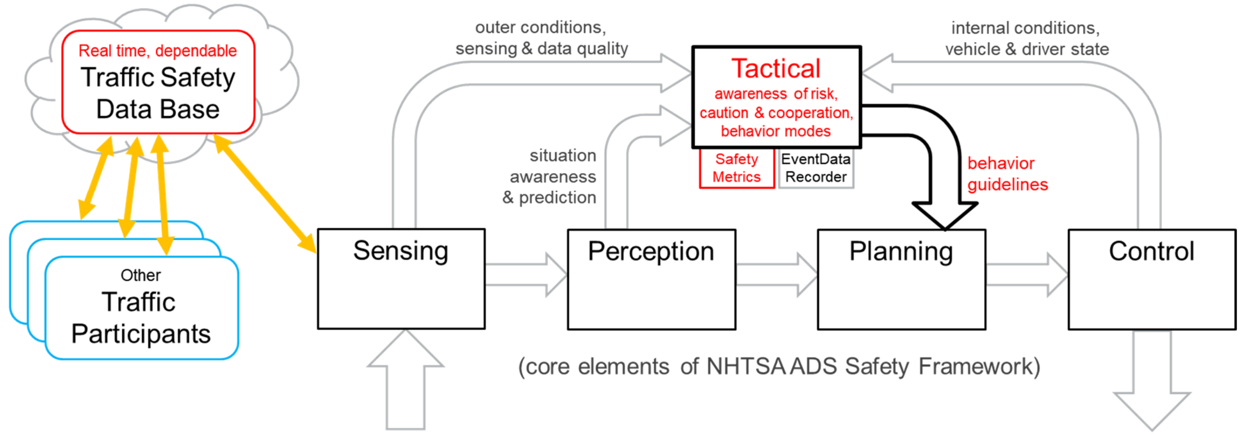
SUMMARY

This document comments the proposed Safety Framework for Automated Driving Systems. The main topics dicussed in this document are:

a) consideration of implicit tasks of a human driver;
b) consideration of risks which arise from traditional vehicle system specifications, which are based on the assumption that a human driver exists;
c) the need for a **quantitative measure for ‚safe driving'** as a general guideline, best as an internationally harmonized regulation;
d) the proposal of **a fith core element called ‚tactical behavior module'**, as a supervisor of autonomous driving tasks, generating a tactical behavior guideline to the ADS;
e) a proposal of a **regulation for a ‚Real-time Traffic Safety Data Base'**, as a prerequisite for autonomous high speed highway driving systems with specific properties;
f) a link to the attached document Schöner 2020: **„Challenging Highway Scenarios Beyond Collision Avoidance for Autonomous Vehicle Certification"**, with scenarios derived from a risk-based identification of relevant highway scenarios. These scenarios are proposed as essential scenarios for ADS certification.

Proposed new elements and new regulations of an ADS Safety Framework



ATTACHMENTS

- Mattas K, e.al. 2020. Fuzzy Surrogate Safety Metrics for real-time assessment of rearend collision risk. In: Accident Analysis and Prevention 148 (2020)
- Schöner HP, e.al. 2021. A Safety Score for the Assessment of Driving Style. Preprint; accepted for publication in: Traffic Injury Prevention Journal (2021)
- Schöner HP. 2020. Challenging Highway Scenarios Beyond Collision Avoidance for Autonomous Vehicle Certification. In: Research Gate, DOI: 10.13140/RG.2.2.29355.05926

COMMENTS ON SPECIFIC QUESTIONS

blue text is copied from the requesting document federalregister.gov/d/2020-25930

Questions and Requests

A. Questions about a Safety Framework

• **Question 1.** Describe your conception of a Federal safety framework for ADS that encompasses the process and engineering measures described in this document and explain your rationale for its design.

The term „Automated Driving Systems“ (ADS) encompasses a very wide field. System safety depends very much on the application and normally cannot be ensured by just one technology; normally it is only possible by combining a wide variety of approaches matched to the application. Thus, in order to cover safety of ADS thoroughly it should be broken down into **different application fields**.

Taking the responsible driver out of the game, the system must accomplish many safety tasks which traditionally have been ***implicit*** **tasks of the driver**. The amount of such tasks is completely different for a SAE level 3 system (in which a responsible driver is still available for some ‚high level‘ safety aspects) or for a level 5 system (in which the system needs to take over much wider responsibilities). Such high level aspects include:

- On the upper end: Ensure that the system (including all of its components) is capable of performing the driving task under the prevailing conditions of: the vehicle state, the traffic, the weather, the road conditions, and much more. This task is very different for an ADS which is a) only activated during portions of a highway drive, or b) for a parking procedure in a confined parking lot, or c) during a 24h per day / 7 days per week taxi fleet service without responsible operator close by (as three examples). And this task has to be done not only at the activation of the autonomous state, but continuously during the autonomous driving task. It includes checking whether the conditions for the ODD are still valid, but it is much more than just checking whether the vehicle is situated on a highway or in a parking lot.
- On the lower end: There is a myriad of little responsibilities of the driver, which can be seen in the owner’s manual of a conventional vehicle. Basicly every sentence in the manual which starts with: ‚caution‘, ‚danger‘, or ‚risk of‘ might fall into the safety responibility of the ADS instead of the driver. In order to take safety seriously, for every of these implicit tasks there should be a procedure to make sure that the ADS (or its accompanying support systems) does the same good job as a human driver. If an automous vehicle is operated in a fleet with trained personel such responsibilities can be taken over in a very different way than in a privately owned vehicle.

Exhaustive safety assessment must be very different for different ADS implementations. It seems that a lot of the procedures and measures described in the presently proposed framework is focussing on the pure driving taks, replacing the steering wheel and pedal actions and some of the traffic- and environment-oriented perception, planning and control actions of the driver. Although these are essential for everday operation of the vehicle, they are not providing the complete set of necessary skills of a ‚driver‘. This might only become evident in rare situations (including failures, coincidents of several difficult conditions, etc.), but under the assumption that autonomous vehicles will represent a growing number with a substantial presence in traffic, those rare situations will become relevant for safety. Those aspects are – to my perception – too little covered in the safety framework under discussion.

- **Question 2.** In consideration of optimum use of NHTSA's resources, on which aspects of a manufacturer's comprehensive demonstration of the safety of its ADS should the Agency place a priority and focus its monitoring and safety oversight efforts and why?

- **Question 3.** How would your conception of such a framework ensure that manufacturers assess and assure each core element of safety effectively?

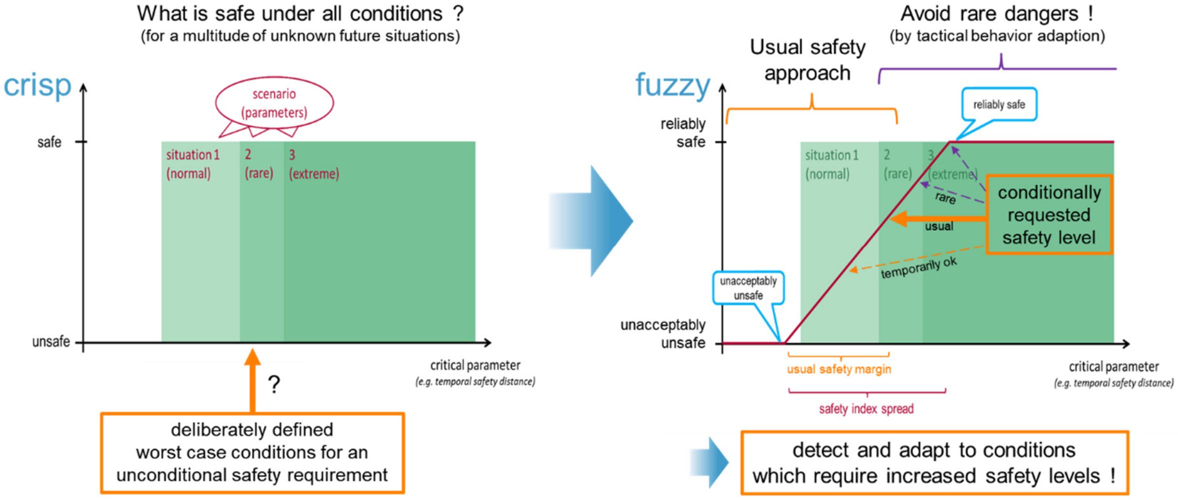
Manufacturers have to make sure that not only the everyday driving skills are met by the ADS, but also a large set of the implicit skills of a human driver. At least a checklist of such skills should be established and published; this list might be quite diverse for different ADS applications.
Very significant risks might be hidden in traditional vehicle system specifications, which are based on the assumption that a human driver exists in the system: State of the art brake systems and steering systems are designed and certified with the assumption that a human driver can and will apply his body force in case of a failure of brake or steering assistance. At least there should be evidence that those safety critical systems have enough redundancy to support the cases in which the driver's force is the last resort in traditional vehicle design. Furthermore, any system designed according to ISO26262 needs a review whether its design is still valid without the driver as a final safety backup. One essential subsystem is the vehicle's electrical power supply system: without well designed redundancy this is a possible source for a single failure to lead to a catastrophic functional collapse of the ADS: a human driver can bring a conventional vehicle with faulty electric power supply to a safe stop; an autonomous vehicle based on a conventional electrical power supply is typically not able to handle this case.
Such requirements might shed a quite new light on the concept of a ‚retrofit ADS' implemented into a conventional vehicle: in my view this is only admissible for a small number of experimental vehicles with additional safety provisions, but definitively not certifiable as a general concept for a large number of different conventional production vehicles.

- **Question 4.** How would your framework assist NHTSA in engaging with ADS development in a manner that helps address safety, but without unnecessarily hampering innovation?

- **Question 5.** How could the Agency best assess whether each manufacturer had adequately demonstrated the extent of its ADS' ability to meet each prioritized element of safety?

A quantitative measure for ‚safe driving' as a general guideline for every traffic participant is missing and needs to be established, best with an international consensus:
For novice human drivers all over the world, the concept of driving examinations has been installed; candidates do not show their emergency braking or evasive steering capabilities, but they show their ability to look ahead, to anticipate dangers and how they manage to keep well clear of collisions with flexible situation-aware safety margins (leaving enough space and time for others to act and respond as well). Such smooth tactical driving skills are more important for safe driving than last second collision avoidance skills. After all, the examinators consider the resulting spatial and timewise safety margins in traffic situations, in combination with the driver's responsiveness and mitigation actions in order to assess mature skills for safe driving. A similar concept needs to be established for autonomous vehicles.
A key point seems to be an objective quantitative measure, which is able to measure the safety margins of a driving style. If such a measure can be established, the average margin against definitively dangerous situations (collisions as an unacceptable limit) can be monitored. Requiring minimum margins (in a distance and time scale) for specific driving situations (larger

for common situations, smaller in rare and emergency situations) can ensure a collision-free driving style with a high probability. In conjunction with monitoring of reaction dynamics and proactive actions, safe handling of situations can be assessed. Used as a control variable, a safety margin can serve as a feed-back value for learning safe behavior with chances for AI to improve the driving skills over time (similar to the learning process of a human driver who has awareness of more or less dangerous situations).


What is safe under all conditions ?
(for a multitude of unknown future situations)
crisp
safe
unsafe
scenario (parameters)
situation 1 (normal)
2 (rare)
3 (extreme)
critical parameter (e.g. temporal safety distance)
?
deliberately defined worst case conditions for an unconditional safety requirement
Avoid rare dangers !
(by tactical behavior adaption)
Usual safety approach
fuzzy
reliably safe
unacceptably unsafe
reliably safe
situation 1 (normal)
2 (rare)
3 (extreme)
rare
usual
conditionally requested safety level
temporarily ok
unacceptably unsafe
usual safety margin
safety index spread
critical parameter (e.g. temporal safety distance)
detect and adapt to conditions which require increased safety levels !

The traditional crisp (binary) safety concept, which just discriminates between ‚safe‘ and ‚unsafe‘ situations, is not suited for providing quantitative safety margins. This concept has been used successfully to decide, whether an emergency driver assistance system needs to intervene or not in a collision-prone situation (after the vehicle has been already conducted into a dangerous situation), but it is no more suited as guideline for a safe continuous driving style. A fuzzy safety score which covers a span between ‚reliably safe‘ and ‚unacceptably unsafe‘ can provide a guideline for a vehicle path planner and controller, as well as a measure for an examinator. For ‚usual‘ situations, a certain safety level can be defined as standard requirement (together with the limit conditions for ‚usual‘). Higher acceptable safety levels under more challenging weather, road or traffic conditions can be requested, requiring that these conditions must be identified by the ADS and lead to tactical behavior changes. The concept of fuzzy (continuous) safety metrics using margins against an ‚unacceptably unsafe‘ condition is also forgiving for misjudgements or unexpected changes of the conditions (the ‚unknown‘ of the SOTIF requirements); on the other hand it allows temporarily lower safety levels for merging situations, for example.
Examples for experimental continuous safety metrics can be found in recent literature (see Mattas e.al. 2020, Schöner e.al. 2021). The topic still needs further research and international harmonization.

• **Question 6.** Do you agree or disagree with the core elements (i.e., “sensing,” “perception,” “planning” and “control”) described in this document? Please explain why.
I agree, that the four core elements are fine for the operational part of the driving task. But as discussed in the answer to question 1, an additional **supervising core element** is missing, **generating a tactical behavior guideline to the ADS**. More details of this concept are discussed under question 7 and 8.
Besides this, for the design of the four operational core elements it must be required, that the complete contribution of a human driver has to be analyzed carefully for each part. This needs to be pointed out more than what has been mentioned in the existing safety framework document. This includes (as some examples): does the sensing system recognize emergency vehicles good

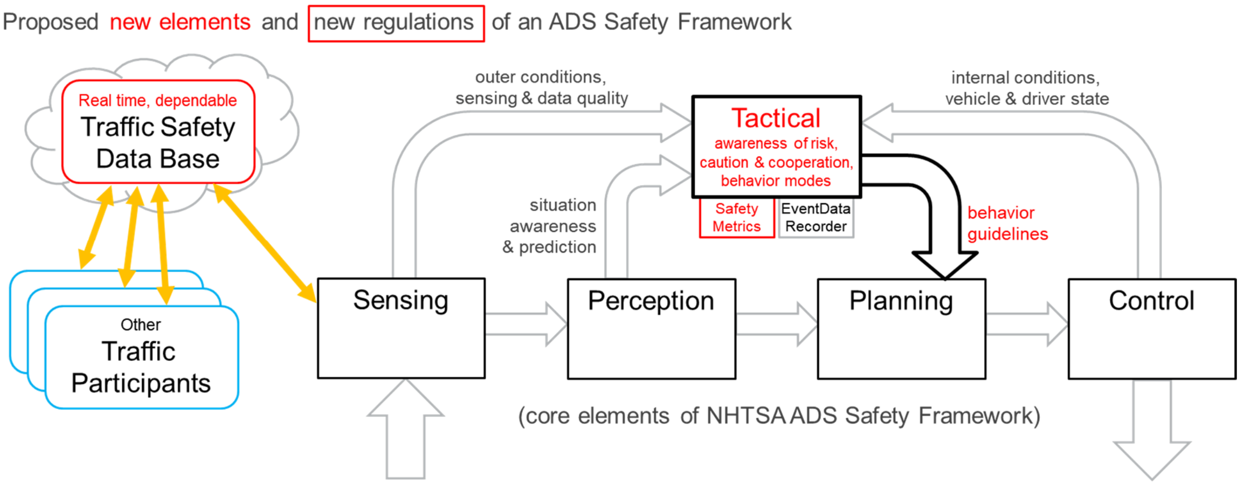
enough (for human drivers they issue an auditory signal as long distance warning; is there an adequate replacement for this safety feature?), does the perception understand warning signals of human first responders or traffic police signals (they are not precisely defined, but evident to a human driver), does the planning understand situations which need cooperation between traffic participants, does the control consider failures or reduced performance of the actuators?

- **Question 7.** Can you suggest any other core element(s) that NHTSA should consider in developing a safety framework for ADS? Please provide the basis of your suggestion.

An additional supervising core element should be considered (see also figure to question 8); it could be called the "**tactical**" **behavior module**.
The role of this additional core element would be to provide an explicit ‚self awareness' of the complete ADS with respect to the implicit responsibilites of the driver like: assessment of the integrity of the system, judgement of the prevailing conditions, overall risk assessment and a conscience with respect to safety of the overall system. The ‚event data recorder' would be one component of this supervising core element, providing the necessary data needed and used for tactical safety decisions.
In cooperation with the other four elements the supervising element needs to conclude, whether the prevailing sensing horizon is sufficiently safe, what kind of system limitations need to be taken into account as guidelines for planning, and whether the conditions for the ODD are met.
The concept of ‚Tactical Safety' (as labeled by CertiCAV and discussed in Schöner 2020) guides behavior changes in order to avoid coming into possibly dangerous situations (changing behavior early and smoothly based on subtle indicators); a tactical behavior guideline would be the output of the supervising element to the other operational elements. This includes the **activation of a more ‚cautious' driving style** (e.g increase safety distances and/or reduce speed depending on unusual prevailing conditions), **emergency modes** (e.g. provide an emergency vehicle corridor) **or ‚limb-home' mode** (e.g. find the next location for safe stopping).
Although many of the tasks of such a tactical behavior module might also be distributed and implicitly included in the other core elements, the explicit allocation of these tasks in this separate module makes the safety relevant functions of this block testable and available to certification audits, independent of complex technology which resides in the other four core elements. The functions and performance of the tactical module, together with clear safety metrics and the measured data of relevant events, makes it possible to assess whether the ADS is aware of risks, how it implements cautious and safe behavior dependent on conditions, and whether the activation of e.g. cautious, emergency or limb-home modes are adequate. With its **ability to respond** to outer and inner conditions, it is the central component for verifying **responsible behavior** of the ADS.

- **Question 8.** At this early point in the development of ADS, how should NHTSA determine whether regulation is actually needed versus theoretically desirable? Can it be done effectively at this early stage and would it yield a safety outcome outweighing the associated risk of delaying or distorting paths of technological development in ways that might result in forgone safety benefits and/or increased costs?

Regulation is needed to establish a common safety basis for all traffic participants. Especially rules, which a single vehicle cannot define, or safety provisions, which a single vehicle would not be able to establish, have to be implemented by a regulation.
I see two specific fields which need, or take significant profit from, regulation:


Proposed new elements and new regulations of an ADS Safety Framework
Real time, dependable
Traffic Safety Data Base
Other Traffic Participants
outer conditions, sensing & data quality
situation awareness & prediction
Tactical
awareness of risk, caution & cooperation, behavior modes
Safety Metrics
EventData Recorder
internal conditions, vehicle & driver state
behavior guidelines
Sensing
Perception
Planning
Control
(core elements of NHTSA ADS Safety Framework)

1. Establish **commonly accepted safety goals based on quantitative measures** which are more useful as a behavior guideline than a binary safe / unsafe definition (which is oriented at collisions). It should quantify the fuzzy notion of ‚endangering‘ and ‚risky‘ behavior and set guidelines for ‚nominal safe‘ behavior, based on (spatial and temporal) safety margins to the crisp collision avoidance in several very specific situations. It should take into account that behavior can be more or less safe on a quantitative scale, and that a lesser safety value is acceptable in certain situations (for a limited time, and with additional care) in order to solve specific traffic flow requirements (like temporarily reduced safety distances during merging situations). On the other hand, more challenging conditions lead to higher safety requirements. Such a quantitative measure should provide the common basis for tactical safety behavior of all traffic participants (see question 5).
2. Establish a guideline for a kind of ‚**Real-time Traffic Safety Data Base**‘, with trustable and time stamped information on actual traffic, weather and road conditions. This data base should be the equivalent to ‚traffic warnings‘ distributed by radio channels to human drivers. But it should be more precise, trustable and with information about completeness and actuality of the information; it should be readable by autonomous vehicles and serve as an input to the ‚tactical behavior module‘ of an ADS (and would as well be usable for tactical safety ADAS systems). The guideline should require such a data base **at least for autonomous high speed highway driving** systems, in which a vehicle-based sensor set cannot guarantee to provide the complete picture of the road ahead (even best sensors cannot see through other vehicles). The exact specification for the data content can be left open for technological development. Ideally, the data collection may be ‚crowd sourced‘ and increase functionality over time. But it would make a big safety step from ‚having a rough idea of the conditions on the road ahead‘ to ‚knowing precicesly what to expect on the road ahead‘. A possible starting concept which collects and provides the driving speed (per lane) on the road ahead is proposed in Schöner (2020, annex E); it fulfills the completeness and actuality requirement with relatively little bandwidth and reality-to-database latency requirements.

Both regulations would improve safety consistently for autonomous vehicles; besides this they would help to improve safety of human driven vehicles as well. Both do not need to have *driverless* vehicles on the road to implement the regulation, they can be introduced in human driven vehicles. The Real-time Traffic Safety Data Base is based on the same technological concepts as HD-maps including hazard warning systems, being in preparation for autonomous vehicles. Map providers can easily expand such systems to provide information about the completeness and actuality of the data along the road, so it can be used for safety purposes. The cooperation of all vehicle (and especially truck) manufacturers is important for a fast introduction. A regulation might be needed for a fast introduction of this safety feature, since a single manufacturer might be hesitating to set the necessary standard without a regulation.

- **Question 9.** If NHTSA were to develop standards before an ADS-equipped vehicle or an ADS that the Agency could test is widely available, how could NHTSA validate the appropriateness of its standards? How would such a standard impact future ADS development and design? How would such standards be consistent with NHTSA's legal obligations?

Both proposed regulations from question 8 do not need to have *driverless* vehicles on the road to implement the regulation, they can be introduced and validated in human driven vehicles. The Real-time Traffic Safety Data Base is based on the same technological concepts as HD-maps including hazard warning systems, being in preparation for autonomous vehicles. Map providers can easily expand such systems to provide information about the completeness and actuality of the data along the road, so it can be assessed for trustability and used for safety purposes. The cooperation of all vehicle (and especially truck) manufacturers is important for a fast introduction, which can be heavily supported by a NHTSA regulation. Without such a regulation vehicle owners might not be interested to report their own speed, but the speed profile along the road is a key to safe tactical behavior.
In the 21st century as an era of communication systems it seems to be adequate to include communication in a safety concept. Leaving communication out of the safety concept for autonomous vehicles on high speed highways could be considered as an omission of due diligence.

- **Question 10.** Which safety standards would be considered the most effective as improving safety and consumer confidence and should therefore be given priority over other possible standards? What about other administrative mechanisms available to NHTSA?

See answer on question 8.

- **Question 11.** What rule-based and statistical methodologies are best suited for assessing the extent to which an ADS meets the core functions of ADS safety performance? Please explain the basis for your answers. Rule-based assessment involves the definition of a comprehensive set of rules that define precisely what it means to function safely, and which vehicles can be empirically tested against. Statistical approaches track the performance of vehicles over millions of miles of real-world operation and calculate their probability of safe operation as an extrapolation of their observed frequency of safety violations. If there are other types of methodologies that would be suitable, please identify and discuss them. Please explain the basis for your answers.

In Schöner (2021) a concept of deriving acceptable safety distances and reaction times from human behavior is proposed and discussed. Safe operation of a vehicle can be done by temporarily going below safety margins (for a good reason), but it should not go often under a threshold of ‚risky'. The concept of tactical behavior change, which adapts safety margins according to prevailing conditions, seems to be more suitable (and closer to human behavior) than requiring a fixed, unconditional safety margin. Nevertheless, such general tactical safety rules have to be established (see question 5).

- **Question 12.** What types and quanta of evidence would be necessary for reliable demonstrations of the level of performance achieved for the core elements of ADS safety performance?

- **Question 13.** What types and amount of argumentation would be necessary for reliable and persuasive demonstrations of the level of performance achieved for the core functions of ADS safety performance?

B. Question About NHTSA Research

- **Question 14.** What additional research would best support the creation of a safety framework? In what sequence should the additional research be conducted and why? What tools are necessary to perform such research?

See answers to question 5 and 8:
A quantitative measure for ‚safe driving' as a general guideline for every traffic participant is missing and needs to be established, best with an international consensus. Experimental fuzzy (continuous) safety metrics can be found in recent literature (see Mattas e.al. 2020, Schöner e.al. 2021). The topic still needs further research and would profit extremely from international harmonization.
A guideline for a kind of ‚**Real-time Traffic Safety Data Base**', with trustable, complete and time stamped information on actual traffic, weather and road conditions, is discussed in the answer to question 8. The core of such a data base are less dynamic states like speed profiles along the road and the location of hazard spots. How this can be extended in order to predict the position of an oncoming vehicle in the wrong lane on a highway, or to implement more dynamic data like emergency braking signals should be evaluated in further research.

C. Questions About Administrative Mechanisms

- **Question 15.** Discuss the administrative mechanisms described in this document in terms of how well they meet the selection criteria in this document.

As discussed with questions 1 and 3, the **implicit tasks** of human drivers should be considered carefully and translated into explicit tasks for autonomous vehicles.
Very significant risks might be hidden in traditional vehicle system specifications, which are based on the **assumption that a human driver exists** in the system.
These risks needs to be identified and at least required from a manufacturer to provide evidence in a safety audit how such human tasks are covered by the autonomous vehicle.

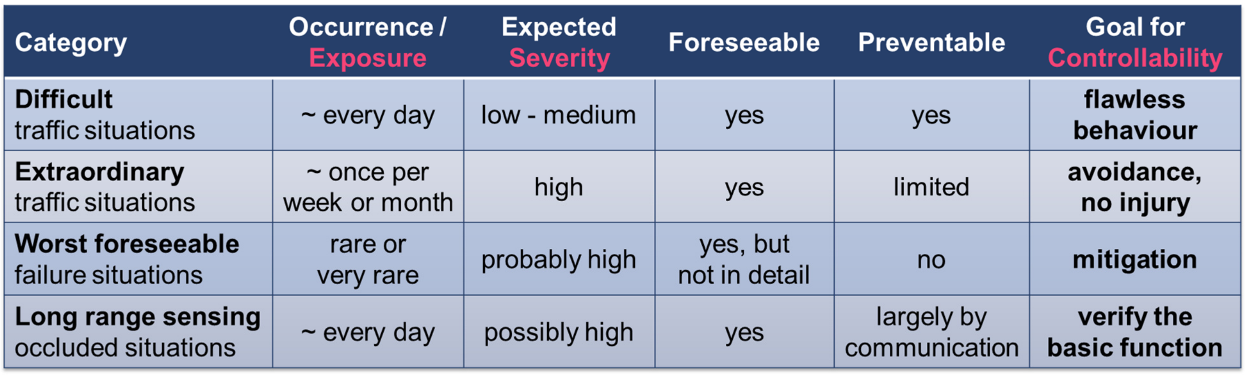
- **Question 16.** Of the administrative mechanisms described in this document, which single mechanism or combination of mechanisms would best enable the Agency to carry out its safety mission, and why? If you believe that any of the mechanisms described in this document should not be considered, please explain why.

- **Question 17.** Which mechanisms could be implemented in the near term or are the easiest and quickest to implement, and why?

- **Question 18.** Which mechanisms might not be implementable until the mid or long term but might be a logical next step to those mechanisms that could be implemented in the near term, and why?

- **Question 19.** What additional mechanisms should be considered, and why?

In the document „**Challenging Highway Scenarios Beyond Collision Avoidance for Autonomous Vehicle Certification**“ (Schöner 2020) a set of some important scenarios is presented with the goal to cover the most important aspects of precautious and cooperative driving on highways. The discussion includes proposals for further specification of parameters for testing purposes.
All of these 16 scenarios have different reasons for being in the scenario set; please refer to this document for the details and for the argumentation of including them (in the annexes of the document).
The collection includes the following highway scenario categories:



Category	Occurrence / Exposure	Expected Severity	Foreseeable	Preventable	Goal for Controllability
Difficult traffic situations	~ every day	low - medium	yes	yes	**flawless behaviour**
Extraordinary traffic situations	~ once per week or month	high	yes	limited	**avoidance, no injury**
Worst foreseeable failure situations	rare or very rare	probably high	yes, but not in detail	no	**mitigation**
Long range sensing occluded situations	~ every day	possibly high	yes	largely by communication	**verify the basic function**

- **Difficult traffic situations**
  These are forseeable traffic situations with some challenges for driving skills, which happen every day; cooperative and predictive behavior helps in reducing the occurrence of danger and collisions out of these situations. The passing criteria for those situations are: flawless behavior, i.e. do not endanger yourself & others and enable smooth traffic flow.
- **Extraordinary traffic situations**
  These are forseeable traffic situations, which might happen once per week or less in the traffic statistics of a region. They are unusual for a single vehicle, but need to be considered for the traffic community in total. The passing criteria for those situations, depending on specifc parameters, are: no collision / at least no injury.
- **Worst case failure situations**
  These failure situations are rare, but foreseeable situations. Because of their probably severe effects, they have to be considered in a safety assessment. The passing criteria for these situations are: show a reasonable mitigation effort, avoiding severe collisions.
- **Long range sensing situations**
  These situations are needed for high speed driving on highways with a safety relevant sensing horizon beyond the practical limits of the own sensors; risk reduction is based on use of communication and cooperative vehicle interaction. Passing criteria here are: showing evidence of basic cooperative capabilities in traffic, suitable supervision of these capabilities and reasonable reaction in case of faults in the cooperative system.

In several annexes the document contains reasoning based on a general risk model and specific calculations which explain the inclusion of the specific scenarios in the set. Besides this, a proposal for a cooperative sensing system (as an approach for the ‚**Real-time Traffic Safety Data Base**‘ in question 8) is included.

• **Question 20.** What are the pros and cons of incorporating the elements of the framework in new FMVSS or alternative compliance pathways?

- **Question 21.** Should NHTSA consider an alternative regulatory path, with a parallel path for compliance verification testing, that could allow for flexible demonstrations of competence with respect to the core functions of ADS safety performance? If so, what are the pros and cons of such alternative regulatory path? What are the pros and cons of an alternative pathway that would allow a vehicle to comply with either applicable FMVSS or with novel demonstrations, or a combination of both, as is appropriate for the vehicle design and its intended operation? Under what authority could such an approach be developed?

D. Questions About Statutory Authority

- **Question 22.** Discuss how each element of the framework would interact with NHTSA's rulemaking, enforcement, and other authority under the Vehicle Safety Act.

- **Question 23.** Discuss how each element of the framework would interact with Department of Transportation Rules concerning rulemaking, enforcement, and guidance.

- **Question 25.** If you believe that any of the administrative mechanisms described in this document falls outside the Agency's existing rulemaking or enforcement authority under the Vehicle Safety Act or Department of Transportation regulations, please explain the reasons for that belief.

- **Question 24.** If your comment supports the Agency taking actions that you believe may fall outside its existing rulemaking or enforcement authority, please explain your reasons for that belief and describe what additional authority might be needed.