

Colour code explanation

Initial inputs from OPI (Officer of Principle Interest) representatives are highlighted in **turquoise**. These reflect preliminary thoughts. The draft text proposal will be highlighted in **yellow** and finalised text will be highlighted to **green** subject to consensus during formal sessions.

Note:

The document uses the terminology “assessor” to refer to the actor performing the safety case assessment and the confirmatory testing. Under UNR, “assessor” is ~~should be~~ understood as the “Type-Approval Authority or the Technical Service on its behalf”. Under GTR, the “assessor” is ~~will be~~ defined by the applicable approval/certification regime.

Paragraph 6.3.2.2.3 (GTR) / 8.3.2.2.3 (UNR)

Pr. High

Topic: Concept of “sufficient” scenarios for fall-back response as part of safety case testing assessment

Regulatory text:

*GTR: 6.3.2.2.3.: “The assessment shall verify that the set of scenarios and situations resulting from the manufacturer’s scenario generation and identification process is suitable for demonstrating the ADS safety case. This includes covering reasonably foreseeable situations and conditions that the ADS will encounter during its real-world operations. **In particular the assessor shall verify that the set of scenarios and situations selected as evidence to support the ADS safety case includes:***

- (a) ***Scenarios and situation in which the ADS needs to initiate a fall-back response (e.g., approaching the ODD limits), and.....”***

*UNR: 8.3.2.2.3.: “The approval authority or its designated technical service shall verify that the set of scenarios and situations resulting from the manufacturer’s scenario generation and identification process is suitable for demonstrating the ADS safety case. This includes covering reasonably foreseeable situations and conditions that the ADS will encounter during its real-world operations. **In particular, the approval authority or its designated technical service shall verify that the set of scenarios and situations selected as evidence to support the ADS safety case includes:***

- (a) ***Scenarios and situation in which the ADS needs to initiate a fall-back response (e.g., approaching the ODD limits), and.....”***

GTR 5.3.2.15. / UNR 7.3.2.15. The safety concept shall describe the manufacturer’s approach to scenario selection to cover the reasonably foreseeable situations and conditions that the ADS will encounter including how the following aspects are covered:

- (a) ***The selection of sufficient scenarios in which the ADS needs to initiate a fall-back response (e.g. approaching the ODD limits),***

Interpretation:

Explanation of the requirement

A fall-back response means:

- “...a system-initiated deactivation procedure or
- an ADS-controlled procedure to place the vehicle in a mitigated risk condition (MRC).”

A system-initiated deactivation is defined as “...a procedure by which the ADS initiates the transfer of performance of the DDT from an ADSF-1 to a fallback user.”

An ADS feature of type 1 (ADSF-1) means an ADS feature which includes an ADS fallback response requiring a fallback user, whereas an ADS feature of type 2 (ADSF-2) does *not* include an ADS fallback response requiring a fallback user.

In the context of this requirement ‘sufficient’ refers to two aspects:

1) The number of scenarios in which a fall-back response is initiated by the ADS feature;

2) The number of users participating in assessing the effectiveness of the user interaction aspect of the ADS fall-back process.

‘Sufficient’ in both 1) and 2) refers to the trust in the outcome of the tests that has been performed on the fall-back response. If either of them is not ‘sufficient’ then the evidence ~~should be~~ is considered insufficient to support the claim.

For example, evaluating a large set of scenarios with a very limited number of participants provides little statistical power and does not support reliable inference. Conversely, evaluating a small subset of scenarios with a large and diverse participant sample may yield statistically meaningful results, depending on the proportion of the total set of relevant scenarios that is covered. The number and diversity of relevant scenarios are inherently dependent on the intended use case and its operational design domain. The choice of the number of participants also depends on what an ADS user needs to do / understand. In case of an ADSF-1 a fall-back user needs to do and understand different things than the passenger of a vehicle with an ADSF-2.

In statistics methods have been developed to assess whether the sample size is ‘sufficient’. These methods focus on the sample size in terms of number of participants. The manufacturer should provide the calculations that lead up to the choice of the number of participants and what are the assumptions used in the calculations.

With respect to the type and number of scenarios the manufacturer is expected to describe the reasoning behind their choice scenarios, why they claim they are representative and the justification for the overall number chosen.

The number of scenarios and/or of participants are only two aspects in the development of a methodologically sound setup in which the interaction between the vehicle and the user is tested. For this kind of testing there are well established practices that indicate what constitutes a methodologically sound setup.

Paragraph 6.3.2.4.1.4 (GTR) / 8.3.2.4.1.4. (UNR)

*Pr. High***Topic: Representativeness of the other road user for safety case testing assessment****Regulatory text:**

GTR: 6.3.2.4.1.4.: “For the specific case of ADS interaction testing, the assessment shall:

- (a) Verify that the people involved are representative of the expected general population of ADS users and other road users where applicable,....”

UNR: 8.3.2.4.1.4.: “For the specific case of ADS interaction testing, the approval authority or its designated technical service shall:

- (b) Verify that the people involved are representative of the expected general population of ADS users and other road users where applicable,....”

Interpretation*Explanation of the requirement*

The concept of representativeness aims to ensure that tests involving other road users provide credible, generalisable evidence about ADS behaviour in real traffic. The reference to “**representative of the expected general population of ADS users and other road users**” is intended to prevent safety claims being based on narrowly selected participants, atypical or overly compliant behaviour, or participants’ choices that do not reflect real-world diversity. It is important that other road users exhibit the full spectrum of normal and unusual behaviours and choices that the ADS might encounter so as to fully investigate the response of the ADS.

This concept is complemented by the notion of *coverage*. Coverage in this context can be considered as a measure of how well the testing “covers” the space that all relevant categories of other road users, interaction types, and traffic situations within the declared ODD are included in the safety case, such that no safety-relevant population or scenario is systematically excluded.

Together, coverage and representativeness ensure the validity of test results across the full range of situations likely to be encountered by the ADS.

In this context, representativeness does not mean a perfect demographic mirror of society. Instead, it means that the characteristics and behaviours of other road users involved in testing are diverse but sufficiently typical of those the ADS can reasonably be expected to encounter, such that conclusions drawn from the tests are valid.

Representativeness has three complementary dimensions:

- 1) Behavioural representativeness, which ensures behavioural patterns of other road users reflect real traffic behaviour (e.g., compliance and non-compliance with traffic rules, reaction times and decision-making variability, assertive, cautious, inattentive, or ambiguous behaviour, natural interaction dynamics (e.g., gap acceptance, yielding, hesitation).
- 2) Population representativeness, which ensures that the group of other road users is representative of the expected population (e.g., age range, physical characteristics relevant to perception (e.g., height), mobility characteristics

(e.g., walking speed, cycling style), driving experience and style (for vehicle drivers).

- 3) ODD-specific relevance, which ensures that the representativeness is always relative to the declared ODD.

Paragraph 6.3.3.1 (GTR) / 8.3.3.1 (UNR)

Pr. High

Topic: Scope of Confirmatory testing

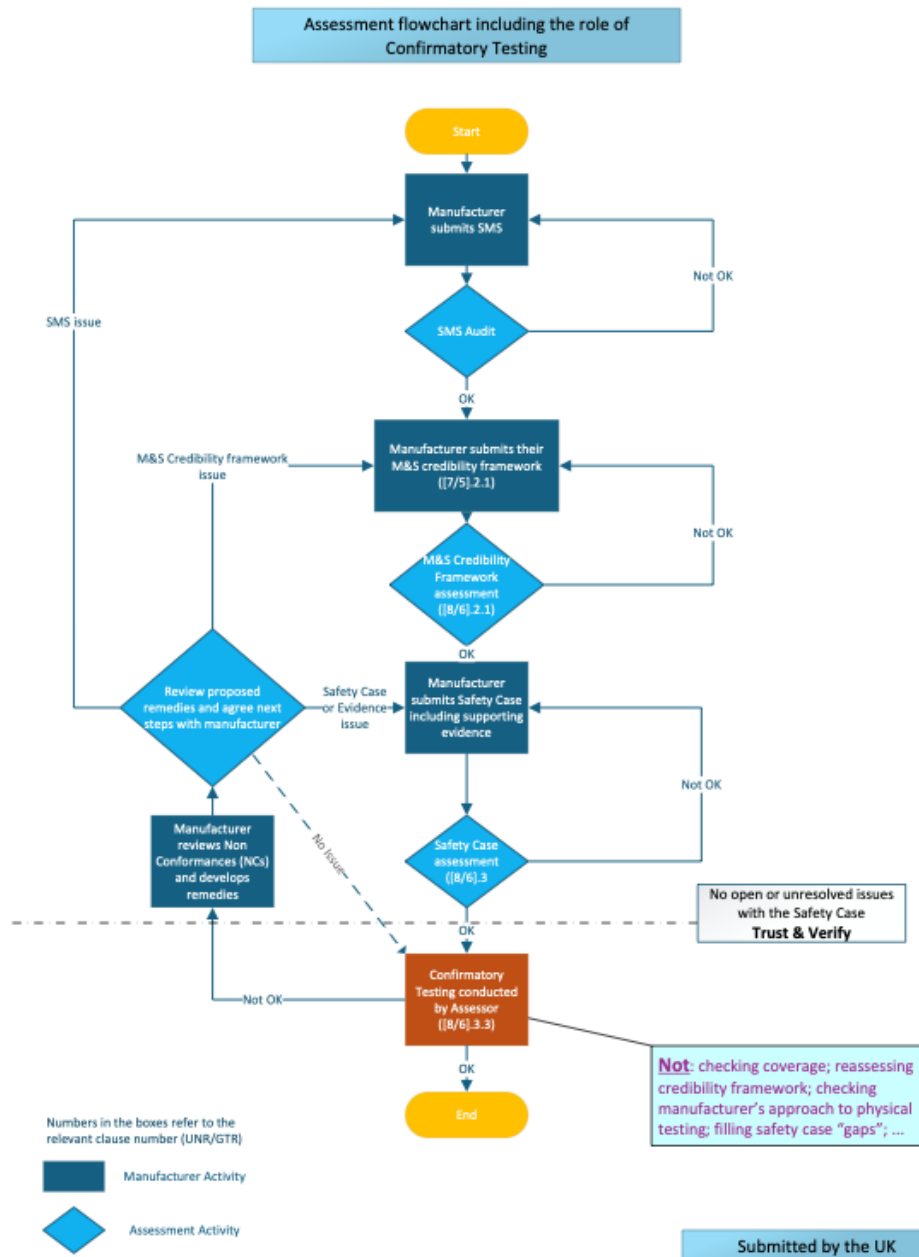
Regulatory text

6.3.3.1.: “At the option of the Contracting Party, confirmatory testing may be required to use one or more test methods and pre-defined and repeatable test protocols to confirm that the evidence provided by the manufacturer accurately represents the ADS performance. The confirmatory tests shall cover a range of driving conditions representative of the ODD, including at least and as appropriate.”

8.3.3.1.: “Confirmatory testing conducted or required by the approval authority or its designated technical service shall use one or more test methods and predefined and repeatable test protocols to confirm that the evidence provided by the manufacturer accurately represents the ADS performance. The confirmatory tests shall cover a range of driving conditions representative of the ODD, including at least and as appropriate.”

Interpretation: ~~Explanation of the requirement~~

The ADS assessment process under this Regulation is centred on the evaluation of a manufacturer’s safety case, where the confirmatory testing appears as the last assessment phase. The flowchart below illustrates a hypothetical ADS certification/approval process with the different steps of the assessment process and clarifies the stage at which confirmatory testing takes place. The flowchart is only provided as an explanatory visual aid and should not be intended as a prescriptive procedural requirement. The approach shown does not preclude different strategies, including the parallelization of some steps.



Confirmatory testing conducted by the assessor should be understood as a targeted and proportionate activity aimed at providing confidence in the credibility and internal consistency of the evidence supplied in support of the safety case, taking place after the latter has been assessed. The confirmatory testing is intended to verify selected evidence and is not meant to replicate in its entirety or replace the manufacturer's validation and verification campaign. The scope of confirmatory testing should therefore consist of representative samples of scenarios or operating conditions that are part of the safety case.

The assessor can accept, allow, or introduce variations to the tested scenarios or conditions, where this is considered necessary to verify the accuracy which the evidence is representative of the ADS performance. Such variations should be

reasonable and compatible with the limitations/abilities of the toolchains¹ and allow comparison with the results provided by the manufacturer or be assessable against the expected behaviour of the system. In this context, reasonable means that the ADS should be tested in scenarios/situations that are likely to occur in the real world, rather than in unrealistic situations that do not conform to the logic of the real world.

Variations may include, for example, adjustments to the initial conditions, parameter ranges, or environmental assumptions within the bounds or at the boundaries of the declared ODD, but may also consider the behaviour at ODD exit or outside the ODD. Variations may also arise inherently from the nature of the testing environment. Considering the real-world confirmatory testing, the assessor is unlikely to be in a position to replicate the manufacturer's test method and protocols. The real-world confirmatory testing may be conducted on test routes different from those of the manufacturer, provided that the route selection related to the ODD of ADS. Similarly, while carrying out proving ground testing, it might not be possible to control all the variables, and there might be stochastic effects in the controlled variables. Whenever variations appear or are introduced, the purpose of the confirmatory testing is to verify the robustness of the ADS and the consistency of the results provided by the manufacturer.

¹ It is expected that the assessment of the credibility and suitability of the testing toolchains takes place before the confirmatory testing to ensure the assessor is aware of any potential limitation and can make an informed assessment of the safety case.