VMAD

Subgroup 2b: Development of Test matrix Automated Vehicles

*Generic Test approach for automated vehicles on a test track and real world testing*

*(Follow-up of SG 2b Case study ALKS)*

Brussels, 25 March 2020

**Introduction**

As a next step in the design of a test approach for automated vehicles on a test track and in real world, a case study has been presented during the VMAD meeting on 16 and 17 January in Tokyo, Japan. In this case study a stepwise method was presented for the design of such a test approach in order to demonstrate safe traffic behaviour on public roads. After the discussion in January in Tokyo, this approach has been further elaborated. The aim of this document is to present a basic model for testing automated vehicles on safety / safe behaviour. The test should also deliver suggestions for improvement of the system(s) / automated vehicle. After acceptance of this approach several steps will have to be taken to further operationalize a validate this test model. These steps will also be described in this document.

**Test matrix**

During a first test phase the automated vehicle will be tested in a virtual test surroundings. either performed by the (car)manufacturer or an independent, third party (following the proposals of subgroups 1a and 2a). In a next step the automated car will be tested on a test track, based on the report(s) of the previous, virtual test phase. The report will describe the behaviour of the automated vehicle / -systems in the virtual surroundings regarding the operational performance, limits and conditions. In fact, these test report(s) will be the input for testing on a test track. In the case study[[1]](#footnote-1) a basic approach for the development of a test matrix or protocol was presented. The test criteria in this matrix to be developed should be based on:

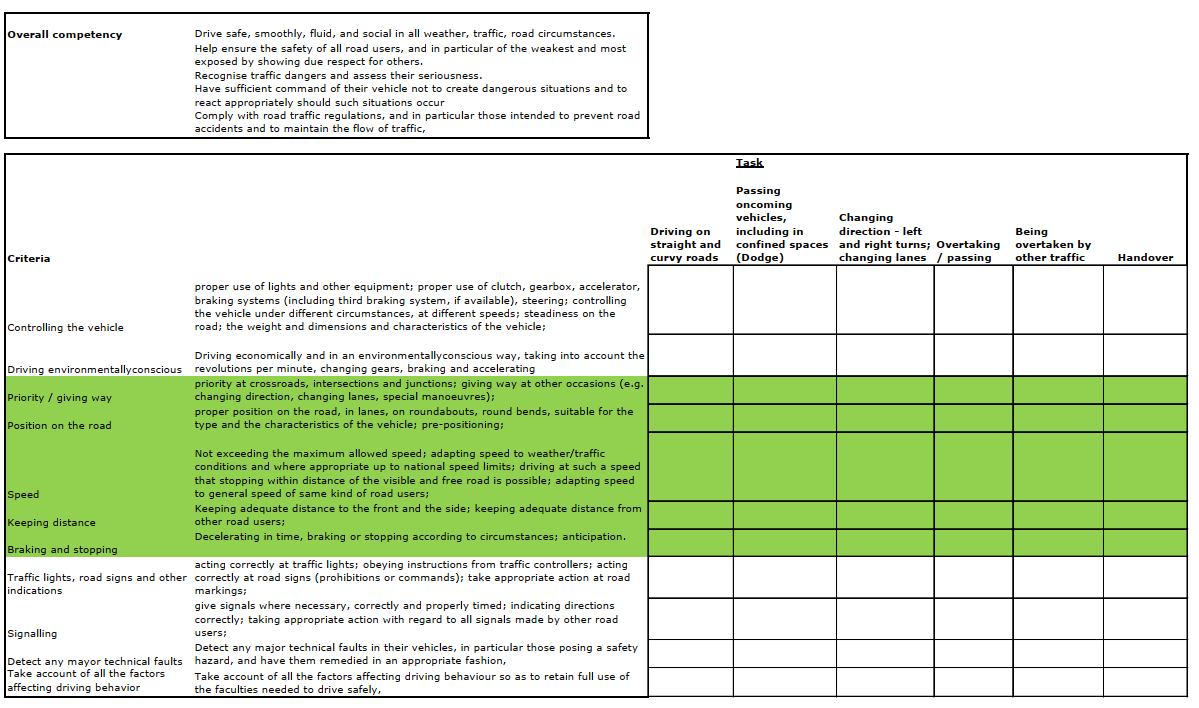
* Safety - awareness of hazards (based on actual behavior of the vehicle), ability to modify behavior in order to prevent accidents happening
* Coverage of the traffic situations (abrupt changes)
* Coverage of the environments ( (outside) urban area, mixed traffic, controlled environment)
* Coverage of circumstances (demanding conditions; weather, congestion)
* Human intervention (need)
* Fluent and smooth driving

**General approach**

It was agreed that the test model first of all should be focused on ALKS (Automated Lane Keeping Systems). When testing a car however it is quite difficult only to test a single system. In the context of the ability to drive concept (and the holistic approach of testing this concept), testing an isolated element or skill has little meaning. Systems interfere with each other and while putting one system “under pressure” (on a test track), other systems should give support, become operational or at least should recognize that (other) hazards might be developing and therefore should intervene and/or warn the driver to take back control (Human Machine Interface). Also the behaviour of the (automated) vehicle should be predictable in an ever proceeding automation in mixed traffic. Exactly for these reasons a generic test model has been designed in order to test the car in a holistic way on (traffic) behaviour in interaction with the (human) driver.

The next page shows the proposal for the test protocol. Step by step each element of the model will be explained. Also attention will be given how the model should be further elaborated in order to create a validated test protocol. Several experts in testing have been working on the design of this model just as driving examiners with expertise in the use of automated systems while driving related to specific traffic situations.

*The basic model for a test protocol:*



**Overall competency**

This generic competency describes the “behaviour” the automated vehicle and/or systems should demonstrate at every moment and in every situation when taking part in traffic. This competency corresponds also with the “safety envelope” as described in the case study mentioned earlier: “…the car (system) should be capable of maintaining a safety envelope in the direct surroundings of the car. This means, dependant of speed, specific (traffic) situations there should be a safe space ( “cushion”) in the direct surroundings of the car in relation to other road users”. All the other competencies, described in the test matrix as proposed are derived from this overall competency. In order to measure this overall competency several tasks have been described and operationalized in the matrix.

While using an analysis of the “technical skills” of the vehicle (as described by the manufacturer), traffic tasks should be selected. It should be clear that blameable shortages in the safety envelope are directly connected to the performance of the system itself (for instance: not reacting on a “merger” is not blameable, because it is not a task of ALKS). This shows again that a holistic approach in testing automated vehicles is relevant just as the criterium “hand over” (of the responsibility for driving from the vehicle to the driver).

**Criteria and (traffic)tasks**

The left column of the matrix shows the measures used to assess the vehicle when taking part in traffic: is / are the system(s) capable to control the vehicle, to keep proper distance, etc. Howthe vehicle / system should perform is described in the second column in the matrix behind the several measures. Each measure will be “connected” to several tasks: how is the performance of the vehicle / system(s) when it performs the task(s)? Also, it is possible to vary the conditions in which the vehicle should perform the specific task(s): weather- and road conditions, traffic circumstances.

The measures and the traffic situations are a first draft; they might be specified and or elaborated during following expert sessions and during try-outs of the model. Even as the standardization of the test: duration of every task, frequency of the tasks, which circumstances etc..

**Combination Tasks and criteria (cells in the matrix)**

Measures and traffic situations will “meet” in the several cells of the matrix. In other words: a value will have to be added to each cell: is the performance of the vehicle / system in that specific (traffic) situation sufficient or not? The values given can be quite simple: pass or fail.

Also a number of points could be given for each measure connected to the traffic situations (for instance between 1 – 10). A minimum number of points to be earned will determine whether the vehicle / system will pass or fail for this situation. The total of the columns will determine whether the vehicle / system will pass the whole test.

Written comments for each cell / column and / or a general written conclusion with feedback (for the manufacturer) could also be an alternative.

Which system would be the best is difficult to say at this moment. The intention is to start with a pass / fail system based on a written procedure for the examiner / tester: how should the vehicle / system perform in the several traffic situations for each specific measure (elaboration of description in column 2 of the matrix)? This approach of a pass / fail system is successful in use nowadays and therefore could be an efficient first step in the development of such a system for automated vehicles.

**Scoring and objectivity of the test**

However scoring of the test could or should start as described above, a first prototype regarding the approach and design of the test(matrix) should be further elaborated, defined and confirmed. This will be done in the next step of this development process during a meeting with traffic and test experts.

However a guarantee for a 100% objectivity is not realistic, continuous attention for the reliability of the test is necessary just as calibration sessions. Also the length of the test just as the number of tests can increase the reliability. While running this process data could be collected, supplemented with expert based acting. Based on both entities, a probabilistic model could be developed in order to guarantee maximum objectivity.

One result will be a procedure (for the testers) to be used for assessing the performance of the vehicle. A basic principle for developing such a procedure could be the so called “TSN system”: The type, seriousness and number of deviation(s) in behaviour/performance will be shown by the vehicle. During an expert meeting such a method (aim: a well argued standard setting(sprocedure) should be designed.

Another approach could be the so-called “Rubric” method: what should be observed to assign a specific value (mostly between 0 – 4 or 0 – 5) to (a) certain performance / behaviour. This should be set by a group of (independent) experts.

Both methods might look as a subjective way of assessing. However by using the system of cross- / inter-assessments (2 or more assessors) the test can be validated and calibrated (inter-objectivity). Training of testers / examiners is also an important tool to create and maintain the reliability of the test.

**Test track: scenarios**

Basically the car should perform well in standard, relatively simple traffic situations. Traffic however is unpredictable and meets variations in intensity and complexity (edge cases). Also the preconditions will differ and are also unpredictable: weather, road markings, obstacles on the road, etc. Performance of the automated vehicle (systems) in situations as described are relevant for safe behaviour of the vehicle and should therefore be part of the test (vehicle “under pressure”): on the test track (testing in a safe, pre-conditioned surroundings), but also in real traffic (in search for specific situations; definition of test routes).

A description of scenarios, selected by traffic and test experts (general approach; to be further elaborated):

Relevant for ALKS:

* Rainy weather; (road) linings cannot be seen properly by the driver.
* Branches, debris, remains of car tyres, on the road.
* No road linings present.
* Performance in case of double linings (road works), or old linings.
* Specific speed zone (60 km/h): no middle lining, only side linings.
* (Sudden) narrowing / widening of the road.
* Things and / or vehicles reaching over the middle lining of the road (exceptional transport).
* Interrupted road linings.
* Tunnels: from dark to light and vice versa.

Relevant for generic test model:

* Oncoming traffic: swerving. Oncoming traffic taking the inside curve.
* Type of pedestrian / cyclist.
* Closed lane. Express lane opened.
* In case of an emergency: which system will be prioritized?
* Blind or impaired person(s); infirmity.
* Cyclists, cycling next to each other (2, 3).
* Other vehicles with parts sticking out.
* A truck next to the ego vehicle in a bend with two lanes.
* Unpredictable behaviour / anticipation on other road users.
* Cars driving too close behind, creating more space in front.
* Situations where lanes should be changed ahead: merging slow traffic.
* Which system has priority?
* Oncoming traffic not following its lining.
* Several situations regarding priority (giving way) with cyclists and mopeds.
* Several (inconsistent) road settings: what is the appropriate (maximum) speed.

While developing the (generic) scenarios, three different types could be distinguished:

* The most common scenarios
* The most dangerous scenarios
* Based on the OD: contextual scenarios

**Real world: in traffic test**

Once the automated vehicle has passed the stage of the test track, an in traffic test will be the next and final test phase. During a certain amount of time the vehicle will drive in real world traffic. The same test matrix will be used while conducting and scoring the test for the automated vehicle.

On a test track the vehicle will be tested on specific scenarios where the circumstances can be varied in order to see how the vehicle will react to specific circumstances (putting the vehicle under pressure). Once entering the borders of its OD we want to know how the vehicle behaves and whether and how it will hand over control to the driver whenever necessary.

In real world traffic however we want to see how the vehicle behaves in its natural habitat. Instead of testing how the vehicle will react in a specific (artificial) situation to a certain situation we want to know how the vehicle (system) deals with a variety of (traffic) situations that might occur all together. In other words: in real world traffic there are many variables at once the vehicle has to respond to.

For the design of this test specific in traffic courses can be designed. It is impossible to test all the elements as mentioned in the test matrix in one or two in traffic tests (in contrary to testing on a test track). This means in traffic tests should be designed in a way there is a chance all the tasks as mentioned in the test matrix might occur. And as already mentioned, the added value of this test is in the combination of tasks to be performed by the vehicle (system) as they appear in daily traffic; sometimes rather complex and unexpected (edge cases). A very important issue in these cases is the interaction between the human driver and the automated vehicle (hand over).

**Annex 1**

**Steps to be taken (follow-up)**

The aim of this document was to design and present a concept (prototype) for a test matrix to be used for testing automated vehicles in combination with the role of the human driver (HMI). Several next steps are necessary in order to validate the test; organization of try-outs; maximalization of the reliability, etc. This annex shows in general an overview of these different stages. The outcome should be finally a validated and reliable test for automated vehicles.

*Expert meeting for development first prototype of test matrix*

During this stage a combination of several experts will be evaluating the test matrix. They will be dealing with (amongst others) the following issues:

* Completeness criteria test matrix
* The usefulness of criteria in relation to the tasks
* Sufficient number and description of traffic situations (tasks)
* Values to be added to the cells of the matrix (scoring system)
* Description of the test and handling the scoring procedure
* Design and description scenarios test track
* Design and description in traffic test courses (in traffic test)

*Selection and training of testers (examiners)*

Training of examiners / testers in using the test and scoring procedure for creating objectivity. Description and operating system inter-objectivity during and after try-outs of the test methodology (see chapter **“**Scoring and objectivity of the test”).

*Try-out*

Once the examiners / testers have been trained and prepared, try-outs can be arranged. On a test track and in real world using the scenarios and in traffic courses.

The try-outs will deliver us tools to improve the validity of the test (are the tasks complete, which tasks will have to be added or deleted, etc). Also there will be attention for further standardization of the test: duration of every task, frequency of the tasks, which circumstances, etc.

1. *Test proposal ALKS test track and real world testing*, Brussels, 29 November 2019 [↑](#footnote-ref-1)