AVSC00008202111

Issued 2021-11
Superseding

Automated Vehicle Safety Consortium™ A Program of SAE ITC

Automated Vehicle Safety Consortium™ Best Practice

AVSC Best Practice for Evaluation of Behavioral Competencies for Automated Driving System Dedicated Vehicles (ADS-DVs)

Rationale

Driving safely is a complex task involving a broad range of skill sets invoked in a vast number of potential scenarios. Automated driving system (ADS) developers are faced with the significant challenge to build an automated driving system-dedicated vehicle (ADS-DV) with the capabilities necessary for driving safely. There is value in assessing a set of behavioral competencies as directional indication of safety performance and as a starting point for additional assessment. This best practice provides an approach to specify ADS behavior by clarifying lexicon surrounding ADS behaviors, enumerating an elemental set of behaviors to consider, and demonstrating how to derive applied metrics to evaluate behavioral competence in practice.

Preface

The Automated Vehicle Safety Consortium™ (AVSC) is an industry program of SAE Industry Technologies Consortia (SAE ITC®) working to quickly publish best practices that will inform and lead to industry-wide standards advancing the safe deployment of automated driving systems (ADSs). The members of this consortium have decades of accumulated experience focused on safe, reliable, and high-quality transportation. They are committed to applying those principles to SAE level 4 and 5 automated vehicles so that communities, government entities, and the public can be confident that these vehicles will be deployed safely.

The Consortium recognizes the need to establish best practices for the safe operation of ADS-dedicated vehicles (ADS-DVs). These technology-neutral practices are key considerations for safely deploying ADS-DVs on public roads. Members of the AVSC intend to support the published principles and best practices in an effort to establish a suggested level for other industry participants to meet. These best practices will serve as a basis to enhance and expedite the formal industry standards development process through SAE International and other global standards development bodies. Effectively implementing these principles can help inform the development of sound and effective ADS regulations and safety assurance testing protocols that will engender public confidence in the efficacy of ADS-DVs.

Comment and open discussion on the topics are welcome in appropriate industry forums. As discussion unfolds, AVSC documents will be revised as significant information and/or new approaches come to light that would increase public trust.

Introduction

ADS developers may decompose the dynamic driving task (DDT) into a generalized set of behavioral competencies that are deliberately broad as a means of evaluating as many DDT subtasks as possible. Subsequently, developers use system engineering techniques to ensure that the decomposition of the driving task maps well to the set of generalized competencies. In other words, the ADS may not always invoke a single competency, but a set of competencies covering a predictable part of the driving task. Although real-world conditions involve complex interactions among numerous systems in various situations, mastery of a broad set of competencies provides evidence of, and increases confidence in, baseline safety performance. Similarly, well-established system engineering techniques inform the validation of the driving capabilities by prescribing test criteria and integrating the results.

This approach, to some extent, resembles a human driver test. A driver's test evaluates human drivers on a small number of samples for a set of behavioral competencies with the assumption that humans can generalize broadly and effectively to execute the driving task in a larger sample. However, as noted in the RAND report on Acceptable Safety for Automated Vehicles, "Although a human-driving test reflects ability to execute key driving maneuvers (behavioral competencies) and a knowledge of driving law, it does not reflect overarching driving skill in every situation." [1] Evaluating behavioral competency in a key set of scenarios for a given ODD contributes to ADS safety performance and provides relevant evidence for AV safety assurance on public roads.

Table of Contents

Rationale
Preface
Introduction
1. Scope
1.1 Purpose
2. References
2.1 Applicable Documents
3. Definitions
3.1 Behavior
3.2 Behavioral Competency
3.3 Dynamic Driving Task (DDT) (SAE J3016_202104)
3.4 [DDT Performance-Relevant] System Failure (SAE J3016_202104)
3.5 Feature (SAE J3016_202104)
3.6 Maneuver
3.7 Object and Event Detection and Response (OEDR) (SAE J3016_202104)
3.8 Operational Design Domain (ODD) (SAE J3016_202104)
4. Defining Behavioral Competencies
4.1 Components of Behavioral Competencies and Safety Assurance
4.2 Elemental Behavioral Competencies
5. Building a Behavioral Competency Evaluation
5.1 Formulation of Behavioral Competency Evaluation Context
5.2 Application of Metrics for Behavioral Competencies
5.3 Aggregate On-Road Testing Assumptions and Thresholds
6. Summary
7. About Automated Vehicle Safety Consortium™
8. Contact Information
9. Acknowledgements
10. Abbreviations
Appendix A. Best Practice Quick Look
Appendix B. References
Appendix C. Behavioral Competency Reference Table

1. Scope

This Automated Vehicle Safety Consortium™ (AVSC) ***Best Practice for Evaluation of Behavioral Competencies for Automated Driving System-Dedicated Vehicles*** **(*ADS-DV*) (AVSC00008202111)** describes an elemental set of ADS behavioral competencies and sample application-specific metrics for safety performance of fleet-operated ADS-DVs. The behavioral competencies are intended to cover SAE level 4 and 5 ADS-DVs operating on public roads, with or without human passengers.

1.1. Purpose

This best practice is intended for use by the technical community (developers, manufacturers, testers, etc.) to aid in the standardization of testing and safe deployment of ADS. It may also be useful to public agencies and stakeholders, including companies supporting ADS developers, standards bodies, and governmental decision-makers, with an interest in better understanding the safety posture of ADS deployments.

This best practice supports public and private organizations in preparing for and deploying ADS-DV systems. For example, this best practice may be used by ADS manufacturers and developers to document the aggregate safety performance of vehicles within the target Operational Design Domain (ODD) prior to and during deployment. It is intended to garner public understanding and acceptance by establishing a common language for discussion around measuring safety performance of ADS-DVs.

2. References

2.1. Applicable Documents

The following publications were referenced during development of this document. Where appropriate, documents are cited.

2.1.1. SAE Publications

Unless otherwise indicated, the latest issue of SAE publications apply. Available from SAE International, 400 Commonwealth Drive, Warrendale, PA 15096-0001, Tel: 877-606-7323 (inside USA and Canada) or +1 724-776-4970 (outside USA), www.sae.org.

AVSC00002202004	AVSC Best Practice for Describing an Operational Design Domain: Conceptual Framework and Lexicon
AVSC00006202103	AVSC Best Practice for Metrics and Methods for Assessing Safety Performance of Automated Driving Systems (ADS)
SAE J3016_202104	Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles

For other referenced documents, see Appendix B

3. Definitions

3.1. Behavior

Specific goal-oriented actions directed by an engaged ADS in the process of completing the DDT or DDT fallback within the ODD (if applicable) at a variety of timescales.

3.2. Behavioral Competency

Expected and measurable capability of an ADS feature operating a vehicle within its ODD.

NOTE: Competency refers to the term “expected” in the definition. Using skills, knowledge, and abilities, an ADS executes behaviors competently according to performance criteria set by the ADS developer.

3.3. Dynamic Driving Task (DDT) (SAE J3016_202104)

All the real-time operational and tactical functions required to operate a vehicle in on-road traffic, excluding the strategic functions such as trip scheduling and selection of destinations and waypoints. Refer to SAE J3016 for a full definition.

3.4. [DDT Performance-Relevant] System Failure (SAE J3016_202104)

A malfunction in a driving automation system and/or other vehicle system that prevents the driving automation system from reliably performing its portion of the DDT on a sustained basis, including the complete DDT, that it would otherwise perform.

3.5. Feature (SAE J3016_202104)

A level 1 through -5 driving automation system’s design-specific functionality at a given level of driving automation within a particular ODD, if applicable.

3.6. Maneuver

Goal-oriented vehicle motion control action undertaken by an ADS to achieve a specific result or outcome.

3.7. Object and Event Detection and Response (OEDR) SAE J3016_202104)

The subtasks of the DDT that include monitoring the driving environment (detecting, recognizing, and classifying objects and events and preparing to respond as needed) and executing an appropriate response to such objects and events (i.e., as needed to complete the DDT and/or DDT fallback).

3.8. Operational Design Domain (ODD) (SAE J3016_202104)

Operating conditions under which a given driving automation system or feature thereof is specifically designed to function, including, but not limited to, operational speed, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics.

4. Defining Behavioral Competencies

An ADS-DV behavioral competency evaluation benefits from comprehensive and consistent terminology to prepare for effective measurement of ADS-DV safety performance. A common lexicon can help build public trust and clearly articulate safety arguments. Like the *AVSC Best Practice for Metrics and Methods for Assessing Safety Performance of Automated Driving Systems* (*ADS*), which proposes a set of metrics and methods for ADS safety performance evaluation, this best practice puts forth a consistent lexicon for ADS-DV behaviors and behavioral competency-related testing concepts that developers can use as a reference. **ADS developers and manufacturers can use this lexicon to help enable consistent communication of safety performance across the automotive industry with respect to behavioral competencies, behaviors, and maneuvers.**

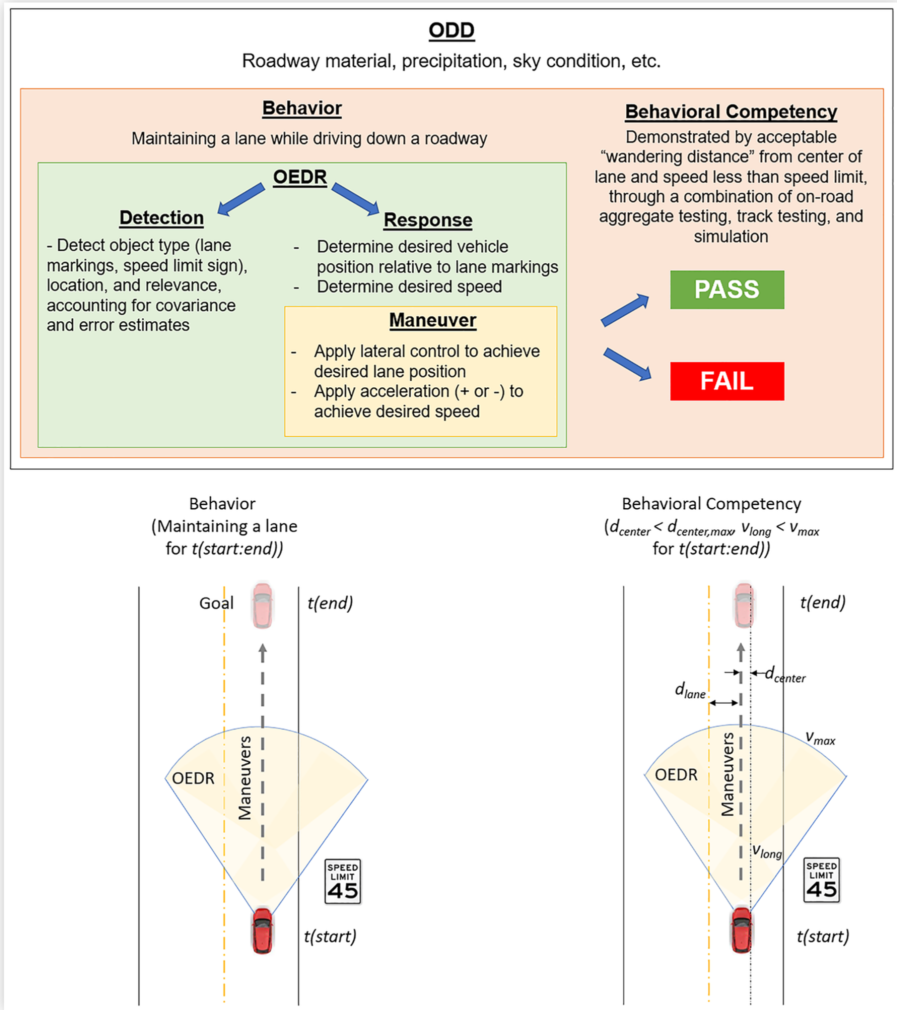
In addition to consistent terminology, this section provides an example of applying behavioral competencies within a safety assurance framework. Behavioral competency evaluation is just one method of providing evidence towards building a safety case for ADS developers and manufacturers. In conjunction with other techniques, using behavioral competencies allows ADS developers and manufacturers to make the case for deployment of ADS-DVs within their ODD (if applicable). Table 1 in 4.2 provides an elemental list of behavioral competencies that ADS developers and manufacturers can use in execution of safety performance testing. Evaluation criteria are then used to demonstrate acceptable ADS-DV behavior for their implementation of the technology internally as well as to external stakeholders.

4.1. Components of Behavioral Competencies and Safety Assurance

Behavioral competency refers to the measurable capability of an ADS to execute goal-oriented actions in a subject vehicle (also known as ego vehicle) while a behavior specifies the goal-oriented action itself. Behavioral competencies represent an abstraction and ontology for partially describing the DDT and can be described using terminology originally proposed in the National Highway Traffic Safety Administration (NHTSA) ADS Testable Cases Framework. In that document, Object and Event Detection and Response (OEDR) is defined as “subtasks of the DDT” and behaviors are “goal-oriented actions directed by an engaged ADS in the process of completing the DDT” [2]. The two concepts are related but different, in that OEDR is specific to DDT subtasks while behaviors can include any action that is goal oriented. Maneuvers as defined in 3.6 can be described as the physical control response an ADS performs to an object or event, and can be characterized as an example of a response the ADS performs as a part of the OEDR process.

To better understand this delineation, Figure 1 illustrates the relationship between a behavior and a behavioral competency using the example of maintaining a lane. On the left part of the diagram, the ADS-DV executes maneuvers to reach its goal. The ADS-DV may be able to execute maneuvers without any existing infrastructure or context. When lane markings, speed limit signs, other road users, etc., are added they create contextual constraints to ADS-DV behaviors executed during DDT.

FIGURE 1 Relationship between behaviors, OEDR, and maneuvers

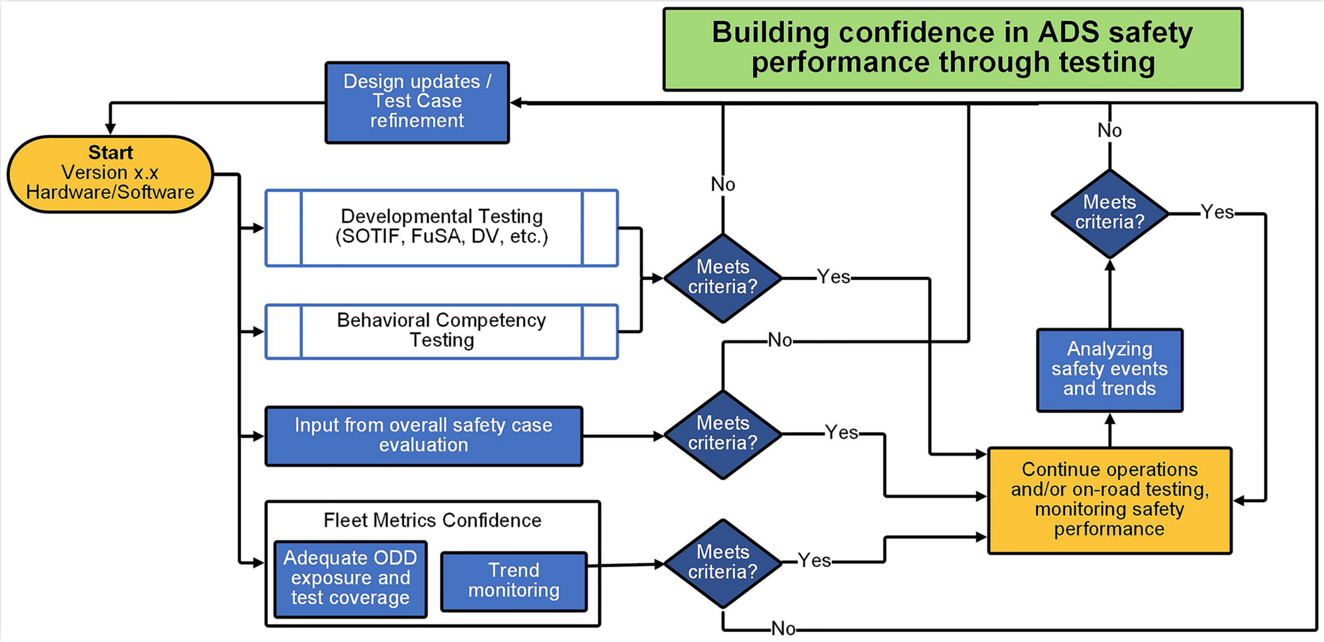


As shown on the right side of the diagram, behavioral competency demonstrates proficiency by layering on success criteria that can be evaluated using metrics with parameters such as longitudinal and lateral distance, OEDR reaction time, and more. An example metric may be "wandering distance" from center of the lane. As discussed in Section 5, the acceptable threshold for the metric depends on the testing approach, for example, the threshold may be a single violation of a specific distance threshold in a concrete scenario conducted on a closed course, or a threshold rate of violations for aggregate on-road testing. As discussed in 5.3, thresholds may be context dependent, for example, the ADS may bias to the left in its lane in order to avoid vulnerable road users entering or exiting parked vehicles on the right.

ADS developers and manufacturers can use behavioral competencies to specify test cases and scenarios. The aggregate evaluation of test cases and scenarios at a system level can than be used to evaluate competency of a behavior. This concept of abstracting behaviors from scenarios and test cases is commonly used in automated system safety assurance frameworks [3]. An overview for integrating behavioral competency testing as part of an ADS safety assurance framework is shown in Figure 2. The system is evaluated through several test methods in parallel; examples of potential testing variations include developmental testing and behavioral competency evaluation.

The flowchart below is conceptual and illustrates how behavioral competency evaluation fits into a broader safety assurance framework. Many of these stages are happening simultaneously and not necessarily in a stepwise fashion. In other words, behavioral competency evaluation may happen in parallel with other forms of ADS testing. In addition, as organizations continue testing and building confidence in ADS safety performance, deployment can take different forms or occur in phases. For example, deployment may start with two in-vehicle fallback test drivers (IFTD) and over time reduced to one IFTD or utilize a remote operator with the goal of expanding ODD coverage.

FIGURE 2 Building confidence in ADS safety performance through testing



Behaviors can span multiple ODDs, and it may not be necessary to test all behaviors for a given ODD. For example, if there are no railroads within the ODD, then detecting and responding to a railroad crossing sign is not required as a behavioral competency. This is accomplished on a case-by-case basis**. ADS developers and manufacturers should identify and utilize behavioral competencies that are relevant to their target ODD**. Although ODD informs what behaviors are required, an ADS developer may choose to restrict the ADS's operating conditions as part of their operational risk management.

In keeping with systems engineering practices, behavioral competency for ADS is first considered in the product development phase, which ISO 26262 describes as ranging from the specification, to design, implementation, integration, verification, validation, production, and release. ADS development testing involves confirming the integrity of the behavioral competency over the operational and environmental range in the ODD. This developmental testing may range from sub-system level testing in simulation to on-road validation of how an ADS-DV integrates into its ODD. Validation may include collecting and analyzing fleet behavioral competency metrics from public road testing to assess generalized conformance with acceptance criteria. Once there is evidence (e.g., from a safety case) that an ADS can operate safely within its ODD, the ADS can advance to subsequent stages in the product development lifecycle. This could include removing in-vehicle safety drivers, limiting assistance from remote operations, and more. If the criteria are not met for testing, fleet metrics, and/or safety case analysis, additional development of hardware and/or software may be needed, the use case for remote operations might be changed, the target ODD may be refined, or other mitigations might be deployed to where the ADS behavioral competency is determined to be sufficient. Finally, safety events and trends are monitored to evaluate and refine ADS safety performance. For example, trend analysis on metrics can help to identify correlations between ADS behavioral performance and operating environment factors. Throughout testing and operations, ADS developers have various criteria to determine whether hardware or software updates may be required to ensure behavioral competency within the ODD.

Refinement of behavioral competencies is expected and can continue beyond development and throughout various stages of testing and commercial operation. Scenarios encountered while driving within the ODD may be fed back to the engineering team to determine if system specification or the definition and testing of behavioral competencies need to be refined to consider system performance under diverse conditions.

4.2. Elemental Behavioral Competencies

In addition to specifying relationships between OEDR, maneuvers, and behaviors, this best practice provides an elemental set of behavioral competencies that ADS developers and manufacturers can use to evaluate ADS safety performance across sets of behaviors and ranges of scenarios. ADS developers can also use the behavioral competencies as a structured starting point for test case development. Each behavioral competency aligns with metrics from the *AVSC Best Practice for Metrics and Methods for Assessing Safety Performance of Automated Driving Systems* [4]. An ADS-DV behavioral competency should be:

- Abstract - The behavioral competency can be decomposed into maneuvers, ODD elements, and OEDR elements with additional levels of specificity.
- Portable - The behavioral competency is valid for other geographic areas with similar ODD elements and parameters.
- Use case neutral - The behavioral competency is valid across use cases (e.g., rideshare, delivery).
- Relatable - Derived from general behavioral competencies required to navigate from one location to another using the existing road and infrastructure elements, and therefore relatable to human driving and understandable.
- Measurable - The behavioral competency can be quantified such that competency can be assessed against ADS developer established metrics.
- Technology neutral - The behavioral competency can be evaluated regardless of the underlying ADS hardware or software.

Several literature sources were evaluated against these criteria, including: PEGASUS [5], California Path [6], MCity [7], NHTSA AV 2.0 [8], Euro NCAP [9], NHTSA Pre-Crash Scenarios [10], Voluntary Safety Self Assessments [11], ASAM OpenX standards [12], CETRAN [13], US Department of Motor Vehicle driver license requirements [14], and existing ADS functionality whitepapers [15]. The AVSC recommended list of elemental behavioral competencies shown in Table 1 was inspired by behaviors of similar construct from these literature sources. ADS developers can select from this list according to their ODD and combine with customized behavioral competencies as appropriate.

While the previous sources provide an in-depth overview of behaviors relevant to Advanced Driver Assistance Systems (ADAS), sub-system level or even transportation-wide system-level behaviors, this best practice focuses on generalized and technology agnostic SAE level 4 and 5 behavioral competencies. The first five behavioral competencies ("Roadway Infrastructure") require formal lane structures for competent execution, which can be in the form of painted lines on the road or existing in a digital map. The "Dynamic Conditions" set represents behavioral competencies that require the ADS to respond to situational influences while already executing one or more of the five roadway infrastructure behavioral competencies.

TABLE 1 AVSC Elemental Set of Behavioral Competencies

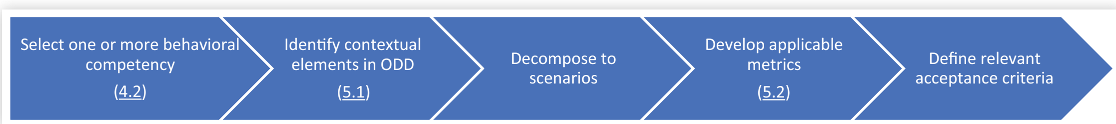
ELEMENTAL BEHAVIORAL COMPETENCIES		
	Behavior	**Specification**
Roadway Infrastructure	Maintaining a lane	Driving along roads predictably and consistently maintaining proper lane position with respect to designated lane markings and speed limits.
	Changing lanes	Lane change (right/left) to establish proper lane position in an adjacent lane, which can include merging and passing into oncoming traffic.
	Navigating intersection	Approaching, driving through, or turning at junctions adhering to traffic control devices, as defined in Manual on Uniform Traffic Control Devices (MUTCD). [16]
	Navigating, entering, exiting unstructured roadways	Approaching, driving through, or turning through roadways that do not have lane markings or clear delineations of traffic directional orientation.
	Navigating pick up and drop off zones and parking situations	Approaching, driving through, or turning to an area where parking may be restricted or prohibited to improve access for short-term curbside operations (including rideshare, airports, parking lots, parallel parking, school zones, act of stopping, VRUs in and out, and markings).
Dynamic Conditions	Responding to vulnerable road users (VRUs)	Maintaining a safety envelope [4] with respect to VRUs
	Responding to other vehicles	Maintaining a safety envelope with respect to other vehicles where another vehicle may be moving from an adjacent lane into the subject lane, ahead of the subject vehicle, either from the same direction or oncoming (e.g., leading, adjacent, encroaching, oncoming, stopped, cut-ins, cut-outs/reveal, wrong direction).
	Responding to special purpose vehicles	Where “special purpose vehicles” include emergency vehicles as defined in the Fixing America’s Surface Transportation (FAST) Act [17], government-owned vehicles, hearses, safety vehicles, school busses, etc.
	Responding to lane obstructions and obstacles	Responding to lane obstructions or obstacles can involve partial or complete lane obstructions with static or dynamic objects but is not meant to capture situations where a formal lane change is required to pass, and which is considered a complete lane blockage.
	Responding to confined road structures	Driving straight through sections of road with limited or no shoulders, potentially restricted or reduced lanes, overhead constraints, atypical reflections, and rapidly changing environmental conditions (lighting, surface conditions, etc.) from the normal roadway.
	Responding to work zones	Navigating work zones can involve detecting the work zones and temporary signage, and responding appropriately, including with respect to speeds, human traffic controllers, and navigating lane overrides or shifts.
	Responding to DDT performance-relevant failure [SAE J3016]	A failure is any malfunction that prevents the ADS from reliably performing the DDT on a sustained basis (as defined in SAE J3016).
	Responding to relevant traffic control devices	Per MUTCD (2021), “traffic control devices include all signs, signals, markings, channelizing devices, or other devices that use color, shapes, symbols, words, sounds, and/or tactile information for the primary purpose of communicating a regulatory warning, or guidance message to road users on a street, highway, pedestrian facility, bikeway, pathway or private roadway open to public travel.” [16]

Furthermore, the behavioral competencies assist with off-vehicle or enterprise-level operations, such as remote assistance in construction zones. This set is not meant to be exhaustive, but rather provides an elemental starting point for current ADS-DV testing needs. Each ADS developer and manufacturer will likely have additional behaviors requiring competency demonstration, depending on their ODD, OEDR, and maneuver specifications. All ADS developers and manufacturers can use this initial set of generalized behavioral competencies for pilot deployments as a means of demonstrating competency for the relevant behaviors for the respective ODD elements. Behaviors may only be relevant within certain ODDs, for example, navigating confined road structures is only applicable to vehicles that include confined road structures in their ODD.

5. Building a Behavioral Competency Evaluation

Evaluation of a behavioral competency involves building out the context in which the ADS-DV will be operating. As discussed in 4.2, the behavioral competency can be decomposed into maneuvers, ODD elements, and OEDR elements. Defining the ODD and OEDR elements is a key step toward defining applicable metrics. In order to build test cases with acceptance criteria for a behavioral competency, it is important to define the context for determining parameters and thresholds. Figure 3 shows a basic framework, which incorporates the context and ties together previous AVSC Best Practice Documents to support a behavioral competency evaluation.

FIGURE 3 Example Framework for Behavioral Competency Evaluation



In the following sections, two behavioral competencies are combined as an example to illustrate these steps. “Maintaining a lane while responding to a lead vehicle” is used throughout to show how elements from the ODD, applicable metrics, and metric thresholds based on context combine to support an ADS-DV safety case.

5.1. Formulation of Behavioral Competency Evaluation Context

Starting from the behavioral competency, relevant OEDR elements are dependent on the ODD specification. ODD elements can be specified using the ***AVSC Best Practice for Describing an Operational Design Domain: Conceptual Framework and Lexicon*** (***AVSC00002202004***), where ODD elements can be chosen based on their relevance to the behavioral competency which is being evaluated. Table 2 provides an example of contextual elements which may be relevant for evaluating the behavioral competency for “Maintaining a lane while responding to a lead vehicle.”

TABLE 2 Example Contextual Elements to specify for “Maintaining a lane while responding to lead vehicle”

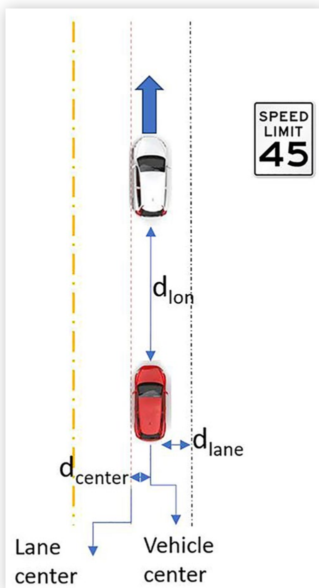
Behavior		Context
Roadway Infrastructure	Maintaining a lane	Lane marking type, width and quality
		Precipitation type and intensity
		Sky condition
		Roadway material
		Road surface condition
Dynamic Conditions	Responding to other vehicles	Presence of lead vehicle only (no cut ins, etc.)
		Size/type of lead vehicle

In some cases, the ODD may require the ADS to be able to detect and respond to other road users, unstructured obstacles, and traffic control devices. More detailed context can be specified, such that the traffic is limited to road users of specific types (i.e., light-duty vehicles, heavy-duty vehicles, and cyclists), states (heading, dynamics, lateral and longitudinal position), and maneuvering capabilities (laterally and longitudinally). Similarly for unstructured obstacles, the ODD may limit the type, dynamics, lateral and longitudinal position, and size of the obstacle [13]. **ADS developers and manufacturers should use contextual elements to help specify OEDR/ODD characteristics applicable to the behavioral competency under evaluation.** Consequently, using a behavioral competency combined with relevant ODD and OEDR contextual factors, ADS developers and manufacturers develop scenarios that are refined into test cases to evaluate ADS behavior under specific circumstances.

5.2. Application of Metrics for Behavioral Competencies

Using a behavioral competencies framework enables ADS developers and manufacturers to focus on metrics related to behaviors that are of relevance to their application. Continuing the example for “Maintaining a lane while responding to a lead vehicle”, the next step in the framework (Figure 3) is to generate applicable metrics relevant to the contextual elements identified in Table 2. To identify relevant metrics, the capabilities or goals for the behavior need to be defined. Capabilities for maintaining a lane while driving down a roadway could include that (1) the subject vehicle should maintain a safety envelope, (2) follow applicable traffic regulations, and (3) exhibit contextually safe vehicle motion control as defined in the ***AVSC Best Practice for Metrics and Methods for Assessing Safety Performance of Automated Driving Systems* (*ADS*) *for more information on metrics* (*AVSC00006202103*)**. Figure 4 provides a graphic to describe the different contextual factors involved and what the parameters specified as part of the applicable metrics represent.

FIGURE 4 Maintaining a lane while driving down a roadway pictorial example



The applicable metrics are test venue agnostic (e.g., can be applied in simulation, test track, or on-road); however, ADS developers and manufacturers may have unique acceptance criteria depending on the ODD or test configuration. These examples have been developed for ADS-DVs, and the parameters (e.g., a_{max_lon} as maximum longitudinal acceleration) are specified in the Institute of Electrical and Electronic Engineers (IEEE) draft standard from the P2846 working group specifying assumptions for models in safety-related automated vehicle behavior [18]. This example application for maintaining a lane assumes only a lead vehicle (e.g., no cut-ins and no other road actors) and a speed limit sign. Other road users and cut-in scenarios are covered in other behavioral competencies defined in Table 2.

TABLE 3 Metrics and example applications of metrics for maintaining a lane.

ADS Safety Metric	Example Application of AVSC Metrics [*AVSC6202103*] for Maintaining a lane while responding to a lead vehicle (Note that acceptance criteria are discussed in Section 5.3)
Safety Envelope	The subject vehicle keeps a minimum separation distance between the subject vehicle and the nearest in-lane leading vehicle (LV) greater than $d_{lon,min}$ (where d_{lon} is the longitudinal distance to another road user). Note that $d_{lon,min}$ is defined based on a number of kinematic variables such as speed, performance, conditions, etc.
Traffic Law - Citations[1,2]	Lane Maintenance (footnote 1): The minimum distance between the subject vehicle and the nearest point on the outer lane boundary on the immediate left or right of the subject vehicle is $\geq d_{lane,min}$. Lane boundary is defined by the lane markings along the roadway. Speed (footnote 2): The subject vehicle maintains a velocity $< v_{max}$ m/s
Contextually safe vehicle motion control	The maximum subject vehicle deviation from the centerline of the vehicle with the center of the lane within $+/-d_{center,max}$ cm. The subject vehicle has a minimum longitudinal acceleration $\geq a_{lon,min}$ m/s^2 and a maximum longitudinal acceleration of $a_{lon,max}$ m/s^2 . The subject vehicle has a minimum lateral acceleration $\geq a_{lat,min}$ m/s^2 and a maximum lateral acceleration of $a_{lat,max}$ m/s^2. The subject vehicle has a minimum longitudinal jerk $\geq j_{lon,min}$ m/s^3 and a maximum longitudinal jerk of $j_{lon,max}$ m/s^3. The subject vehicle has a minimum lateral jerk $\geq j_{lat,min}$ m/s^3 and a maximum lateral jerk of $j_{lat,max}$ m/s^3.

5.3. Defining Acceptance Criteria and ODD-Relevant Thresholds

Context for the behavioral competency evaluation goes beyond the structure of the scenario in which a metric is measured and includes the setting and relative frequency of the measurement. For example, a safety envelope violation is scenario dependent (e.g., on-road testing in denser traffic) and may not require system updates, whereas one safety envelope violation under nominal conditions (no non-compliant actors) during controlled track testing may be an indication that a change is necessary.

There are numerous approaches to define acceptance criteria for behavioral competency evaluation, and the approach may differ depending on the specific context. ADS developers may disaggregate the metrics by behavioral competency and look at each case separately and determine acceptance criteria based on a combination of tests. As a result, recommending specific values or thresholds to the parameters outlined in the examples is currently not appropriate given the significant variations in ODD and/or context. ADS manufacturers and developers should utilize the results of several metrics, evaluate them in context, and use the findings in combination with other elements of their overall safety performance evaluation (e.g., risk management plans, systems engineering verification and validation) in their deployment decisions.

ADS developers and manufacturers could use one or more of the following options to define acceptance criteria (note that this list is not exhaustive but simply provides examples of criteria):

- Apply rates to a single metric in specific applications (e.g., one safety envelope violation every 1,000 miles driven)
- Aggregate multiple metrics to evaluate holistic performance (e.g., comparing the number of safety envelope violations to traffic law citations and contextual vehicle motion control violations)
- Analyze violations individually to determine whether the ADS exhibited contextually safe motion control

[1] Example: California Vehicle Code Division 11 Chapter 3 Article 1. Driving On Right Side [Line 21658]:Whenever any roadway has been divided into two or more clearly marked lanes for traffic in one direction, the following rules apply: A vehicle shall be driven as nearly as practical entirely within a single lane and shall not be moved from the lane until such movement can be made with reasonable safety.

[2] Example: California Vehicle Code Division 11 Chapter 7 Article 1. [Line 22350]:No person shall drive a vehicle upon a highway at a speed greater than is reasonable or prudent having due regard for weather, visibility, the traffic on, and the surface and width of, the highway, and in no event at a speed which endangers the safety of persons or property.

It is important to note that thresholds may be non-zero values and that metric violations are expected when operating within the ODD. In other words, achieving a successful passing rate for a metric during controlled tests does not guarantee that the applied metric will always be met on-road (i.e., success is probabilistic, not deterministically guaranteed). In applications where trend analysis is not the appropriate acceptance criteria for a given metric, ADS developers and manufacturers may analyze these situations individually to establish contextually safe behavior or implement operational risk controls (e.g., constrain the ODD). In the example of maintaining a lane, the ADS may violate the thresholds of the metrics associated with the behavioral competency in exceptional situations such as a car door opening or encountering an unstructured obstacle along the centerline of the lane such as a hay bale falling off a lead truck.

6. Summary

Behavioral competency testing fits into a broader safety assurance framework to help build evidence towards a safety case. The breakdown of terminology consisting of behaviors, maneuvers, and OEDR clarifies how the different concepts interact and should be used in the context of behavioral competency testing. The elemental set of behavioral competencies contribute to the evaluation of aggregate safety performance of the dynamic driving task (DDT) by an ADS-DV. It is assumed that additional behavioral competencies may be applied in combination with this set to supplement evidence of safety performance at different stages of product development and deployment based on the ODD. Consistent with other AVSC best practices, this document supports industry-led, voluntary approaches in the standards development community and is expected to evolve as technology matures.

This elemental set can enhance communication of an ADS-DV safety posture among stakeholders and accelerate public acceptance. Public agencies may use this document to better understand the safety posture of ADS-DV deployments and how ADS developers and manufacturers define and evaluate applicable metrics for a behavioral competency in different test venues. In addition to the technical development community, other audiences considered in the development of this best practice include standards bodies, public agencies, and other decision-makers that may influence the deployment of ADS-DVs.

7. About Automated Vehicle Safety Consortium™

The objective of the Automated Vehicle Safety Consortium™ is to provide a safety framework around which automated vehicle technology can responsibly evolve in advance of the broad use of commercialized vehicles. The consortium will leverage the expertise of its current and future members and engage government and industry groups to establish safety principles and best practices. These technology-neutral principles are key considerations for deploying SAE level 4 and 5 automated vehicles on public roads.

AVSC Vision:

Public acceptance of SAE level 4 and 5 automated driving systems as a safe and beneficial component of transportation through industry consensus.

AVSC Mission:

The mission of the Automated Vehicle Safety Consortium™ (AVSC) is to quickly establish safety principles, common terminology, and best safety practices, leading to standards to engender public confidence in the safe operation of SAE level 4 and 5 light-duty passenger and cargo on-road vehicles ahead of their widespread deployment.

The AVSC will:

- Develop and prioritize a roadmap of pre-competitive topics;
- Establish working groups to address each of the topics;

- Engage the expertise of external stakeholders;
- Share output/information with the global community;
- Initially focus on fleet service applications.

8. Contact Information

To learn more about the Automated Vehicle Safety Consortium™, please visit https://avsc.sae-itc.org.

Contact: AVSCinfo@sae-itc.org

9. Acknowledgements

The Automated Vehicle Safety Consortium™ would like to acknowledge the contributions of the member organizations during the development of this document:

Aurora Innovations, Ford, Honda, Lyft, Motional, Toyota, and VW.

10. Abbreviations

ADS - Automated Driving System
ADS-DV - Automated Driving System - Dedicated Vehicle
ASAM - Association for Standardization of Automation and Measuring Systems
AVSC - Automated Vehicle Safety Consortium™
CETRAN - Centre of Excellence for Testing and Research of Autonomous Vehicles - NTU
DDT - Dynamic Driving Task
Euro NCAP - European New Car Assessment Programme
FAST - Fixing America's Surface Transportation
IEEE - Institute of Electrical and Electronics Engineers
IFTD - In-vehicle Fallback Test Driver
MUTCD - Manual on Uniform Traffic Control Devices
NHTSA - National Highway Traffic Safety Administration
ODD - Operational Design Domain
OEDR - Object and Event Detection and Response
PEGASUS - Project for the Establishment of Generally Accepted quality criteria, tools and methods as well as Scenarios and Situations
SAE - Society of Automotive Engineers
VRU - Vulnerable Road User

Appendix A. Best Practice Quick Look

AVSC Best Practice for Evaluation of Automated Driving System Behavioral Competencies

- Defining Behavioral Competencies ADS developers and manufacturers can use behavioral competencies as part of a safety assurance framework to generate application-specific metrics and can use the elemental set within this best practice as a starting point.
- Components of Behavioral Competencies ADS developers and manufacturers can use the recommended lexicon to enable communication of safety performance across industry in terms of behavioral competency and implement behavioral competency testing as part of a safety assurance framework to build evidence towards a safety case.
- Elemental Behavioral Competencies ADS developers and manufacturers can use the elemental set of behavioral competencies as starting points for behavioral competency testing.
- Generating Applications of Metrics for Behavioral Competencies ADS developers and manufacturers can generate application-specific metrics as a first step towards executing behavioral competency testing and to track safety performance across test venues.

Appendix B. References

1. Blumenthal M., Fraade-Blanar L., Best R. and Irwin J., “Safe Enough: Approaches to Assessing Acceptable Safety for Automated Vehicles,” The RAND Corporation, 2020. [Online]. Available: https://www.rand.org/pubs/research_reports/RRA569-1.html. [Accessed 1 April 2021].
2. Thorn E., Kimmel S. and Chaka M., “A Framework for Automated Driving System Testable Cases and Scenarios,” September 2018. [Online]. Available: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13882-automateddrivingsystems_092618_v1a_tag.pdf. [Accessed 1 April 2021].
3. Lascara B., Lacher A., DeGarmo M., Vempati L. and Zimmerman R., “Behavioral Competency Model for Safety Assurance of Automated Aviation Systems,” The Mitre Corporation, 14 June 2019. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.2019-3256. [Accessed 1 April 2021].
4. Automated Vehicle Safety Consortium, “AVSC Best Practice for Metrics and Methods for Assessing Safety Performance of Automated Driving Systems (ADS),” 2021. [Online]. Available: https://avsc.sae-itc.org/principle-6-5471WV-46540KJ.html?respondentID=30092323#our-work.
5. German Federal Ministry for Economic Affairs and Energy, “PEGASUS Method: An Overview,” 14 June 2019. [Online]. Available: https://www.pegasusprojekt.de/files/tmpl/Pegasus-Abschlussveranstaltung/PEGASUS-Gesamtmethode.pdf. [Accessed 1 April 2021].
6. University of California PATH Program, “Peer Review of Behavioral Competencies for AVs,” February 2016. [Online]. Available: https://www.nspe.org/sites/default/files/resources/pdfs/Peer-Review-Report-IntgratedV2.pdf. [Accessed 1 April 2021].
7. Peng H. and McCarthy R., “MCity ABC Test: A Concept to Assess the Safety Performance of Highly Automated Vehicles,” January 2019. [Online]. Available: https://mcity.umich.edu/wp-content/uploads/2019/01/mcity-whitepaper-ABC-test.pdf. [Accessed 1 April 2021].
8. U.S. Department of Transportation, “Automated Driving Systems - A Vision for Safety 2.0,” September 2017. [Online]. Available: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf. [Accessed 1 April 2021].
9. European New Car Assessment Program (Euro NCAP), “2018 Automated Driving Tests,” October 2018. [Online]. Available: https://www.euroncap.com/en/vehicle-safety/safety-campaigns/2018-automated-driving-tests/. [Accessed 1 April 2021].
10. Najm W., Smith J. and Yanagisawa M., “Pre-Crash Scenario Typology for Crash Avoidance Research (DOT HS 810 767),” U.S. Department of Transportation National Highway Traffic Safety Administration, April 2007. [Online]. Available: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/pre-crash_scenario_typology-final_pdf_version_5-2-07.pdf.
11. National Highway Traffic Safety Administration, “Voluntary Safety Self-Assessment,” March 2021. [Online]. Available: https://www.nhtsa.gov/automated-driving-systems/voluntary-safety-self-assessment. [Accessed 1 April 2021].
12. Association for Standardization of Automation and Measuring Systems, “ASAM e.V. Adopts Standards for Highly Automated Driving,” 10 September 2018. [Online]. Available: https://www.asam.net/news-media/press-releases/detail/news/asam-ev-adopts-standards-for-highly-automated-driving/. [Accessed 1 April 2021].
13. de Gelder E., Op den Camp O. and de Boer N., “Scenario Categories for the Assessment of Automated Vehicles,” CETRAN, 21 January 2020. [Online]. Available: http://cetran.sg/wp-content/uploads/2020/01/REP200121_Scenario_Categories_v1.7.pdf. [Accessed 1 April 2021].
14. Zakhareuski A., “DMV Driving Test in 2021: Your Complete Guide to the Road Test,” driving tests, 11 November 2020. [Online]. Available: https://driving-tests.org/road-test/. [Accessed 11 November 2020].
15. Toyota, “Toyota Automated Driving Whitepaper,” August 2020. [Online]. Available: https://automatedtoyota.com/docs/10/1a0b46b2c789165fdf296ad6d74e796f2b57837f9/ATwhitepaper.pdf.
16. “MUTCD Information Site,” Manual Uniform Traffic Control Device, 2009. [Online]. Available: https://mutcd.info/. [Accessed 1 April 2021].
17. U.S. Department of Transportation, Federal Highway Administration, “Bridges & Structures,” 3 November 2016. [Online]. Available: https://www.fhwa.dot.gov/bridge/loadrating/161103.cfm#:~:text=An%20emergency%20vehicle%20as%20defined,other%20hazardous%20situations%20(23%20U.S.C.&text=The%20gross%20vehicle%20weight%20limit,under%20section%20127(r).. [Accessed 1 April 2021].
18. IEEE P2846 Working Group, “Draft Standard for Assumptions for Models in Safety-Related Automated Vehicle Behavior,” IEEE P2846, 2021.

19. Alaska Department of Administration, Division of Motor Vehicles, “Road Test,” [Online]. Available: https://doa.alaska.gov/dmv/akol/roadtest.htm. [Accessed 24 April 2021].
20. State of California Department of Motor Vehicles, “California Driver Handbook,” 1 January 2020. [Online]. Available: https://www.dmv.ca.gov/portal/uploads/2020/06/dl600.pdf. [Accessed 1 April 2021].
21. State of Connecticut Department of Motor Vehicles, “Taking the Road Test at a DMV Office,” [Online]. Available: https://portal.ct.gov/DMV/Licenses/Licenses/Testing---Road-Test. [Accessed 14 April 2020].
22. Florida Highway Safety and Motor Vehicles, “Class E Knowledge Exam & Driving Skills Test,” [Online]. Available: https://www.flhsmv.gov/driver-licenses-id-cards/licensing-requirements-teens-graduated-driver-license-laws-driving-curfews/class-e-knowledge-exam-driving-skills-test/. [Accessed 18 March 2021].
23. Illinois Department of Motor Vehicles, “Illinois 2021 Rules of the Road,” [Online]. Available: https://www.cyberdriveillinois.com/publications/pdf_publications/dsd_a112.pdf. [Accessed 10 April 2021].
24. Maryland Department of Transportation Motor Vehicle Administration, “Driving Skills Test,” [Online]. Available: https://mva.maryland.gov/about-mva/Pages/info/26300/26300-28T.aspx. [Accessed 22 November 2020].
25. New York Department of Motor Vehicles, “Schedule and take a road test,” [Online]. Available: https://dmv.ny.gov/schedule-road-test. [Accessed 13 December 2020].
26. Ohio Bureau of Motor Vehicles, “Driver License & ID Cards,” [Online]. Available: https://www.bmv.ohio.gov/dl-gdl.aspx. [Accessed 18 December 2020].
27. Pennsylvania Department of Motor Vehicles, “Tests: Knowledge Test,” [Online]. Available: https://www.dmv.pa.gov/Driver-Services/Driver-Licensing/Driver-Manual/Chapter-1/Pages/Tests.aspx. [Accessed 15 March 2021].
28. Texas Department of Public Safety, “Schedule your Driving Test Appointment,” [Online]. Available: https://www.dps.texas.gov/section/driver-license/schedule-your-driving-test-appointment. [Accessed 26 March 2021].
29. Virginia Department of Motor Vehicles, “Road Skills Test,” [Online]. Available: https://www.dmv.virginia.gov/general/#skillstest.asp. [Accessed 1 October 2020].
30. State of Wisconsin Department of Transportation, “Class D road test study guide,” [Online]. Available: https://wisconsindot.gov/Pages/dmv/license-drvs/how-to-apply/drivingskillstest.aspx. [Accessed 27 October 2020].
31. Aurora, “Putting safety first: Aurora’s Safety Playbook,” [Online]. Available: https://aurora.tech/safety. [Accessed 30 March 2021].
32. Argo AI, [Online]. Available: https://www.argo.ai/wp-content/uploads/2021/04/ArgoSafetyReport.pdf.
33. Ford, “A Matter of Trust: Ford’s Approach to Developing Self-Driving Vehicles,” [Online]. Available: https://media.ford.com/content/dam/fordmedia/pdf/Ford_AV_LLC_FINAL_HR_2.pdf. [Accessed 12 January 2021].
34. Lyft, “Lyft Self-Driving Safety Report,” 2020. [Online]. Available: https://self-driving.lyft.com/wp-content/uploads/2020/06/Safety_Report_2020.pdf. [Accessed 13 August 2020].
35. Motional, “Safety: The Driver at Motional,” 15 January 2021. [Online]. Available: https://motional.com/news/safety-the-driver-at-motional. [Accessed 16 February 2021].
36. Toyota, “Automated Toyota,” [Online]. Available: https://automatedtoyota.com/. [Accessed 1 April 2021].
37. Waymo, “Waymo Safety Report,” September 2020. [Online]. Available: https://storage.googleapis.com/sdc-prod/v1/safety-report/2020-09-waymo-safety-report.pdf. [Accessed 25 September 2020].

Appendix C. Behavioral Competency Reference Table

Table 4 represents references considered by AVSC in creation of the elemental list recommended for ADS developers and manufacturers in behavioral competency evaluation. These references show similar constructs yet should not be considered as equivalent to the definitions presented in Table 1.

TABLE 4 Behavioral Competency Reference Comparison

Behavior	NHTSA [2]	VSSA reports	Cal PATH [6]	DOT driver tests	CETRAN [13]	MCity [7]
Maintaining a lane	X	X	X	X	X	X
Changing lanes	X	X	X	X	X	X
Navigating intersections	X	X	X	X	X	X
Navigating unstructured roadways, entering/exiting unstructured roadways	X	X	X	X	X	X
Navigating pick up and drop off zones and parking structures	X	X				X
Responding to VRU	X	X	X		X	X
Responding to other vehicles	X	X	X	X	X	X
Responding to special purpose vehicles	X	X	X			X
Responding to lane obstructions and obstacles (*static, dynamic, including animals*)	X	X	X		X	X
Responding to confined road structures (*tunnels, bridges, parking garages, etc.*)	X	X			X	
Responding to work zones (*including workers*)	X	X	X		X	X
Responding to DDT performance-relevant system failure	X	X	X			X
Responding to relevant dynamic traffic signs and signals	X	X			X	

The "**DOT driver tests**" column was developed using a subset of states driver tests from the following states:

- Alaska [19]
- California [20]
- Connecticut [21]
- Florida [22]
- Illinois [23]
- Maryland [24]
- New York [25]
- Ohio [26]
- Pennsylvania [27]
- Texas [28]
- Virginia [29]
- Wisconsin [30]

The **"VSSA reports"** column is composed of a combination of Voluntary Safety Self-Assessment reports from the following entities:

- Aurora [31]
- Argo [32]
- Ford [33]
- Lyft [34]
- Motional [35]
- Toyota [36]
- Waymo [37]