Advances in Reliability, Safety and Security, Part 5 – Kołowrocki, Kosmowski (eds.)

Advances in Reliability, Safety and Security

ESREL 2024
Monograph Book Series

Exploring Safety-Related Metrics To Assess Human-System Interactions In Heavy-Duty Automated Vehicles

Anna Cosmin-Spanoche[a,b], Camila Correa-Juillian[a,b],
Xin Xia[c], Ali Mosleh[b], Jiaqi Ma[c,d]

[a]*Department of Mechanical and Aerospace Engineering, University of California, Los Angeles, USA*
[b]*B. John Garrick Institute for the Risk Sciences, University of California, Los Angeles, USA*
[c]*Mobility Lab, University of California, Los Angeles, USA*
[d]*Department of Civil and Environmental Engineering, University of California, Los Angeles, USA*

Abstract

Automated driving technologies are becoming increasingly common across various applications in the transportation industry. In recent years, there has been a growing interest in expanding these applications towards commercial heavy-duty operations, aiming to increase operational hours and reduce fatal collisions. Currently, multiple companies are involved in the development, testing, and small-scale deployment of heavy-duty automated vehicle (HD-AV) systems. With the emergence of new Automated Driving System (ADS) technology additional risks are introduced to commercial fleet operations. Currently, HD-AV fleet operations are planned as a joint effort of multiple human and machine agents, including an onboard safety driver and a fleet operations center. HD-AV operations can potentially cover a range of applications, including middle-mile, drayage, long-haul, etc. each with distinct operational profiles and safety requirements. In each of these, the interactions between agents contribute to the complexity of the operations and the design of safety requirements. Most notable among these are the interactions between the safety driver and the ADS, and these interactions must be modelled to construct an in-depth safety analysis. This work presents a discussion of current ADS and human-related safety metrics and suggests potential metrics that can build upon these to assess human-ADS interactions.

Keywords: automated driving systems, human-system interactions, safety metrics

1. Heavy-duty vehicles and automated driving systems

Automated driving technology is becoming more common in various transportation applications worldwide, including passenger "robotaxi" transport, driver support features integrated in passenger vehicles, and commercial transport applications. This wide range of potential large-scale applications have underscored the importance of adequately identifying, characterizing, and estimating the risks introduced by this new technology in the transportation environment. The degree of driving automation across applications is varied, depending on the Dynamic Driving Task (DDT) division and allocation between human and machine agents. The Society of Automotive Engineers (SAE) classifies the extent of a vehicle's automation into six levels, ranging from 0-5. Levels 0-2 refer to driving assistance features, where the human driver remains in charge of the DDT. Levels 3-5 gradually transfer the DDTs to the ADS (SAE International, 2021). The ADS agent is referred to as "the hardware and software that are collectively capable of performing the entire DDT on a sustained basis, regardless of whether it is limited to a specific Operational Design Domain (ODD)". Distinctions between Level 3 and Level 4 reside in that in the former the human driver is still responsible for acting as a fallback-ready user. Similarly, Level 4 and Level 5 differ in that the operation – including DDT fallback – is not restricted to specific ODD requirements.

The development of HD-AVs is a rapidly growing field, with over 20 unique companies in various stages of innovation internationally. Heavy-duty vehicles are defined as Class 7-8 vehicles, which are vehicles weighing over 26,001 pounds (11,794 kg). The classes are denoted by the limitations of Class 7 being vehicles between

26,001 and 33,000 pounds (11,794 to 14,969 kg), and Class 8 being vehicles over 33,000 pounds (14,969 kg). The functions that heavy-duty vehicles perform are varied, ranging from construction vehicles (such as dump trucks and cement mixers), garbage trucks, and large transit buses. However, their most common use is in commercial transport operations, which can be either short-haul or long-haul. These are further categorized into drayage, freight, less-than-truckload, and intermodal trucking, among other distinctions.

Motivations for the development of ADS include the potential reduction of collisions caused by human error, an increase in commercial transport efficiency, and an eventual progression toward 24/7 operations. A common statistic from a National Highway Traffic Safety (NHTSA) report states that 94% of collisions can be attributed to human error (Singh, 2015). In addition, various media sources have inferred that replacing humans with automated driving technologies would reduce over 90% of crashes (Bonnefon, 2021). However, this report notes that this estimate does not imply that the drivers are at fault in all the associated collision scenarios (Zhai et al., 2023). A more realistic estimate of crash reduction when replacing humans with automated driving technologies may, in fact, be closer to 50% (Shetty et al., 2022). Nevertheless, in the near future, humans will continue to interact with ADS systems as drivers, passengers, or supervisors, inside and outside of the vehicle.

In the United States, the safe operation of heavy-duty vehicles is principally overseen by NHTSA, who publish yearly safety reports detailing crash and fatality statistics. In 2021, there were almost 14 million medium- and heavy-duty vehicles registered. In this year, “large trucks accounted for 9% of all vehicles involved in fatal traffic crashes” despite only making up 5% of all registered vehicles that year. In addition, 72% of fatal crashes involving a large truck resulted in an injury or fatality to occupants of another vehicle, compared to 11% being occupants of the truck. (National Highway Traffic Safety Administration, 2023). To increase reliability and reduce incidents while also increasing operational coverage, many heavy-duty operators are exploring ADS as an option. However, efforts must be directed towards understanding and assessing the new risks that arise by increasing the number of heavy-duty vehicles on the road, expanding their operational hours, and using new technologies.

Currently, there are multiple prospective models of HD-AV operations, with levels of driving automation ranging from 2-4 depending on the specific application. Many of these contemplate interactions with human operators and safety drivers, either monitoring or completing sections of the drive. For instance, shorter-term applications such as “middle-mile” operations consider safety drivers to complete sections of the driving operation prior to entering and after exiting the ODD. Additional to the safety driver, these commercial operation contexts may also consider the active participation of a remote operations center, and, in some cases, an additional onboard safety operator. These human operators not only are expected to interact with the HD-AV during planned sections of the operation, but also to collaborate with the vehicle in emergency situations, manage unplanned control transitions (takeovers and handovers), and conduct post-incident procedures. Therefore, it is of interest to analyze the safety implications of the interaction of these human agents with the ADS-equipped heavy-duty vehicles. To support claims that HD-AV systems are safe, data on these systems needs to be collected, and a set of universal metrics needs to be defined to conduct fair comparisons and safety assessment.

Most metrics proposed to assess the performance of an ADS focus on the system itself and its interaction with other road users, such as Time-to-Collision and Aggressive Driving (Wishart et al., 2020). However, these do not consider the presence of a safety driver and/or operator and the interactions that take place between these human operators and the ADS within their vehicle. In addition, there has been research done with regards to takeover requests (TORs) and the metrics associated with them. These studies have analyzed characteristics like attention, fatigue, and reaction time, which are all relevant for HD-AV systems. Other studies have focused on emergency situations, and their associated task load and complexity of take-overs. When providing a full view of operational scenarios, modelling emergency situations is also important and highlights the criticality of safety metrics. To develop effective metrics to assess interactions between humans and the ADS system, analysts and developers should rely on a combination of quantitative and qualitative metrics. For instance, metrics assessing TORs are based on not only the operator’s reaction time, but also the quality of the takeover. They are used to identify system and operation design elements that hinder or enhance the performance of the safety driver and the ADS (DeGuzman et al., 2021; Morales-Alvarez et al., 2020; Zhang et al., 2019).

The aim of this work is to explore potential human-system interaction and collaboration metrics in the context of HD-AV operations. This work is structured as follows: a general reference fleet that covers most commercial HD-AV operations is developed, forming a system breakdown and definition of operational phases based on system interactions. This allows for analyzing the human-system interactions within this system and identifying relevant metrics using these interactions. Based on this, a system breakdown and operational phases are defined, which express the system interactions and tasks of each agent. To analyze the safety of human-system interactions we use this breakdown and denote the interactions between human and ADS that are the most critical and use this to construct metrics.

2. HD-AV reference fleet definition

In the US, commercial trucking corporations are mostly structured under two different paradigms. The first model is led by owner-operators, who are independent contractors that lease their services to trucking companies or directly to shippers. The second consists of fleet operators, which rely on companies that own and operate their own fleet of trucks and are responsible for hiring drivers to operate these vehicles. Operators of a commercial truck must possess a valid Commercial Driver's License (CDL) and must undergo extensive background checks and training prior to operation. The Federal Motor Carrier Safety Administration (FMCSA) regulates standards relating to commercial trucking, including creating appropriate CDL tests and enforcing regulations.

For the purpose of this work, a reference fleet was defined based on a current sample of companies developing and testing in the HD-AV space. These include Aurora, Kodiak Robotics, TuSimple, Torc Robotics, and Ike (currently acquired by Nuro). Characteristics such as ODD restrictions, vehicle sizing, and human-autonomy teams were selected to construct a representative model of HD-AVs in the industry. In addition, we ensure they follow guidance from NHTSA for design of ADS systems. Although many existing HD-AV companies have significant differences in their designs and processes, we aim to construct a general case that is representative of the designs and processes of the short-term development of this industry (National Highway Traffic Safety Administration, 2017).

The reference fleet consists of retrofitted Class 8 commercial vehicles owned by fleet operators, who are responsible for developing and implementing the ADS (in coordination with an ADS developer), training safety drivers, and monitoring operations through a remote operations center. These vehicles operate on highways and on/off ramps, which are denoted the "middle mile" for commercial goods transport, as depicted in Figure 1. A summary of characteristics of the reference fleet are given in Table 1.

Fig. 1. Depiction of "middle mile" ADS operations.

Table 1. Reference Fleet Characteristics Summary.

Reference Fleet Characteristic	Description
Operational Profile	Highway and on/off ramp operations – "Middle mile" for commercial goods transport.
	Fleet operator owns vehicles and is responsible for training safety drivers.
	Pre-shift inspection checklists and special maintenance requirements for ADS components.
Vehicle Characteristics	Retrofitted Class 8 commercial vehicles (>33,000 lbs./14,969 kg)
ODD Restrictions	Location restricted through geofenced maps.
Road & Weather Conditions	Highway roads; Fair weather; High visibility conditions; Clear to mild rain
Human-Autonomy Teams	Case 1: 1 trained onboard safety driver, 1 trained onboard safety operator, remote operations center, ADS
	Case 2: 1 trained onboard safety driver, remote operations center, ADS
ADS Capabilities	Take input from sensor data to perform DDT task, automatic detection of moving to fallback/MRC state, notify driver/operator if approaching limits of ODD.

The reference HD-AV fleet operates in a restricted ODD with location constrained by geofenced maps. Their ODD is also restricted by weather conditions, with operations taking place in high visibility and fair-weather conditions, with the most severe weather condition being light rain. Although the idea of dedicated autonomous lanes has been proposed, ideally HD-AVs should be able to operate in mixed traffic scenarios, so those conditions are considered here. These vehicles operate at a nominal Level 4 of autonomy, which denotes that the vehicle itself must be able to initiate fallback autonomously if triggered by ODD exits, vehicle safety-critical failures, and other emergency situations. In the event the HD-AV exists the ODD under unplanned circumstances a DDT fallback is triggered until a Minimal Risk Condition (MRC) is achieved. In addition, operations consider the presence of a trained safety driver onboard and, in some cases, a trained safety operator, to oversee operations and intervene in the vehicle's operation for safety reasons (Aurora, 2021; Ike, 2019; Kodiak, 2020; Torc, 2021; TuSimple, 2020).

2.1. System breakdown and high-level tasks

To analyze the reference fleet's operations, the operations of the HD-AV system are broken down into different functional agents. The HD-AV system includes four main agents: the ADS, safety driver, safety operator (optional), and remote safety operator. These agents and their main task categories are defined in Figure 2 and described in depth in the following sections. This system and operational breakdown were created in accordance with the safety hazard identification methodology for automated driving systems fleets (Correa-Jullian et al., 2024).

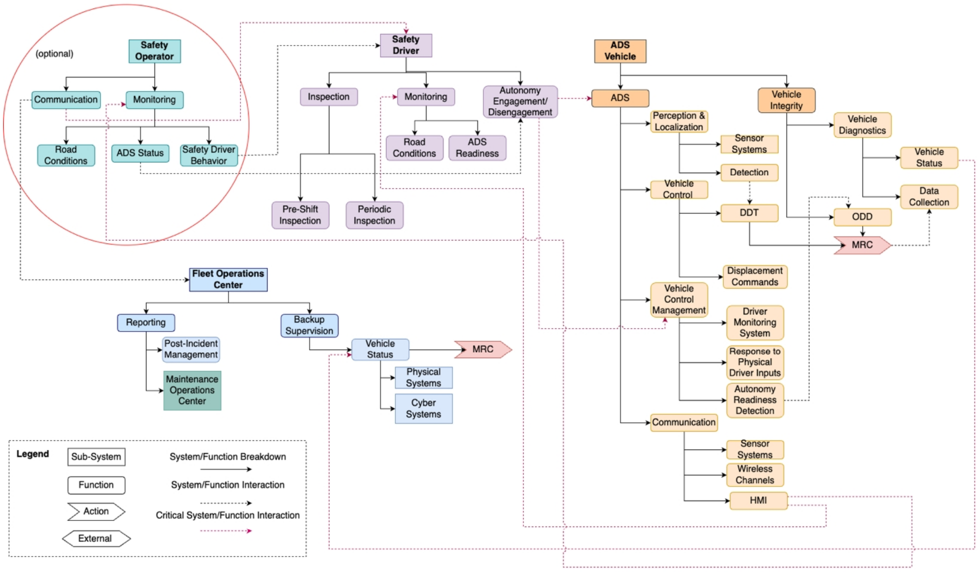

Fig. 2. System functional breakdown.

2.1.1. ADS

The first agent present in the HD-AV system is the ADS itself, which is the software and hardware responsible for performing the DDT within the limits of the ODD. The ADS has a nominal Level 4 of autonomy, which means that it designed to function without the need for a human to take over the vehicle while operating within the ODD. The ADS can perform the DDT fallback to reach MRC if required. However, current operations still require a safety driver to perform sections of transit outside of the ODD (i.e. before and after middle mile operations) and provide backup in emergency cases. In addition, the ADS contains a driver monitoring system to assess driver attention. A description of the ADS high-level tasks is shown in Table 2.

2.1.2. Safety driver

The safety driver is a commercial vehicle operator who possesses a valid CDL and has undergone training for commercial driving operations and interactions with the built-in ADS. It is expected that the safety driver is trained in identifying the ODD requirements, ADS limitations, and is instructed on emergency procedures. Their high-level responsibilities include driving the vehicle outside its ODD and engaging and disengaging the autonomous driving phase. A description of the safety driver's high-level tasks is shown in Table 3.

2.1.3. Safety operator

The safety operator is an additional human agent onboard the vehicle in the passenger's seat whose responsibility is to monitor road conditions and the state of the HD-AV. These operators may interact with dedicated Human Machine Interface (HMI) display to identify any potential issues internally with the ADS or externally in order to warn the safety driver. In addition, they serve as a party that enables communication between the safety driver and remote operations center. In the event the fleet operations do not include a safety operator, their tasks are incorporated into the safety driver's responsibilities. A description of the safety operator's high-level tasks is shown in Table 4.

2.1.4. Fleet operations center

The fleet operations center is a physical space in which hired operators monitor the HD-AV fleet in a control room environment. Each remote operator may be tasked to monitor multiple HD-AV systems through a dashboard and provide warnings to safety drivers and/or safety operators on board. In addition, remote operators receive incident notifications automatically and play a role in traffic, route, and accident management. A description of the high-level tasks of the remote operators at the fleet operations center is shown in Table 5.

Table 2. ADS Task Categories and Descriptions.

Task Category	Description
Inspection	Collect and process sensor input data coming from GPS, camera, radar, and LiDAR systems.
Vehicle Control	Plan and implement the DDT (Dynamic Driving Task) while the vehicle is under computer control.
	Issue actuation commands to the vehicle, including steering, throttle, braking, and indicator commands.
	Respond to physical inputs from Safety Driver.
	Assume fallback state when vehicle begins to exit ODD.
Vehicle Control Management	Determine autonomy readiness based on road conditions and ODD.
	Monitor Safety Driver behavior through driver monitoring system.
	Inform Safety Driver/Operator about autonomous system state.
Communication	Communicate with Safety Driver and Operations Center if there is an unexpected event while ADS is activated.
Vehicle Diagnostics	Assess and report status of vehicle subsystems, both related to ADS and non-automated systems.

Table 3. Safety Driver Task Categories and Descriptions.

Task Category	Description
Inspection	Conduct pre-trip inspection of safety-critical vehicle systems.
	Inspect truck and trailer every time the truck stops.
Monitoring	Monitor road and behavior of vehicle.
Autonomy Engagement/Disengagement	Engage and disengage vehicle's autonomy system.
	Take control of the vehicle in case of a disengagement.
	Manually drive the vehicle when it is outside of its ODD.
Communication	Communicate with FOC about issues.

Table 4. Safety Operator Task Categories and Descriptions.

Task Category	Description
Communication	Communicate ADS intentions, status, and misbehavior to Safety Driver.
	Communicate with operations center.
Monitoring	Monitor the operation of the ADS via HSI internal display.
	Monitor HSI for missed detections, false detections, unsuitable motion plans, and poor data quality.
	Warn safety driver to disengage autonomous system.
	Record notes about system and road conditions and incidents for post-shift debrief.

Table 5. Fleet Operations Center Task Categories and Descriptions.

Task Category	Description
Backup Supervision	Perform a backup support role for Safety Driver and Safety Operator.
	Monitor fleet of vehicles and their statuses from control room environment via dashboard.
Reporting	Communicate with Safety Driver/Safety Operator about potential obstacles and risks coming ahead.
	Respond to alerts of Safety Driver inattention.
	Warn safety driver to disengage autonomous system.
	Respond to accident scenarios, alert emergency services, and participate in post-accident debrief.

2.2. Operational profile

The operation of the reference fleet and the functions each of the human and machine agents perform may vary depending on certain conditions. For this work, the high-level tasks of the agents described in the previous section are organized into an operational phase diagram depicted in Fig. 3. A brief description of the five phases is presented specifically for middle mile operations.

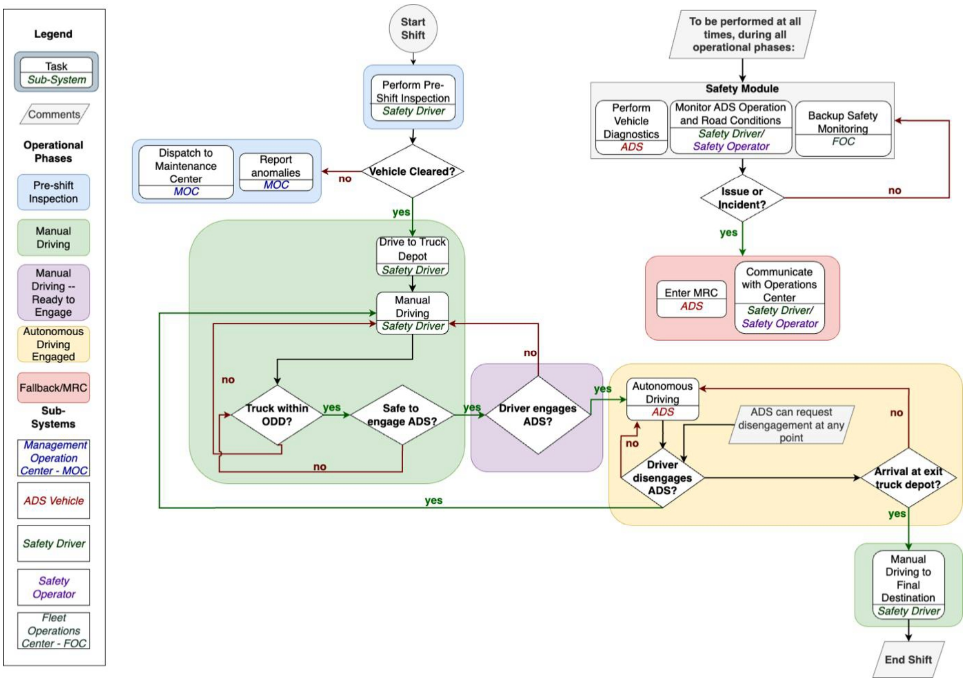

Fig. 3. Diagram of the operational phases and subsystems.

2.2.1. Pre-shift inspection

The first operational stage of a HD-AV is a pre-shift inspection, which is performed by the safety driver prior to beginning their driving shift. This pre-shift inspection is typical for commercial trucks and involves inspecting safety-critical components (i.e. brakes, fluid levels, and tire pressure) and tracking this analysis through a standardized checklist. With the introduction of the ADS, this pre-shift inspection now includes a thorough inspection of the retrofitted ADS sensors, connectors, and mounts. Additionally, the safety driver is required to conduct a similar inspection every time the truck stops, for instance: during refueling, mealtimes, and breaks. (Kodiak, 2020). If the vehicle is not in adequate condition, it is reported to the fleet operations center and goes to the maintenance operation center, whose job it is to conduct any repairs or tests that will make the vehicle operational. If the vehicle is in adequate condition, it is then approved to move on to begin the shift and proceed to Manual Driving.

2.2.2. Manual driving

The first section of the driving stage is Manual Driving. The Manual Driving phase refers to points when the safety driver is in full control of the vehicle during its DDT and is driving it outside of the bounds of it ODD. This stage can involve navigating non-highway roads leading to the truck depot, or any stage during the middle-mile journey in which the truck is not in its ODD and the safety driver is in control of the vehicle. The remote operations center monitors the vehicle's location, status, and any upcoming obstacles on the road. This stage can be interrupted by any incidents outside expected operations, e.g., system failures, unplanned route changes. In this stage, the safety driver is responsible for fallbacks and incident management. If an incident occurs, this stage is interrupted, and the fleet operations center is notified. The driver monitoring system, which is a component of the ADS, monitors the drivers to ensure that they are fully aware and attentive on the driving task at hand.

2.2.3. Manual driving—ready to engage

Manual Driving—Ready to Engage denotes when the vehicle has entered its ODD and the ADS has determined that it is safe to activate the autonomous driving phase, but the safety driver remains in control. This stage is communicated to the safety driver through a combination of visual and auditory messages. In this stage, the safety driver decides whether to engage Autonomous Driving. This stage can also occur if the safety driver disengages Autonomous Driving for any reason while the vehicle remains in its ODD. The driver has the option to reengage it while still in the Manual Driving—Ready to Engage phase. The remote operations center continues to monitor the vehicle's location, status, and upcoming obstacles. Like the Manual Driving phase, the safety driver is responsible for fallbacks and incident management. If an incident occurs, this stage is interrupted, and the fleet operations center is notified.

2.2.4. Autonomous driving engaged

This stage is when the ADS is performing the DDT and is controlling the steering, brake actuation, and road monitoring. The vehicle enters this stage when the safety driver activates the autonomous driving function via a button. In this stage, the safety driver remains in the car and monitors road conditions for any unexpected scenarios. While the vehicle is in this stage, the ADS is responsible for fallback, but the safety driver can also intervene and take manual control at this stage. For instance, this can occur when the triggers for fallback are not autonomously detected. The ADS at this stage also detects when the vehicle is nearing the limits of its ODD and alerts the driver with visual and auditory cues in this case, notifying them to resume manual control. Additionally, if the ADS detects that the driver is inattentive, the safety operator and fleet operations center are notified and they can proceed accordingly. If the vehicle exits its ODD and the safety driver has not taken control, or if an incident occurs, then the vehicle enters the Fallback/MRC stage of operations.

2.2.5. Fallback/MRC

The fallback stage can be triggered by several incidents, for instance: internal system failures, breaches of the ODD environment, rapid changes in weather or road conditions, or incidents with other vehicles or pedestrians on the road. In the event a DDT fallback is triggered, it is expected that the ADS plans and implements a DDT fallback strategy, and achieves an MRC – i.e., the vehicle implements a safe stop, unless the safety driver intervenes to resume Manual Driving. After an MRC occurs, the post-incident procedures are triggered and the safety driver and fleet operator decide the course of action to take, whether that be remotely assisting the vehicle, or recovering it for return to the maintenance operations center. In addition, a post-shift debrief is conducted to collect information about the fallback trigger and the result of the MRC.

3. Metrics to assess human-system interaction

Safety metrics are an effective framework to analyze overall system safety in tandem with other indicators of risk, and are used in a variety of industries, such as maritime, nuclear, and aviation. Metrics can either be considered lagging—based on past incidents, or leading—measuring non-accident scenarios that point to the likelihood of an accident (Reiman & Pietikäinen, 2012). While crash data is widely used in safety analysis, crashes are relatively rare occurrences, so it is often difficult to find meaningful patterns due to the limited data points. Thus, surrogate safety metrics (SSMs) have been developed to measure traffic conflicts, which are statistically linked to crashes, but do not necessarily result in a crash (Bin-Nun et al., 2023). Research has been conducted towards assessing human driving performance, including a defined driving inattention scale and behavior analytics. In addition, there have been several SSMs developed to analyze ADS performance (Automated Vehicle Safety Consortium, 2021; Wang et al., 2021). Fig. 4 depicts existing sources for safety indicators for both humans and ADS.

When considering humans operating non-autonomous vehicles, there have been several onboard driver monitoring systems as part of Advanced Driver Assistance Systems (ADAS) developed. These systems aim to assess the degree of attention the driver has when manually driving, through physical cues and eye movements which are analyzed through computer vision (Masala & Grosso, 2014). In addition there are retroactive self-reported assessment methods, namely the driving inattention scale (ARDES), which dictates the propensity of drivers to make attention-related errors. Originally developed in Argentina, ARDES has been validated to observe cross-cultural driving inattention patterns across various countries including Spain, the UK, and the USA (Castro et al., 2024).

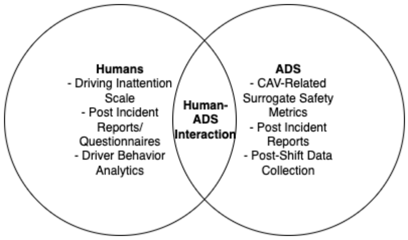

Fig. 4. Depiction of metric sources for humans and ADS

Several SSMs have been proposed to measure and contrast the driving safety performance of autonomous vehicles, some stemming from traffic engineering metrics created for human-driven vehicles, and others proposed following results from real-life and simulator environments. Some metrics proposed through this analysis include Time-to-Collision (TTC), which is the time until a collision between two vehicles in the scenario environment if they maintain present velocities. In addition, Aggressive Driving (AD), a binary metric, assesses whether an autonomous vehicle performs “repeated maneuvers (longitudinal and lateral accelerations) above specified thresholds completed by the ego vehicle that are defined as less safe” (Wishart et al., 2020).Other suggested metrics include, for instance, totally the number of instances when ADS does not react to a pedestrian or cyclist in a specific ODD.

However, there has been limited research done into metrics assessing the quality of interactions between human operators and the ADS system—thus there is a need to explore metrics that can help to explain these interactions and compare them across systems. Driving simulators, such as OpenCDA and CARLA, have been helpful to study aspects of driving including vehicle functions and behavior, as well as human factors. They are relatively low-cost, safe, and have the potential to more adequately model potential incidents where there is limited data (Dosovitskiy et al., 2017; Xu et al., 2021, 2023). Not only do simulators allow for the collection of data regarding vehicle performance, but when experimenting with safety drivers, certain factors of the human-system interaction stage can be assessed. Several studies have observed variables like ADS takeover decision-making, trust calibration, and organizational factors that lead to perception of ADS through data sources like physiological state and post-experiment questionnaires (Chu et al., 2023).

3.1. Proposed metrics

One of the main topics that has been assessed in ADS research is the degree of trust that human operators inside and outside the vehicle have on the system. (Yang et al., 2018). Although somewhat of a qualitative assessment, trust can be estimated in a numerical sense through a series of proposed metrics. Based on the previous discussion, some preliminary metrics are proposed as follows:

- *Divergence from ADS decisions:* This metric refers to the number of times during the Autonomous Driving Phase that the safety driver steps in and conducts a maneuver that differs from what the ADS was proposing to do. This can be accomplished by moving the steering wheel, braking, or pressing the “Disengage” button at a nominal stage within Autonomous Driving when there is not an incident or accident scenario that requires interception.

- *Time spent in manual driving—ready to engage phase:* This metric refers to the amount or percentage of time spent in the Manual Driving—Ready to Engage phase, which is when the ADS notes that autonomous driving is safe, but the safety driver chooses to not engage autonomy.
- *Time spent in autonomous driving engaged phase:* This metric represents the amount of time spent in the Autonomous Driving Engaged phase, where the Safety Driver is purely monitoring the ADS and not intervening. This metric can be used in tandem with the ADS SSMs in order to assess ADS performance during this stage of operation.
- *Reason for ADS disengagement:* This metric represents the reason the safety driver disengaged the ADS, as recorded through post-shift questionnaires.

Another consideration of human system interaction for heavy duty vehicles involves the team of human operators onboard in the case of both a safety driver and a safety operator. The dynamics between these two operators can affect the decision-making and action-taking processes regarding the ADS. For instance, if a safety operator notes an upcoming new road condition or potential for an incident, but the driver ignores them and is reliant on either their own perspective or that of the ADS sensors, these disagreements may develop into more complex hazard scenarios. Although we can also assess the disagreements between safety driver, safety operator, and ADS in a numerical manner denoting number of times a disagreement occurs, it is likely that a more complete picture will be formed through a post-shift debrief, in which questionnaires will assess the motivation behind such a disagreement.

These metrics are critical to assess the system and operational design of the HD-AV system, and because of data gleaned from these metrics, functions of the system or triggers between phases may be modified. Furthermore, the human-system interaction metrics can be refined by conducting Probabilistic Risk Assessment (PRA). This assessment relies on tools like event sequence diagrams (ESDs), fault trees (FTs), and, more recently, Bayesian Networks (BNs) to model and assess risk scenarios. Another method has also been developed to study the interactions of human-systems in autonomous systems based on task analysis. This concurrent task analysis (CoTA) has been applied to autonomous marine vessels, and its hierarchical model can be used to identify the most important tasks that the human agent and ADS system need to perform. This can then be used to establish which metrics are the most important and inform the creation of new metrics (Ramos et al., 2020a, 2020b).

4. Discussion and concluding comments

The autonomy level of current commercial HD-AV systems is commonly advertised at an SAE Level 4, which is defined as the ADS performing the entire DDT and all associated fallback procedures. However, in current testing operations there is always a safety driver in the vehicle who can intervene in case of emergencies. The active role of the safety driver is not expected to be removed due to current autonomous vehicle legislation and recent events in driverless passenger transport applications (Mickle et al., 2023). Thus, the effective level of autonomy may decrease to Level 3 according to current J3016 definitions; however, since the ADS nominally plans and implements the DDT fallbacks, further discussions may be required to assess intermediate levels of automation. There has been criticism on the currently defined SAE Levels, especially that their strictly linear progression overlooks the necessary hybridity of interactions between human and automated technologies (Hopkins & Schwanen, 2021). We can refer to the HD-AV system as existing at a nominal Level 4 of automation; namely it has characteristics of a Level 3 system, however at a nominal state it can perform all the tasks expected of a Level 4 system. As HD-AV companies more clearly define roles and tasks we can reassess the autonomy state and the presence of safety drivers and operators.

Since the interplay between safety drivers and the ADS plays a large role in the division of tasks and operation of heavy-duty systems, there is a need to include safety metrics about how these humans and systems interact. This can help inform the designs of ADS systems in turn; for instance, making HMIs more effective for safety drivers, developing more robust driver monitoring systems, and designing alert functions. In addition, observing these metrics can help with designing HD-AV operations for both efficiency and safety without promoting one at the expense of the other. A focus on human-centered metrics has been applied in various industries like nuclear, aviation, maritime, etc., to develop effective human reliability assessments (HRAs) which are conducted through models such as Phoenix, SPAR-H, and IDAC. Existing HRA methods can be expanded upon to also incorporate autonomous vehicles in their models and potentially be used to expand current PRAs conducted to inform the design and regulation of these automated driving systems.

Acknowledgements

This research is partly supported by the University of California Institute of Transportation Studies. The work presented in this paper remains the sole responsibility of the authors.

References

Aurora. 2021. Voluntary Safety Self-Assessment.
Automated Vehicle Safety Consortium. 2021. Best Practice for Metrics and Methods for Assessing Safety Performance of Automated Driving Systems (ADS). www.sae.org.
Bin-Nun, A. Y., Lizarazo, C., Panasci, A., Madden, S. & Tebbens, R. J. D. 2023. What do surrogate safety metrics measure? Understanding driving safety as a continuum. Accident Analysis and Prevention, 195. https://doi.org/10.1016/j.aap.2023.107245.
Bonnefon, J.F. 2021. The car that knew too much: Can a machine be moral? MIT Press.
Castro, C., Doncel, P., Ledesma, R. D., Montes, S. A., Barragan, D., Oviedo-Trespalacios, O., Bianchi, A., Kauer, N., Qu, W. & Padilla, J. L. 2024. Measurement invariance of the driving inattention scale (ARDES) across 7 countries. Accident Analysis and Prevention, 195. https://doi.org/10.1016/j.aap.2023.107412.
Chu, M., Zong, K., Shu, X., Gong, J., Lu, Z., Guo, K., Dai, X. & Zhou, G. 2023. Work with AI and Work for AI: Autonomous Vehicle Safety Drivers' Lived Experiences. Conference on Human Factors in Computing Systems - Proceedings. https://doi.org/10.1145/3544548.3581564.
Correa-Jullian, C., Ramos, M., Mosleh, A. & Ma, J. 2024. Operational safety hazard identification methodology for automated driving systems fleets. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability. https://doi.org/10.1177/1748006X241233863.
DeGuzman, C.A., Kanaan, D., Hopkins, S.A., Donmez, B. 2021. Takeover request (TOR) effects during different automated vehicle failures. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations. https://doi.org/10.1080/15472450.2021.1891536.
Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A. & Koltun, V. 2017. CARLA: An Open Urban Driving Simulator. http://arxiv.org/abs/1711.03938.
Hopkins, D. & Schwanen, T. 2021. Talking about automated vehicles: What do levels of automation do? Technology in Society, 64. https://doi.org/10.1016/j.techsoc.2020.101488.
Ike. 2019. Safety Report.
Kodiak. 2020. Kodiak Safety Report 2020.
Masala, G. L. & Grosso, E. 2014. Real time detection of driver attention: Emerging solutions based on robust iconic classifiers and dictionary of poses. Transportation Research Part C: Emerging Technologies, 49, 32–42. https://doi.org/10.1016/j.trc.2014.10.005.
Mickle, T., Metz, C., & Lu, Y. (2023, November 3). G.M.'s Cruise Moved Fast in the Driverless Race. It Got Ugly. The New York Times.
Morales-Alvarez, W., Sipele, O., Léberon, R., Tadjine, H. H. & Olaverri-Monreal, C. 2020. Automated driving: A literature review of the take over request in conditional automation. In Electronics (Switzerland), Vol. 9, Issue 12, pp. 1–34. MDPI AG. https://doi.org/10.3390/electronics9122087.
National Highway Traffic Safety Administration. 2017. Automated Driving Systems 2.0: A Vision for Safety.
National Highway Traffic Safety Administration. 2023. Traffic Safety Facts - 2021 Data: Large Trucks.
Ramos, M. A., Thieme, C. A., Utne, I. B. & Mosleh, A. 2020a. A generic approach to analysing failures in human – System interaction in autonomy. Safety Science, 129, 104808. https://doi.org/10.1016/J.SSCI.2020.104808.
Ramos, M. A., Thieme, C. A., Utne, I. B. & Mosleh, A. 2020b. Human-system concurrent task analysis for maritime autonomous surface ship operation and safety. Reliability Engineering & System Safety, 195, 106697. https://doi.org/10.1016/j.ress.2019.106697.
Reiman, T. & Pietikäinen, E. 2012. Leading indicators of system safety - Monitoring and driving the organizational safety potential. Safety Science, 50(10), 1993–2000. https://doi.org/10.1016/j.ssci.2011.07.015.
SAE International. 2021. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. SAE Standard J3016.
Shetty, A., Tavafoghi, H., Kurzhanskiy, A., Poolla, K. & Varaiya, P. 2022. Automated Vehicle Safety and Deployment: Lessons from Human Crashes. 2022 International Conference on Connected Vehicle and Expo (ICCVE), 1–6. https://doi.org/10.1109/ICCVE52871.2022.9742994.
Singh, S. 2015. Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey (Traffic Safety Facts Crash Stats. Report No. DOT HS 812 115). Technical Report, United States National Highway Traffic Safety Administration.
Torc. 2021. Innovating Safety and Efficiency - Torc Safety Report.
TuSimple. 2020. TuSimple Safety Report Version 2.0.
Wang, C., Xie, Y., Huang, H. & Liu, P. 2021. A review of surrogate safety measures and their applications in connected and automated vehicles safety modeling. Accident Analysis and Prevention, 157. https://doi.org/10.1016/j.aap.2021.106157.
Wishart, J., Como, S., Elli, M., Russo, B., Weast, J., Altekar, N. & James, E. 2020. Driving Safety Performance Assessment Metrics for ADS-Equipped Vehicles. SAE Technical Papers, 2020-April(April). https://doi.org/10.4271/2020-01-1206.
Xu, R., Guo, Y., Han, X., Xia, X., Xiang, H. & Ma, J. 2021. OpenCDA: An Open Cooperative Driving Automation Framework Integrated with Co-Simulation. 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), 1155–1162. https://doi.org/10.1109/ITSC48978.2021.9564825.
Xu, R., Xiang, H., Han, X., Xia, X., Meng, Z., Chen, C. J., Correa-Jullian, C. & Ma, J. 2023. The OpenCDA Open-Source Ecosystem for Cooperative Driving Automation Research. IEEE Transactions on Intelligent Vehicles, 8(4), 2698–2711. https://doi.org/10.1109/TIV.2023.3244948.
Yang, X. J., Pradhan, A. K., Tilbury, D. & Robert, L. 2018. Human Autonomous Vehicles Interactions: An Interdisciplinary Approach.
Zhai, S., Gao, S., Wang, L. & Liu, P. 2023. When both human and machine drivers make mistakes: Whom to blame? Transportation Research Part A: Policy and Practice, 170, 103637.
Zhang, B., de Winter, J., Varotto, S., Happee, R. & Martens, M. 2019. Determinants of take-over time from automated driving: A meta-analysis of 129 studies. Transportation Research Part F: Traffic Psychology and Behaviour, 64, 285–307. https://doi.org/10.1016/j.trf.2019.04.020.