feedback from others to help you be clear about your own ability to drive safely and responsibly</li> <li>- when to seek professional help</li> <li>- the advantages of having regular driver development sessions with a competent instructor to keep up to date and remove bad habits</li> <li>- the advantages of having an initial input from a competent instructor if you return to driving after a break or you change to an unfamiliar vehicle</li> </ul>	partially applicable: in-use regulator / operator (manufacturer)	WP1/ WP5: As the safe behaviour of the ADS is linked to the ODD it is important to ensure that the actual operating environment remains representative of the ODD. Monitoring for modifications that could negatively affect the ADS must be part of the operators or in-use regulator's responsibility. For the operator the monitoring and reporting of any incidences must be part of their SMS. The manufacturer

<p><b>Unit 5.2:</b> <b>Keep up to date with changes</b></p>	<p>Element 5.2.1: Keep up to date with changes</p>	<ul style="list-style-type: none"> <li>- demonstrate that your understanding of the meaning of road signs and markings is current</li> <li>- demonstrate that your understanding of the law on the use of a vehicle on public roads is current</li> <li>- keep up to date with changes to vehicle technology especially if you change the vehicle you are using</li> <li>- safely operate any technology that is fitted to any vehicle you drive including disabling it where appropriate</li> <li>- respond correctly to any changes in the documents required to use a vehicle on the road</li> <li>- take all steps needed to maintain your entitlement to a licence for the type of vehicle you are driving</li> </ul>	<ul style="list-style-type: none"> <li>- where to find information about changes to signs, markings and legislation, such as:                             <ul style="list-style-type: none"> <li>&gt; The Highway Code updates</li> <li>&gt; GOV.UK</li> <li>&gt; government publications</li> <li>&gt; motoring organisation websites</li> </ul> </li> <li>- where to find information about changes to vehicle technologies, for example:                             <ul style="list-style-type: none"> <li>&gt; manufacturers' websites</li> <li>&gt; trade magazines and websites</li> </ul> </li> <li>- where to find instructions on the safe operation of technology fitted to a vehicle</li> <li>- where to find information about changes to registration, MOT, or tax rules, such as:                             <ul style="list-style-type: none"> <li>&gt; GOV.UK</li> <li>&gt; government publications</li> <li>&gt; motoring organisation websites</li> </ul> </li> </ul>	<p>partially applicable: in-use regulator / operator manufacturer</p>	<p>needs to be part of any required response action (e.g. by providing updates to the ADS)</p>
---	--	--	---	---	--

*Table 47 National Driving Standard Role 5 Summary*

**Role 6: Demonstrate developed skills, knowledge and understanding**

Reference		Performance standards You must be able to:	Knowledge and understanding requirements You must know and understand:	Applicable to	Comment
<p><b>Unit .1 Demonstrate developed understanding of The Highway Code and the national standard for driving</b></p>	<p>Element 6.1.1: Demonstrate developed understanding of The Highway Code and roles 1 to 5 of the national standard</p>	<ul style="list-style-type: none"> <li>- the Highway Code and how to apply its rules</li> <li>- the national standard for driving and how to apply its elements</li> </ul>	<ul style="list-style-type: none"> <li>- the subject areas covered in The Highway Code</li> <li>- the rules set out in each section of The Highway Code where failure to comply is a criminal offence</li> <li>- the principles of the general guidance given within each section of The Highway Code</li> <li>- the subject areas covered in roles 1 to 5 of the national standard</li> <li>- the competences that:                             <ul style="list-style-type: none"> <li>&gt; each of the roles require</li> <li>&gt; a safe and responsible driver should be able to demonstrate</li> </ul> </li> </ul>	<p>ADSE/ manufacturer / operator</p>	<p>Compliance to the Highway Code is a key contributor to ensuring safe behaviour of an ADS but it is necessary to ensure that these rules, which are specifically defined for human drivers, are reviewed in the context of automated driving. This is in the scope of the WP2 activities. In WP1 the proposed Technical requirement 2 includes compliance with Highway Code rules where applicable (e.g., right of way rules to be part of DDT functionality).</p>

	<p>Element 6.1.2: Demonstrate developed driving competence</p>	<ul style="list-style-type: none"> <li>- make progress on the road safely and responsibly:               <ul style="list-style-type: none"> <li>&gt; while driving a variety of category B vehicles</li> <li>&gt; in urban and rural environments</li> <li>&gt; on any class of road</li> <li>&gt; at various times of day</li> <li>&gt; in differing lighting and weather conditions</li> </ul> </li> <li>- apply a systematic approach to driving to make sure that, at all times, you:               <ul style="list-style-type: none"> <li>&gt; are aware of what is happening on the road around you, including behind you</li> <li>&gt; respond to all road signs and markings correctly</li> <li>&gt; prioritise emerging and actual hazards and plan to deal with them effectively</li> <li>&gt; select and maintain a suitable position on the road</li> <li>&gt; travel at a suitable speed</li> <li>&gt; select the appropriate gear to be able to manoeuvre your vehicle, respond to hazards and minimise the environmental impact of your vehicle</li> </ul> </li> <li>- give a verbal commentary without reduction in driving competence, about:               <ul style="list-style-type: none"> <li>&gt; what you see on the road around you</li> <li>&gt; what you understand about the principles of eco-driving</li> <li>&gt; how you adjust your driving in response to what you have seen and what you understand</li> </ul> </li> <li>- maintain appropriate attitudes</li> </ul>	<ul style="list-style-type: none"> <li>- the main differences in set up and technology that are found in category B vehicles</li> <li>- how to modify your driving to take into account the differing conditions and hazards that you are likely to find:               <ul style="list-style-type: none"> <li>&gt; on different classes of road</li> <li>&gt; in different levels of traffic</li> <li>&gt; at different times of the day and night</li> <li>&gt; in differing lighting and weather conditions</li> </ul> </li> <li>- how to apply a systematic approach to driving, such as Mirrors, Signal, Manoeuvre – Position, Speed, Look (MSM-PSL)</li> <li>- how to balance and combine the demands of safe driving and the principles of eco-responsible driving</li> <li>- how to give a simple talk while you're driving that allows an observer to understand how you:               <ul style="list-style-type: none"> <li>&gt; scan your environment</li> <li>&gt; plan how to make progress</li> <li>&gt; make sure you're safe</li> <li>&gt; minimise the environmental impact of your vehicle</li> </ul> </li> <li>- how to deal safely and effectively with any negative attitudes and emotions that might be triggered while driving, such as anger at other drivers who you believe are driving dangerously</li> <li>- the main differences in set up and technology that are found in category B vehicles</li> <li>- how to modify your driving to take into account the differing conditions and hazards that you are likely to find:               <ul style="list-style-type: none"> <li>&gt; on different classes of road</li> <li>&gt; in different levels of traffic</li> <li>&gt; at different times of the day and night</li> <li>&gt; in differing lighting and weather conditions</li> </ul> </li> <li>- how to apply a systematic approach to driving, such as Mirrors, Signal, Manoeuvre – Position, Speed, Look (MSM-PSL)</li> <li>- how to balance and combine the demands of safe driving and the principles of eco-responsible driving</li> <li>- how to give a simple talk while you're driving</li> </ul>	<p>partially applicable: ADSE/ manufacturer / operator</p>	<p>Scope of WP5 is to set out what the in-use monitoring must monitor in order to determine the continued safety of the AV over time. There will be requirements on the operator to ensure that these monitoring activities are implemented and executed as required - their SMS will require those processes to be set up.</p>
--	--	---	--	--	---

		to all other road users at all times	<p>that allows an observer to understand how you:</p> <ul style="list-style-type: none"> <li>&gt; scan your environment</li> <li>&gt; plan how to make progress</li> <li>&gt; make sure you're safe</li> <li>&gt; minimise the environmental impact of your vehicle</li> </ul> <p>- how to deal safely and effectively with any negative attitudes and emotions that might be triggered while driving, such as anger at other drivers who you believe are driving dangerously</p>		
--	--	--------------------------------------	---	--	--

Table 48 National Driving Standard Role 6 Summary

## 10.6 Appendix 6: Safety Management Systems

This is an appendix to Section 7.1 of this report, and is broken down into subsections (A to D), with some of these being further decomposed.

### 10.6.1 Appendix 6 – A: Background to SMS Structure and Processes

An SMS does not have a universally-defined structure. There are common factors and key processes or requirements that are critical for an AV SMS to include. However, each organisations SMS will be unique, as it must be designed specifically for use within that organisation. There will also be differences determined by which industry the organisation is from (e.g. rail, road, aviation), but they must all be comprehensive and detailed. An SMS can be used for operational and systems safety but should still align with the overall management of safety, security and environment within the organisation. The AVSC (2021) defines 4 key elements to an effective SMS:

- **“Safety Policy and Objectives (SPO):** Establish or enhance safety practices with a clear safety policy, safety roles and responsibilities, and organisational safety objectives.
- **Safety Risk Management (SRM):** Proactively manage risk using safety risk assessments.
- **Safety Assurance (SA):** Monitor, analyse, and measure overall safety performance, including effectiveness of its safety risk controls, safety management, and associated processes.
- **Safety Promotion (SP):** Regularly conduct activities that inform, educate, and heighten the safety awareness of employees.”

This structure is given in specific reference to SAE Level 4/5 AVs. The naming of sections is not critical, but the inclusion of the key elements is.

The structure above must be conceptualised into a process or cycle through which an SMS is continually assessed, updated, and improved. The British Standards Institution (BSI, 2018) gives guidance and requirements for operational health and safety systems, with Figure 46 representing an example of a simplified but effective process for SMS maintenance.

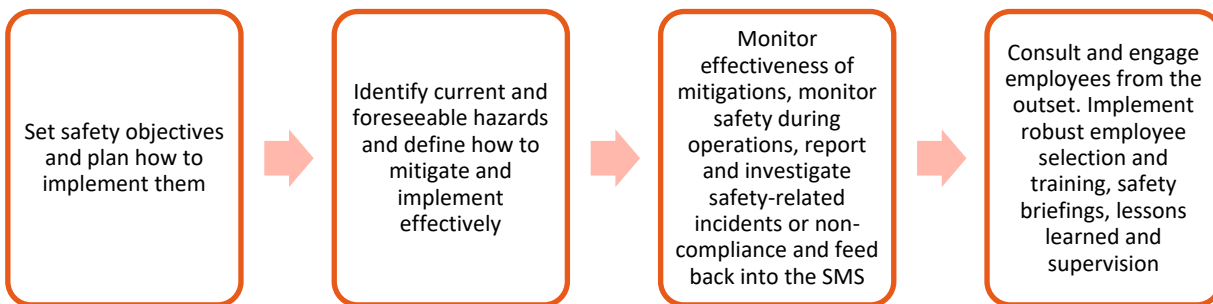


Figure 46: Example of a simplified SMS process. Source: (BSI ISO 45001:2018)

If an organisation is following this basic structure, it is already making progress on producing a robust SMS and maintaining an effective safety culture. This should not be a one-off event, but a continual process of “Plan, Do, Check, Act” (BS ISO 45001:2018). More detail on continual monitoring and improvement is given in Appendix 6 - D.

The University of York’s Vehicle Regulations Objective (2021) give further detail on the information to be documented and reviewed when maintaining and updating an SMS. These give specific examples of processes and procedures that are required to be followed to be maintained effectively. The SMS should cover the entire safety process of commercial AV deployment, demonstrated by the list below:

- “a) A process for safety assessment of design, verification and change relating to the vehicle, covering software, hardware, subsystems and data.

- *b) Procedures and mechanisms for responding to test failures, incidents, accidents, and hazardous failures. [Appendix 6 - D1]*
- *c) Processes, procedures, competencies, certifications and training for vehicle design, manufacture, maintenance, and upgrade activities. [Appendix 6 - C]*
- *d) Processes for responding to directives from regulators, including making design changes and communicating to users/operators of the vehicles.*
- *e) Processes for updating the safety documentation to allow for regular review and re-issue as appropriate". Appendix 6 - D]*

To expand upon list item *b)*, this further requires identifying the test failures, incidents, accidents, and hazards through appropriate processes. This is also the case for traffic infractions, behavioural competency issues and violations of the safety case. If a review of an organisation's safety processes finds processes and procedures like this in assessing and updating safety objectives, then this is good initial evidence of an effective safety culture, especially in the constantly evolving context of AVs. However, leadership, management, consultation, and involvement are also critical, so all employees feel ownership of safety. It is recommended that the basic structure outlined above is followed for an organisation's SMS development, and is adapted for their specific needs and processes.

## 10.6.2 Appendix 6 – B: Creation and Maintenance of an SMS

An organisation may choose to produce a defined SMS if it had not previously established one, or it may already have an SMS in place but wish to update it to a new and more appropriate structure. It is recommended that this process begins with 'mapping' the organisation and their processes as they relate to AV safety. This can involve the current safety/risk responsibilities across organisational roles and how they feed into the current safety reporting structure. These can then be applied to an SMS structure that provides for the safety needs of the AVs. This will also identify the safety practices that are effectively implemented already, and possibly employees who take a key role in promoting safety culture, to produce the 'building blocks' from which to develop a better SMS. The scope of an SMS should be defined before, and adapted during, formulation. This needs to consider the type of activities conducted by the organisation and how complex they are, and will determine the proportionate level of detail needed when defining safety processes while the SMS is being mapped and created.

The importance of mapping structure and key safety roles (Appendix 6 - B2) is even more pertinent in larger organisations with more diversified safety processes; however it is always evidence for a robust SMS. This mapping of the organisation and processes can equate to a gap analysis of safety needs, as opportunities for improvement are likely to be discovered when analysing the entire safety reporting/maintenance structure. Identification and improvement of these can immediately be set as 'safety objectives' or targets for the new SMS to achieve.

### 10.6.2.1 Appendix 6 - B1: Overall SMS Documentation Approach

When an organisation is creating or 'mapping' an SMS, it is important that they have an awareness and understanding of current good practice, and have consulted the existing internal documentation to get an accurate representation of the organisation's current safety practices. Knowledge management is key in such a rapidly changing industry like AV research and deployment. As a result, documentation needs to be able to be rapidly updated and information shared to maintain alignment with good practice and new standards, and to ensure safety. Documentation within organisations is decided by many factors, including SAE Level (Level 4+ AVs in particular), size of the organisation, organisational role within the AV deployment and current documentation processes of the organisation. The size of the organisation may affect the documentation processes of the SMS but should not affect how thorough or detailed the SMS is; the organisation should still practice effective safety oversight, regardless of size. Therefore, it is up to the organisation to decide what of their current or future documentation should be included within the creation of and updates to the SMS through consulting and complying with regulations and requirements. However, this documentation must follow the recommendations and guidance.

It is important that the documentation reflects the commercial deployments the organisation is currently involved in, and that employees understand how to apply the safety processes to their work within these deployments. It



is also important to reflect current good practice, emerging good practice and learning from other, comparable industries. Where data exists (such as for near misses, incidents), this should be fed into improved, safer practices within the SMS. Documentation needs to be concise, unambiguous, periodically reviewed, updated to reflect lessons learned, version controlled and part of change management process. Documentation should be written in plain language and understandable by people using them. People also learn in different ways, so a combination of diagrams and written English guidance is preferable. The University of York's Vehicle Regulations Objective (2021) recommends documentation or information to be included in an SMS including:

- "a) A process for safety assessment of design, verification and change relating to the vehicle, covering software, hardware, subsystems and data.*
- b) Procedures and mechanisms for responding to test failures, incidents, accidents, and hazardous failures.*
- c) Processes, procedures, competencies, certifications and training for vehicle design, manufacture, maintenance, and upgrade activities.*
- d) Processes for responding to directives from regulators, including making design changes and communicating to users/operators of the vehicles.*
- e) Processes for updating the safety documentation to allow for regular review and re-issue as appropriate".*

To expand upon list item *b)* above, this also requires identifying the test failures, incidents, accidents and hazards through appropriate processes. This is also the case for traffic infractions and violations of the safety case.

An SMS needs to cover the entire safety process within the organisation, and therefore requires organisation-wide input into the safety recommendations and processes; this also includes experience from employees (Appendix 6 - D4). An SMS should consider system design, configuration, integration and operation, as well as all relevant information that should be used across policy, risk assessment, organisation, role setting and updating of the SMS. This requires a thorough approach, and a positive safety culture would assist in maintaining this thoroughness. Evidence of reviewing a variety of forms of safety documentation is therefore a requirement when assessing the effectiveness of an SMS.

### 10.6.2.2 Appendix 6 - B2: Roles, Responsibilities and Key Safety Personnel

As stated previously, an organisation should not only evidence the presence of an effective SMS, but also the application and use of it across the organisation and employees. An important part of this is "*communicating the safety responsibilities of the organization's personnel and ensuring they have the necessary competencies to perform duties relevant to the operation and performance of the SMS*" (FAA, 2018). This means ensuring that the key safety roles within the organisation are occupied by those with appropriate expertise and that other employees are also informed of, understand and are trained to perform their responsibilities (this education/ awareness need is expanded upon in Appendix 6 - C). Communication with other organisations with safety responsibilities for deployment is also key; for example, this could be a component supplier and/ or a licenced fleet operator. Similarly, communication around safety responsibilities between those organisations, through training, user manuals etc., is also key to ensuring that employees are educated and that the information in these manuals is appropriate to the current iteration of the SMS.

It is recommended that key safety personnel are established and take responsibility for the maintenance and continued updating/ awareness of the SMS within the organisation. These will also be the personnel who promote a positive safety culture by ensuring participation in and adherence to appropriate safety policies. Although the entire organisation should comply with the SMS, it is these key safety personnel who should ensure it is properly updated, disseminated, and complies with relevant standards while following best practice. It is also the responsibility of senior management and executives to be informed and responsible for updating and promoting SMS effectiveness, rather than sole ownership belonging to the key safety personnel. This overall ownership will promote a strong safety culture, because all employees should be aware of their responsibility to conduct their work safely. Those investigating the SMS and safety culture of an organisation during approval can use the actions and involvement of these key personnel to assess how well the organisation carries out safety operations.

In the context of LSAVs, the key safety personnel and senior management should have a good operational knowledge of commercial AV safety policy, standards and best practice, because this ensures the recommendations and policies promoted are appropriate and will have a tangible impact on the organisation's safety. This could involve information such as knowledge of the Operational Design Domain (ODD) and Target Operating Domain (TOD) of any AVs currently being deployed. Essentially, the safety personnel should be

nominated from technically proficient and involved parts of the workforce so that the oversight of the SMS and safety culture is well informed.

These key safety personnel, and how any safety incidents (Appendix 6 - D1) are discussed, should also evidence consideration of the Law Commissions' 'Duty of Candour' proposed within their Automated Vehicles Joint Report (Law Commissions,2022). Here, the 'nominated person' is responsible for signing off safety cases and reports, analogous to being key safety personnel. The duty of candour states that safety incidents leading to injury or other serious issues are accurately reported by the nominated persons, and that senior managers and relevant organisations are notified, for example when licencing vehicles. Though a "no-blame safety culture" is suggested, if the nominated person does not *"take active steps to ensure that the information submitted to the regulator is correct and complete"* then they are liable to be charged criminally should the lack of candour lead to a punishable offence. Organisations should demonstrate that their SMS can fulfil this Duty of Candour at all levels of the organisation and ensure it is followed by all employees and especially by the nominated safety personnel; it should also be embedded into their SMS. This ensures liability and responsibility related to the SMS is maintained, especially in relation to AV incidents that cause serious injury or death.

Though responsibility and criminal liability should be considered, the Law Commissions' report also encourages a move away from blaming human error where possible. It should instead focus on identifying and addressing software/ technical errors within AV operation. A relevant quote from the report is shown below:

*"[We] proposed a move away from the current emphasis on the criminal prosecution of human drivers. Instead, we proposed that the in-use safety assurance scheme should investigate breaches of traffic rules by AVs driving themselves and apply a flexible range of regulatory sanctions on ASDEs...We do not think that an individual should be penalised for a breach that was brought about by the ADS. "*

This means that organisations should continually assess the safety of the SMS, mainly through the operation of the ADS and its functionality/ software. This will utilise data from the safety and incident reporting processes that address incidents through monitoring incident data and software performance, rather than focussing on human error. The Law Commissions suggest sanctions be placed on ASDEs (broadly equivalent to 'manufacturer' within this report) in the event of inadequate safety oversight, because these will better promote effective in-use safety assurance, with monetary fines - called "civil penalties" in the report - not always being appropriate to apply to organisations as a whole. They are defined as part of the set of regulatory sanctions but are separate in their use. Rather than place most of the blame immediately on nominated safety personnel, these sanctions will address errors on a wider, functional and regulatory level. There are situations defined where individual penalisation is appropriate, such as cases of gross negligence leading to safety incidents involving AVs, and in these cases the Law Commissions do suggest investigating criminal liability. However, for cases of errors within the safety processes set by an organisation, placing of wider sanctions is more appropriate. It is recommended that organisations work to follow this advice to maintain a modern SMS in line with current thoughts on SAE Level 4 AV regulations and ensure that the updating and reporting of safety incidents is done through appropriate processes.

These needs and responsibilities of key safety personnel also show that it is not only their appointment that needs to be evidenced, but also their operation and actions within the organisation. An organisations' management are ultimately responsible for safety culture and naming a "safety lead" or similar role. They will play an important role in supporting the organisation in the upkeep of an SMS, and feeding back to the organisation's management, but should not be solely responsible for the ownership of the SMS.

### 10.6.2.3 Appendix 6 - B3: Tailoring an SMS

An SMS should be bespoke and specific to the organisation in question, and should be applicable to their safety needs and processes, together with the inherent complexity of the technology, behavioural competencies and operating environment of the vehicle concerned. Typically, the more complex these factors are, then the more comprehensive and detailed safety oversight is needed, especially relating to software and data monitoring of the AVs in-use. Safety oversight of all AVs should be detailed and comprehensive, but more complex software and data monitoring will affect the safety oversight required.

The bespoke nature of an SMS can be evidenced through several different indicators, as described in the following sub-sections.

### 10.6.2.3.1 Appendix 6 - B3.1: Company Lexicon

A “company lexicon” is the naming conventions and keywords unique to an organisation and how they structure their business. For example, there may be specific job titles within the hierarchy that do not correspond to those outside the organisation. There may be document types or names produced internally that must be named using specific authoring standards, or teams may be organised in a functional system that is unique to the organisation.

An SMS should take this into account when it is created and used, otherwise there are likely to be errors associated with the dissemination of the safety practices and documentation. If an SMS uses a generic structure with generic titles and processes, it will not be as effective. The mapping of previous safety processes onto a novel SMS in Appendix 6 - B assists with this. An SMS should be able to document the specific processes and roles associated with safety in the organisation and not utilise generic statements or non-specific roles. An example of an ineffective statement or policy is:

“It is the responsibility of technicians working on automated vehicles to report when a safety incident occurs”.

Which technicians are reporting this? Who are they reporting to? What documents or channels need to be used to report incidents? These need to be detailed and specific to the company lexicon within the SMS to ensure proper use and integration into safety culture. This is an indicator that is easily assessed, because by reviewing the SMS and other organisation documentation, there should be a clear overlap or correspondence. If there is not, the SMS should be updated to match the lexicon used across the organisation and the specific hierarchy and roles.

### 10.6.2.3.2 Appendix 6 - B3.2: Safety Objectives

Organisations should have a well-defined and specific set of safety objectives (SOs). These are relatively high-level goals or aims that the organisation is working towards, and should represent the most pertinent focus or issues of safety work currently taking place within the organisation. These safety goals will be set in collaboration between senior management and key safety personnel. They should be consistent with the work taking place and the scenario in which progress towards any goal can be assessed. As with the entire SMS, they should be bespoke to an organisation. An example of an ineffective or ill-defined SO could be:

“Increase operational safety of automated vehicles”.

Here there is little detail as to how the goal is to be achieved and what the specific outcomes are. A well-defined example will be closer to something like:

“Ascertain corrections to be made to operational design domains of urban driverless pods through incident reports”.

This has a clear objective (correct ODD) and method (reviews of incident reports). The SMS for Small Organisations (SMICG, 2015) suggests that, alongside safety performance indicators below, SOs are the overall goal or end aim, whereas the SPIs are how progress towards these goals are to be addressed. SPIs must be set using the “SMART” system of “Specific, Measurable, Achievable, Relevant and Timed”. The objective need not contain these details, as it is a high-level goal. SPI’s can be seen as the SMART element of SO’s.

The SOs should be relatively high-level but should involve deeper research into the safety needs of the organisation and should be updated according to changes. They should match closely the projects and current work the organisation is engaged in so as to be as accurate and useful as possible. BS ISO 45001:2018 states that *“the OH&S policy and related OH&S objectives are established and are compatible with the strategic direction of the organisation [and] ensuring the integration of the OH&S management system requirements into the organisation’s business processes”*. By reviewing these SOs in reference to the aims and operations of the organisation, a strong safety culture can be demonstrated, especially through the dissemination and awareness of these objectives and how they apply across the workforce (Appendix 6 - C). If employees are aware of the SOs, and the objectives match the current AV deployments, then this provides good evidence of a robust safety culture.

### 10.6.2.3.3 Appendix 6 - B3.3: Safety Performance Indicators

In Appendix 6 - B, one of the recommended operations within an SMS is to *“Monitor, analyse, and measure overall safety performance”* (AVSC, 2021). This goes together with SOs to measure and track overall progress of the organisation towards the safety objectives. In an SMS, SOs are made “measurable” by safety performance indicators (SPIs). These are specific measures of the performance of safety processes or documentation, and

can be key indicators of both the effectiveness of an SMS and the integration of a safety culture in an organisation. SPIs should be traceable to the current SOs and specifically measure how close the organisation is to achieving these goals.

SPIs should be more detailed than the SOs since they should accurately measure the “safety performance” they reference. They should be formulated using “SMART” principles, as they are what make an SO able to be assessed. The same recommendations apply as were described for SOs with regard to being specific and bespoke to the organisation’s current processes, because a generic set of SPIs may not be able to be accurately measured against when using more complex data.

A generic and ill-defined example of an SPI is:

“Number of safety meetings”.

Measuring the number of safety meetings may give an indication as to how often safety is being discussed, but there is no reference in the SPI about whether more/less are being aimed for and why. More specific examples are as follows:

“Reduction in the number of vehicle safety reports relating to issues with the vehicle X’s operation in project Y” or “Ascertain corrections to be made to operational design domains of urban driverless pods through post-trial reviews of incident reports. Measure by documenting post-trial reviews and recording lessons learned. Assess changes made by end of Q3”.

These are examples of SPIs that explain what is being measured and what to measure against. They also state the AV deployment/work focus they are related to and should be able to assess progress towards the appropriate SOs.

SPIs should also be set in reference to the type of data analysis or in-use analysis being used by the organisation. It is likely that as the complexity of the ADS and its operating environment increases, the complexity of data and therefore SPI goals will increase proportionately, and develop and change more frequently.

Looking at the crossover between SMS and safety culture, as is seen with the setting of SOs, evidence of a positive safety culture within an organisation includes a focus on continually assessing and considering SPIs in safety and commercial AV work. If an organisation is continually working to improve performance against an SPI and employees are aware of which SPIs relate to them (Appendix 6 - C), then their culture will very likely promote strong safety assurance.

#### 10.6.2.3.4 Appendix 6 - B3.4: Differentiation Between Collaborating Partners

When organisations are working together on projects involving commercial AVs, it is important that they communicate and collaborate on safety policy and process throughout to maintain a safe working environment. This promotes a good safety culture across the partners, since they are taking tangible steps to improve safety not only for their own workers but those of collaborating organisations. By sharing learning experiences from incidents, the entire industry can improve future safety responses and monitoring through continuous feedback. It is also important for organisations such as manufacturers and fleet operators to share safety responsibilities including through training, user manuals and safety cases informed by each organisation’s own safety culture. However, partners having an SMSs for these processes which are too closely aligned could result in gaps in internal safety governance, since they could involve documentation and safety concerns that are not relevant to their own organisation. When working together, the important and relevant parts of the various SMSs should fall in line to ensure consistent application, but they should not be changed to be indistinguishable from each other.

For example, a commercial LSAV manufacturer and operator should be aware of each other’s safety processes to allow for a good understanding of how to safely operate an LSAV once it is handed over. There should be collaboration through sharing and compliance with safety policies especially relating to in-use LSAV operation. However, even if they are regular collaborators, the organisations should have distinct SMSs. The policies of a commercial LSAV manufacturer will not sufficiently cover the other operations of an operator, and vice-versa. It is recommended that a shared safety framework is used when collaborating, but the individual SMSs should not be required to achieve perfect alignment with each other. This differentiation should be assessed when examining for effective SMS and safety culture, since simply replicating the SMS of another organisation shows a lack of concern towards having a well-defined and appropriate safety policy in an organisation.

### 10.6.3 Appendix 6 – C: Education and Awareness of an SMS

Information in this section can be read alongside Appendix 6 - B2, since the roles and responsibilities of key safety personnel have already been discussed. These roles and responsibilities extend across an organisation, because a key to a positive safety culture is employee involvement in the SMS and awareness of the SMS. An important part of a fully integrated SMS is “communicating the safety responsibilities of the organisation’s personnel and ensuring they have the necessary competencies to perform duties relevant to the operation and performance of the SMS” (FAA, 2018). This means that the use of the SMS and information about it needs to be *appropriately* shared across an organisation, and employee safety competency should be assessed. Commercial AV organisations employ a variety of roles including trial managers, maintenance workers and test/software engineers, and so provision of SMS information and training needs to be diversified to be effective.

This does not mean that all employees/workers must be shared the entirety of the SMS and be expected to understand all of the processes, reporting and rationale behind it; this is the role of “key safety personnel” that has been previously explored. Instead, they should be given the information and responsibilities relevant to their expertise and role, such that this information applies to them specifically. This can involve training or workshops on how to relate and apply their work to the SMS, or how the SMS will influence their work. This dissemination of information ensures that the SMS applies across the organisation and that its processes are followed, rather than being known by key safety personnel alone. The training put forward by key safety personnel must consider all levels of the organisation, not just the production of an overall SMS framework. The SOs within the SMS should be deliverable and visible across the organisation, to be able to properly define a robust and effective safety culture.

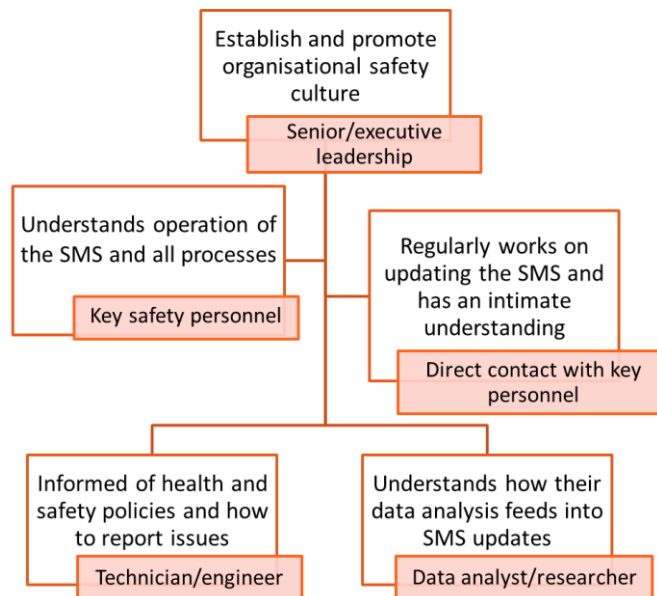


Figure 47: Examples of employee SMS education needed.

Since the way in which education and awareness training “must consider all levels of the organisation”, this includes tailored training for all members of staff from key safety personnel and senior leaders to hands-on technicians and new starters. Therefore, this includes “the design of work areas, processes, installations, machinery/equipment, operating procedures and work organization” (BS ISO 45001:2018) and operations both of employees in the organisation and of other external workers or partners who are also in attendance at the workplace. An organisation assessing an SMS should check that it refers to all appropriate levels of the organisation, and its SOs and SPIs may also refer to improving safety in these areas (Appendices 6 - B3.2 and 6 - B3.3). It should also be clear that employees are knowledgeable on their role within the SMS, how their work is important and any safety goals that they should be aware of. They should also know the channels through which they can report any safety incidents or improvements to safety policy they encounter (Appendix 6 - C and Appendix 6 - D). Figure 47 summarises the differing needs of employees with regards to SMS training.

Overall, the priority of awareness and education regarding an SMS and safety culture is ensuring that employees have the appropriate knowledge, skills, and motivation to apply their tasks to the organisational SMS and help

maintain the core values of the safety culture. It is recommended that organisations provide regular training and education opportunities that incorporate the appropriate level of detail for the projects they are a part of. For example, a data analyst examining the performance output of projects should be better informed about how to address SPIs relative to the data they are reviewing, whereas an LSAV maintenance worker should be trained on how to prevent, monitor, and report safety incidents concerning AVs and physical injury. Having appropriate training and education provision is clear evidence of a positive safety culture.

The differences in tailored training can also be affected by the size of an organisation. A small or very small organisation (SMICG, 2015) will likely have less diversified needs when disseminating information about the SMS. It is up to the organisation's key safety personnel to assess the needs of individuals in these organisations, to determine whether any information they share with members is irrelevant and could disrupt their work towards safety goals in their areas of focus. For larger organisations (25+ employees), the SMS, key safety personnel and employees should demonstrate a diversified awareness and provision of SMS information, so they can evidence that they are effectively integrating this into their safety practices.

## 10.6.4 Appendix 6 – D: Continuous Monitoring and Improvement

An SMS is not a static, single-use document; it is a dynamic set of policies, goals, processes, and measures that evolves alongside the work an organisation engages in. Therefore, there is a need for evidence of continual monitoring, assessment and updating of the documentation, to maintain an effective SMS and demonstrate effective safety culture. This is true across multiple industries that use/ require an SMS, but this continual updating is particularly important for commercial AVs due to the innovative and developing nature of the technology and industry. This does not always mean more detail, as an SMS should not be overly complex or difficult to navigate, but instead may mean more frequent updates to reflect the changing nature of commercial AV technology, use cases and assurance methods. BS ISO 45001 (2018) gives the 4 processes that are key to an SMS as “Plan, Do, Check, Act”, with “Check, Act” explained as “*monitor and measure activities and processes with regard to the OH&S policy and OH&S objectives and report the results [then] take actions to continually improve the OH&S performance to achieve the intended outcomes*”. It is also important that these changes are both reactive (incident reporting) and proactive (employee consultation, operational planning).

Safety audits around the entire SMS and specific safety cases can also be conducted. These audits should be conducted by key safety personnel and will identify whether key mitigations are being implemented in response to specific safety case documentation, or examine the overall function of the SMS within the organisation, rather than merely auditing the content within the SMS itself. These audits must consider the entire function of the organisation, including formulation (Appendix 6 - B), documentation and updates (Appendix 6 - D), performance (Appendix 6 - D3) and employee integration (Appendix 6 - C).

These updates to the SMS should be present for all parts, not just for goals alone. Every organisational SMS should be unique, and this is facilitated by continual update and monitoring, and each organisation will change and differ through its own work, even if this is collaborative work with other partners.

Below are examples regarding when updates need to take place, or possible updates that could be made, depending on the event preceding or requiring the attention of key safety personnel who update the SMS outside an overall safety audit.

### 10.6.4.1 Appendix 6 - D1: Incident Monitoring and ODDs/TODs

Incident and data monitoring are both extremely important in maintaining a strong SMS. If these are not monitored correctly, then it is difficult to create an accurate picture of the safety performance of an organisation. An “incident” could be both a safety incident to do with a vehicle (e.g., injury, near miss or non-compliance with standards) but also a “trigger” event that leads to the requirement to review and update the SMS. Triggers can include a change in management, moving to a new location or an organisation restructure. All of these will mean updates are needed for the SMS to be suitable for the organisation; for example, a change in the nominated key safety personnel, or how the different roles after a restructure take responsibility for achieving safety goals, would need to be reflected within the SMS.

A phrase important to define is “organisational direction”. This is the focus of the work the organisation is currently involved in, which can change over time and thereby require changes to the SMS, but a good safety culture should be resilient to these changes. For example, a change in organisational direction would be a commercial AV organisation shifting focus from automated passenger vehicles in urban areas to off-highway material transport vehicles. This change in direction can precipitate change like the other “triggers” defined above.

For commercial AV organisations, it is important for them to establish from their data methods what constitutes an incident and how these are checked for (and later reported, Appendix 6 - D1). A significant part of this comes from their data analysis and what data is being used. For example, if they are studying speed data, route compliance data or journey time data, then there must be methods in place to identify when an incident has occurred, explain why it occurred and whether there are any takeaways or learnings from the data around the incident. There should be defined processes, or at least reference to these processes/ analyses in the SMS documentation, so that when an incident has occurred, employees are able to look at the data and determine where the incident stemmed from.

If there are defined processes, then these need to be updated within the SMS in line with the types of data and analysis being used. For example, if there is reference to a process of using positional/ object proximity data to analyse a safety incident in the SMS, but this type of data is no longer being used by the organisation, then this process for incident monitoring has no use within current safety work. These incident monitoring methods should be kept up to date with the current focus of the organisation, demonstrating how the use of innovative data analysis in the commercial AV industry presents the need for more frequent updating than in other industries. This is also important when organisations work with project partners and may adopt aggregated datasets or share large amounts of information/ processes. Though they should not entirely replicate each other's SMS (Appendix 6 - B3.4), it evidences a good safety culture when they can immediately discuss incident monitoring using these aggregated datasets to ensure that there are as few gaps in oversight as possible.

Safety of software and effectiveness of in-use monitoring must also be a dedicated consideration within appropriate organisations who operate commercial AVs, since during operation is when the most significant safety issues are likely to occur. This could include safety issues resulting in physical injury or worse. Within an SMS, reference should be made to this in-use monitoring and how the data is used and documented, since through this, appropriate safety measures and mitigations can be defined by analysing frequency of incidents and the most frequent types of incidents. The performance of this software should be audited in relation to safety incidents and also regularly during overall safety audits (Appendix 6 - D).

ODDs and TODs (see Section 4.1 of this report) also require consideration within an SMS. By having an SMS that is frequently updated to be in-line with organisational direction and current data analysis, it can be ensured that the COD experienced by the vehicle in service remains compatible with the TOD defined for the deployment, and hence with the ODD, with any CODs experienced by the system that lie outside the TOD triggering an investigation, and potentially an update to documentation. This is important for incident monitoring, because a well-defined TOD means that TOD exit can also be well-defined. A TOD exit occurs when an ADS operates outside of its defined conditions and/or deployment requirements. This can include running at a speed over its defined limit within an area, travelling onto routes where it has not been approved for use, encountering situations that were not previously anticipated to be possibilities ('black swan events'). These could cause serious safety incidents, such as collisions.

An SMS needs to include appropriate processes that are both proactive and reactive in assisting the prevention of AVs exiting TODs. This includes processes before deployment, such as reviewing software performance, route eligibility and vehicle/hardware condition, to ensure that the LSAV can run within the requirements of the TOD. This can also include evidence from previous trials, including telematics, dash cam footage, human operator report or any other form of appropriate information recording that the organisation uses. Through this being documented, more appropriate and well-controlled ODDs and TODs can be established, since any mitigations or solved issues used in these circumstances can be proactively implemented. Reactively, reviews and safety audits of the conditions leading up to a TOD exit can be used to determine causes, and therefore future mitigations to employ when operating a commercial HAV/ADS.

It is also appropriate to produce emergency response plans (ERPs) for use in the event of certain incidents during deployment. These should be formulated in collaboration between manufacturers, operators and other stakeholders. The ERPs define the steps the organisation should take in event of an incident, including contacting emergency services, contacting senior management/ executives and protection of other employees/ people around the incident. These ERPs should be updated with lessons learned from previous incidents in order to be specific and effective.

A positive safety culture in this instance should be built around an awareness of results and learning opportunities from previous incidents and monitoring, and use these learning opportunities to update and educate others on how to apply the SMS to ADS work. This includes identifying and documenting these learning opportunities, and how they were implemented, through change management and changes to the SMS. This can utilise a 'lessons learned' summary post-incident, and also the maintenance of an ongoing, live record of updates and underlying

rationale. Organisations should also communicate these updates to relevant parties in relation to the safety processes they affect, including working environment and software updates that are required/ recommended.

### 10.6.4.2 Appendix 6 - D2: Safety Risk Assessments and Safety Reporting

A further “proactive” measure that can be used to update an SMS is safety risk assessments (SRAs). These are formal assessments used to examine and evaluate the safety risks of certain defined projects, implementations, and operations. In the case of commercial AVs, they can be used to identify the significant hazards and risks that may occur in the next phase of deployment so that mitigations and safety controls can be applied. Risk assessments can lead to rejection or suspension of deployment progress if the identified hazards are intolerable. They can also be used to assess “trigger” events, defined in Appendix 6 - B as a change in management or company restructure that may present hazards to employees that need to be assessed. Operational risk assessments for LSAV deployments were considered within Section 6.1 of this report.

A strong SMS can be evidenced by SRAs, as this enables the organisation to address safety concerns proactively and means that the SMS can be updated through the results of these assessments. If an SMS did not employ the use of frequent SRAs, then it would not be adequately checking the safety of the operations that the organisation is taking part in. This also highlights evidence of a good safety culture, since key safety leadership is about ensuring continued safety of the organisation by initiating these assessments and helping to mitigate against safety incidents both before and after they occur. SRAs are a key tool in assessing whether an organisation is taking safety seriously, and form strong evidence of effective implementation of an SMS.

SRAs are also required to be conducted by law. The Health and Safety Executive (HSE) requires that organisations protect their employees from harm and protect others who may be harmed through the actions of the organisation. The Management of Health and Safety at Work Regulations (1999) states:

- “Every employer shall make a suitable and sufficient assessment of—
  - (a) the risks to the health and safety of his employees to which they are exposed whilst they are at work; and
  - (b) the risks to the health and safety of persons not in his employment arising out of or in connection with the conduct by him of his undertaking”.

The Act states that the assessment should define the hazards, the risk of the hazard occurring, and how to plan and control for these hazards. Therefore, it is not only advised that organisations use SRAs in a strong SMS, but SRAs are also required by GB law.

However, not all safety incidents can be totally controlled for, and they could occur during commercial AV use. Therefore, organisations should have in place effective systems for reporting these safety incidents. Safety incident reporting is a reactive way to control for risk, compared to SRA, which are proactive. Safety incident reporting must take place after an incident has occurred. If there is no process for reporting safety incidents in an organisation, this should be a cause for concern, and evidence that there is not an effective SMS or safety culture in place. Such lack of a feedback mechanism would lead to concerns about an organisation’s ability to take responsibility for adapting software, repairing sensors or applying learnings from any other form of in-use monitoring from these incidents.

A way to facilitate safety reporting is through having a dedicated safety reporting form (SRF) that all relevant employees are aware of and trained in how to complete. This means that both major and minor incidents have a channel to be reported through and that safety personnel are aware of these incidents. These SRFs should contain information that allows for identification of the incident, description of the incident, names of the responsible parties, and then feedback given by key safety personnel together with plans on how to proceed. These forms should include the following in the appropriate detail and format for the organisation:

- Date, time, and location of the event
- The name of the reporter and role/organisation area
- Their role in the project
- Details of the incident, not limited to:
  - Context and events that happened before the incident



- How the incident occurred and who was affected
- The severity of the incident (possibly through a defined rating scale)
- Whether people were injured, or property damaged
- Whether the general public was involved
- Steps taken post-incident as safety controls and processes were followed
- Any suggestions as to the cause of the incident or errors that lead to it
- Suggestions on how to prevent future incidents and how serious future similar incidents could be.

There should also be a system of documenting report numbers and references, to evidence safety reporting over time and allow back-checking on previous reports and subsequent responses. This reporting should be performed by the personnel overseeing deployment during the incident, but there is also a role for key safety personnel to take for the reports, who should document information including and not limited to:

- Recommended actions to be taken to prevent future incidents after review, including improved oversight of errors that lead to the incident and improving human handover
- The responsibility for the incident and whether there may be criminal liability (Appendix 6 - B2)
- Resources required for changing or updating the SMS to account for the incident
- Whether the change requires a software, vehicle, or human factors change
- Follow-up actions to take including by whom and when (SM ICG, 2015).

Having an operational and well-defined method of safety reporting is a required feature of an SMS and can be clearly evidenced through templates and guidance documents. Employees should also be educated and aware of these safety reporting methods so that when they occur, negative impacts can be mitigated and be used as learning opportunities in the future.

A positive safety culture will likely make substantial use of safety reporting, even for minor or regular/ mandatory reports. Use of reports means that the organisation is taking responsibility for the incidents and not 'sweeping them under the carpet', instead being open about how to improve their safety policies so that the necessary changes can be implemented more effectively.

### 10.6.4.3 Appendix 6 - D3: Updating Safety Objectives and Performance Indicators

Appendices 6 - B3.2 and 6 - B3.3 explain the importance of creating relevant and operationalised SOs and SPIs. Like the rest of the SMS, these need to be dynamic and develop as the focus and data used by an organisation change. It is likely that many SOs will not change as frequently as SPIs, since the general safety goals for an organisation will typically be more consistent, including reducing the number of safety incidents dependent on certain products and projects. From an earlier example, an organisation may alter its focus from automated pod transport in urban areas to off-highway materials transport. This would trigger a significant need to change any SOs relating to the populations affected or the design of the AVs, since the applications of their processes will now be different. These SOs should be continually assessed, and any changes documented, both to maintain relevant SOs and to provide a history of SOs; the latter will be of value to identify whether previous SOs have been met, are outstanding, or need to be reinstated. This should again be the responsibility of key safety personnel, since they hold primary responsibility for updating the SMS documentation and application.

SPIs are suggested to be more frequently and specifically updated as these are means by which the SOs are measured and assessed; this is especially pertinent for commercial AV deployment. Appendix 6 - D deals with incident monitoring and refers to adjusting operations by way of new data analysis and data types; this is relevant to SPIs as well. SPI data should be checked through the safety performance on projects, such as number of safety incidents or certain other outputs. It must be made sure that these SPIs are relevant to the data that the organisation is currently collecting and monitoring. Due to the innovative nature of AV research and technology, new updates may be frequent, as new software and new methods of monitoring will be developed and introduced regularly. Therefore, the dynamic nature of an SMS should be evidenced by the relevance of the SPIs for an organisation, and how these SPIs change depending on what is being assessed.

SPIs must be used to assess how well an organisation is performing against their safety objectives and must also be regularly assessed themselves as to how relevant they are to the work an organisation is involved in. They should be checked against measures including incident frequency and type, because any SPIs that no longer accurately represent the contemporary work programme will very likely not be promoting any improvements in safety. This is recommended to be both a regular process, but also to be performed in response to changes to projects, methodology and other SMS updates, such as to safety reporting methods and objectives. One key trigger that might result in frequent updates is changes to the ODD and TOD definitions resulting from new permutations being identified, which may result in the need for SPIs to be adapted accordingly.

It is recommended for organisations to be able to show evidence of documented changes, versions, and differences in SPI application across time. If an organisation can demonstrate that they have adapted SPIs consistently in response to safety concerns, changes in organisational objectives and safety failures, then they can be seen as having a positive safety culture.

SPIs are one of the most important considerations when assessing the suitability of an SMS for commercial AVs, since reliance on data analytics and innovation will require frequent work to keep these SPIs relevant to the performance of organisations. Frequently and accurately updated SPIs are also good evidence of strong safety culture, in the same way as safety reporting in Appendix 6 - D.2, since not only is safety performance being assessed, but the relevance and performance of the assessment measures themselves. This demonstrates a commitment to continued safety oversight, as does good performance against the set SPIs.

#### 10.6.4.4 Appendix 6 - D4: Employee Consultation

Improvements and changes to the SMS and contributions to safety culture originate from the performance of the safety processes and data measures described previously, but it is also important for organisations to utilise the contribution of employees in improving SMSs and their application. This goes together with employee education and awareness (Appendix 6 - C) in making sure that employees not directly attached to key safety personnel have both the opportunity (channels) and the ability (through education and training) to provide safety-related feedback.

Consulting employees at all levels about the SMS and how they work within it is key to a positive safety culture. This consultation can come in various forms, including through regular catchups/ workshops and through dedicated channels that are always available for employees to contact key safety personnel. It is important that there is *“a mechanism to regularly and accurately report safety concerns, including provisions for employees to submit feedback on the process”* (SAE, 2021). This need not be anonymous internally unless requested, as it needs to relate to the specific areas the employees work in for it to be actionable feedback. Any non-anonymised issues or feedback must not be brought up for reprisals, as a positive safety culture should work towards *“fostering a voluntary, cooperative, non-punitive environment for the open reporting of safety concerns”* (FAA, 2018). This also directly relates to the Law Commissions' proposed 'duty of candour' (Appendix 6 - B2), as employees being able to report on their concerns within a no-blame culture moves an organisation towards an effective overall view of safety management.

All employees at all levels/ areas of the organisation should have this opportunity to raise safety concerns, because some issues may not be effectively recorded by safety audits, risk assessments or post-incident reporting. These are more likely to be qualitative measures of change (streamlined processes, ease of use, new ideas) than quantitative (data analytics and measures), since the consultation is focussed on the experience of the employees within the organisation. For example, a technician working on AVs may see that non-technicians are not following the operational health and safety rules in a garage when they are viewing the AVs, and suggest that they be informed and trained on this before visiting an active garage. These are experiential issues with the SMS that would not be identified without direct consultation of employees, demonstrating the need for these channels. Formats for such reporting should allow a level of flexibility; this will help avoid situations where staff are unable to provide a comprehensive account due to very prescriptive boxes in a form not including suitable categories to capture key information.

Documentation and identification of issues is the first step, but integrating this into the continually evolving and improving SMS is key to this being effective. A strong safety culture uses employee consultation as another means to update and improve the SMS according to the opinions and real-world application of processes and principles from the SMS. This consultation depends on the employees' knowledge and understanding of the SMS as it relates to them, and how straightforward it is to report their concerns and have them documented.