Chapter 20
Safety Performance Assessment of Assisted and Automated Driving in Traffic: Simulation as Knowledge Synthesis

Thomas Helmer, Klaus Kompaß, Lei Wang, Thomas Kühbeck, and Ronald Kates

20.1 Introduction

20.1.1 Driver Assistance and Automation

Based on comprehensive, sensor-based detection of vehicle surroundings, the vision of self-driving vehicles has already become a reality in research vehicles (e.g., [1, 2]). Today, advanced driver assistance systems (ADAS) or partial automation, in which the vehicle assumes the role of a “copilot,” can assist the driver in demanding situations, support him in complex or hazardous situations, and, if necessary, perform maneuvers automatically. Higher levels of automation can, at least temporarily, free the driver entirely from the driving task and enable him to fulfill other tasks. Thus, automation could cause a considerable improvement in the quality of individual mobility [2].

In addition to individual mobility, collective and socioeconomic impact of vehicle automation have a central role regarding further development and dissemination of those functions. In highly developed and industrialized countries, especially within metropolitan areas, satisfaction of the “mobility demands” often requires additional infrastructure and induces impacts on environment and quality of life. High percentages of automated vehicles could enable technological strategies that
increase the capacity of streets and the “mobility supply” and thus allow a more environmentally friendly and efficient use of resources.

Reliability and safety criteria are key factors for homologation and operation of advanced driver assistance systems and especially automated driving functions. Only when the overall safety level of a technology reaches (or exceeds) a generally accepted threshold should secondary criteria, such as comfort or efficiency, be considered.

Since the individual customer is hardly able to independently evaluate the safety features when purchasing a vehicle, he has to rely on trustworthy and objective evaluations by third parties. In addition to automotive manufacturers and suppliers, public decision-makers, regulatory agencies, insurance companies, and consumer protection organizations are key stakeholders in the evaluation of vehicle safety. In particular, demonstration of equivalent or superior safety in all operating scenarios is an important prerequisite for the introduction and homologation of novel technologies, such as automated driving functions (ADF).

Ideally, it should be possible to derive quantifiable predictions of safety impacts for new technologies in such a way that all stakeholders can agree on their relevance, objectivity, validity, and reproducibility—despite their sometimes divergent interests. A prerequisite for this is a careful definition of an appropriate safety metric with respect to the specific context.

Thus, for example, in the context of pedestrian protection, the reduction of pedestrian injury by mitigating or avoiding collisions is the primary objective. The quality of pedestrian protection system could be objectively quantified by a metric which relates to injury reduction (see below).

An objective metric of traffic safety also includes possible side effects [8, 13, 22, 33]: preventative pedestrian protection, for example, can result in occasional unwanted system actions due to the underlying physics. As a consequence, secondary risks can occur (e.g., rear-end collisions following an emergency braking maneuver). With the help of an objective metric, the ratio of desired safety improvements to unwanted side effects can be quantified. Based on this ratio, the function can be optimized with respect to traffic safety (Fig. 20.1).

Several important aspects of vehicle safety such as integrity of communication, data protection, and technical reliability are beyond the scope of this paper. The focus here is on the effects of ADAS and current as well as future fully operational automated systems on traffic safety.

Although ADAS are generally designed to minimize well-known accident risks or mitigate the consequences of unavoidable accidents, the optimal design of ADAS in terms of traffic safety is a complex task. ADAS usually rely on sensor-based decision algorithms, which often depend on certain parameter settings. However, any interaction with the driver or interference with the driver’s task can in principle have a positive or negative impact on traffic safety, for example:

- **Lane departure warning**: A typical criterion for an alarm or intervention in case of impending lane departure would be “time to lane crossing” (TLC), calculated using sensor data. It could be assumed that earlier warnings (large
TLC—thresholds) are “safer.” However, too many unnecessary warnings could lead to negative feedback, such as a lack of acceptance.

- *Emergency braking systems:* It could be assumed that earlier system interventions avoid more rear-end collisions. However, some drivers might still execute an evasive maneuver, even if accident avoidance by braking was no longer possible (this is a well-known dilemma regarding braking and steering). Moreover, if the system reacts in such situations too early with a correspondingly strong intervention, the risk of accidents for the driver (loss of control) as well as for subsequent vehicles (rear-end collisions) can arise.

### 20.1.2 Assessment and Optimization of ADAS and ADF as Key Processes During the Development

The examples illustrate that design of advanced driver assistance systems or automated driving functions is a complex task, especially with regard to optimization of traffic safety.

One approach for structured analysis of impacts on traffic safety of existing ADAS is retrospective analysis of accident databases linked to vehicle equipment data (if feasible). However, large samples from accident databases and long observation times [8] are required for retrospective statistical evaluation of accident avoidance due to ADAS. Statistical challenges for evaluation of a particular ADAS arise due to multifactorial data characteristics: vehicles in the sample may differ in several aspects, not only regarding the equipment with the system in question. Vehicles from different manufacturers may have different functional characteristics, interaction concepts with the driver, or activation conditions. Differences of driver population or exposure to traffic scenarios can be correlated with the equipment. Such properties of retrospective studies act statistically as confounders and make an unambiguous interpretation of results and conclusions about causality challenging.
Comprehensive statements regarding changes in traffic safety ("effectiveness") are therefore often problematic [13, 22].

In addition, evaluation of the consequences of unwanted system actions is practically impossible based on retrospective studies alone, since the relevant accident types in general relate to a range of possible secondary effects.

Summarizing, retrospective studies can enable ex post evaluations, but can hardly be used for any optimization of ADAS or ADF, due to the long feedback loop of the development, retrospective evaluation, and redesign.

Interestingly, retrospective surveys are also considered as second best for assessment of new treatments or therapies in health care: the standard of quality is rather a prospective, randomized, controlled trial in a representative sample.

There is an urgent need for reliable and valid predictions of traffic safety for design and optimization of ADAS and especially for ADF. The evaluation and optimization of such systems using traffic safety predictions should be an integral part of the development process.

However, truly prospective, randomized, controlled, and representative studies of ADAS are not feasible on public roads due to ethical and practical considerations. This dilemma could be resolved using virtual, simulation-based traffic safety predictions of ADAS and ADF.

### 20.2 Overall Safety Assessment

#### 20.2.1 Safety and Economy

Currently, passive, active, and integral safety functions contribute to the goal of improved vehicle safety [23, 25]. In the future, automation is also expected to exert a decisive influence on vehicle safety. It is increasingly necessary to assess and compare competing, technically feasible safety measures regarding their effectiveness in order to define appropriate priorities. Therefore, a comprehensive approach for safety assessment of new technical concepts is required.

Since automation affects the entire traffic flow, it can also have an impact on individual driving comfort and efficiency. Thus, in assessing ADF, comfort as well as social and economic aspects will play an important role in addition to safety.

#### 20.2.2 Conflicting Objectives in Vehicle and Traffic Safety

In assessing effectiveness of ADAS and ADF, both positive influences as well as risks and side effects need to be considered and included [8, 13, 22, 33].

Considering active safety systems, some conflicting objectives are well known, for example, during the precrash phase: if a system classifies a risk situation
very late, “false negatives” or too late system actions can occur, reducing the effectiveness of the ADAS. Conversely, if a system responds to a potential hazardous situation very early, “false positives” can result (see Table 20.1): in these cases, the driver could possibly have avoided the accident even without system intervention. Excessive frequency of false positives can reduce acceptance and even result in negative feedback on traffic safety [8, 13, 20].

For future automated systems, even more complex trade-offs between conflicting objectives can be expected during the design phase.

The comparison of effectiveness and conflicting objectives is part of an integrated development and assessment process of advanced vehicle safety systems. Current methods based on single tests usually cannot properly take conflicting objectives into account [22, 33].

In order to quantify the effectiveness of ADAS in a target scenario by a virtual experiment (see below), an appropriate metric for characterization of the safety benefits is required. An ideal metric includes both avoided accidents and reduced injury severity as well as reduced fatalities in the remaining accidents. Probability models can be used to derive injury severity based on detailed accident characteristics in the target scenario. To calculate the metric, two steps are required: first, calculating the impact of the system on accident characteristics in the target scenario and, second, applying a conditional probability model of injury severity depending on these characteristics.

To model injury severity—quantified as MAIS (maximum abbreviated injury scale), for example—depending on accident characteristics, there are several complementary approaches. A commonly used method is the construction of statistical models (e.g., regression models) from existing accident databases (e.g., [11, 14–16]). Another approach is “co-simulation.” Here, a representative sample of time series from accident simulations is generated and analyzed using a high-resolution crash simulation, which is capable of rapidly calculating injury indicators [9, 35].

As mentioned above, certain interventions of ADAS, such as automatic braking or evasive maneuvers, can induce serious side effects into traffic, for example, loss of controllability.

Milder side effects, for example, decrease in user acceptance, can also be caused by controllable interventions or warnings, if they are perceived as superfluous (false positive). Since ADAS can only fulfill their purpose once they are activated, a high alarm rate, lacking comprehensible justification for the driver, can lead to disproportionate non-utilization rates or frequent lack of response to warnings.

---

**Table 20.1** Categorization of possible system actions (following [21])

| Objective risk | System action     |
|----------------|-------------------|
|                | Yes   | No                |
| Yes            | True positive | False negative   |
| No             | False positive | True negative    |
Such unintended side effects impact the overall quality of an ADAS and should therefore be considered as part of an assessment. For this purpose, it is necessary to classify false-positive actions and other unintended side effects regarding their frequency and severity.

A key characteristic, borrowed from medicine, is the NNT (“number needed to treat”). In terms of ADAS, NNT can be defined as the ratio between all system actions and true positives, i.e., desired system actions [5, 11, 31]. NNT can be calculated separately for each type of system action (warnings, interventions).

It is useful to distinguish false positives occurring due to technical limitations from those occurring due to situational uncertainties (such as a possible mitigation of a hazardous situation by other traffic participants). A camera-based sensor, for example, can cause false positives if a phantom object is detected due to light and shadow that apparently poses a hazard. Eliminating technically induced false positives may sometimes be accomplished without compromising effectiveness using, for example, test drives.

However, a portion of “unnecessary” system actions arises, as described above, due to essential conflicting objectives and cannot be eliminated without impact on effectiveness. Therefore, a systematic balancing in terms of “benefits” and “costs” is useful during design of most ADAS: each system action is assigned a value for “costs,” and each avoided adverse event (e.g., accident) is assigned a value for “benefit.” The safety effect of an ADAS can then be optimized based on this balance (e.g., using NNT).

In practice, it is hardly feasible to carry out such an optimization based solely on empirical driving tests or other classical testing setups. Due to the rarity and variability of accident scenarios, it is virtually impossible to obtain a statistically reliable proportion of avoided accidents in the test time available. Therefore, it can be considered current practice to design advanced driver assistance functions to have a false-positive rate below a given threshold. The dependence of effectiveness on control parameters (such as threshold TTC value for triggering) may show trends in individual experiments; however, this approach does not allow optimization with respect to effectiveness and false positives. True ADAS optimization requires feedback between the assessment and development processes. Finally, the assessment of ADF will also require meaningful methods allowing an integration of optimization into the development process [8, 23, 33].

1“Costs” in this context is used as collective term for unintended side effects, not necessarily in the monetary sense.
20.3 Design and Optimization of ADAS Using Virtual Experiments

20.3.1 Paradigm of Design of Virtual Experiments

The concept of virtual experiment plays a central role in the comprehensive safety assessment of ADAS. A typical design of a virtual experimental uses the paradigm of a prospective, randomized, controlled, and representative trial. The difference to a “real” experiment is the substitution of the actual traffic flow with a simulated traffic flow and the representation of other “real” components by simulated components. Quality and validity are central requirements in this paradigm.

As in prospective, randomized trials, “treatments” (here one or more variants of an ADAS) are compared to a reference (control or “baseline”)—for example a vehicle without the system. In the virtual trial design, target variables are statistically “captured” and used to calculate the relevant metric of traffic safety, analogous to empirical tests. Effectiveness of a system to the baseline can be quantified objectively using the metric. Just as in a randomized, controlled trial, the detection of small differences of effectiveness using statistical tests requires correspondingly high numbers of cases or long observation periods in virtual trials. Unlike in empirical studies or test driving, high virtual sample sizes can be produced with relatively modest resources.

20.3.2 Representation of Safety-Relevant Processes in Simulation

In principle, all relevant dynamic and human processes are represented as time series of states within the simulation. The simulated changes of those states can have both “deterministic” and “stochastic” characteristics. For example, the yaw rate of a vehicle on a dry surface is primarily deterministically dependent on steering wheel angle and speed, but slippage introduces an effectively stochastic component. The duration from a collision warning to braking by the driver can vary widely within a population of drivers and therefore must be regarded as stochastic. Simulation models must therefore be capable of representing not only deterministic but also stochastic properties.

Using safety assessment for the design and development of ADAS must therefore consider the stochastic properties of all relevant technical and human processes.
20.3.3 Knowledge Synthesis and Integration of Other Test Domains

All safety-relevant processes are modeled in sufficient detail in the simulation. The strategy is to capture the effects of ADAS on these processes. These processes are related to exposure variables, traffic flow, and dynamics of the driver-vehicle unit including human factors, technical systems, and so forth. There are many possible interactions within a driver-vehicle unit and between traffic participants and their surroundings. For ADF, the interaction between driver and vehicle may possibly be very limited (Fig. 20.2). Traffic safety is a result of all elements of the driver-vehicle control loop [26], together with other factors, such as infrastructure or regulations.

For technical systems, the entire process chain must be represented. The process chain generally consists of sensors, traffic environment modeling, algorithms (logic), and vehicle dynamics controllers and actuators (Fig. 20.3). Additionally, the system impacts and feedback loops on driver, vehicle, and traffic are modeled.

The models must in particular take the stochastic nature of physiological, psychological, and physical phenomena into account. Each warning or any direct intervention of an ADAS may trigger or affect a driver’s individual response.

In this paradigm, simulation serves as a form of knowledge synthesis. The models are synthesized to obtain realizations of all process chains and finally the safety metric. The validity of any result from the virtual testing of ADAS will depend, of course, on the quality of the underlying knowledge base (i.e., the models). The development of an appropriate knowledge base therefore represents one of the central tasks of a holistic assessment approach.

Fig. 20.2 Possible interactions of driver-vehicle-surroundings including ADAS or ADF; each including possible interactions with another human driver (after [3])
A wide range of existing data sources on vehicle, traffic, and human factors is used for modeling: accident databases and traffic surveys, NDS (“naturalistic driving studies”), FOT (“field operational test”), static traffic observations (e.g., highly instrumented research intersections), as well as classical testing methods such as tests on test tracks, laboratory experiments, and software, hardware, and “vehicle in the loop” (see, e.g., [7, 11, 12, 34, 36]).

In order to create relevant scenarios with a representative frequency in the simulation, “exposure models” are required. Exposure models, i.e., insights about the frequency of certain constellations of risk factors, are supported by traffic surveys, NDS, and FOT, for example.

Test methods are subject to certain ethical and practical limitations, for example, they need to be risk-free. Although classical testing methods alone hardly enable overall conclusions about effectiveness and the optimal design of ADAS, they provide valuable information to describe human or technical factors in a limited context. Functions can be tested on test tracks under varying conditions (road classes, weather) using a real vehicle with full experimental control. The obtained findings are used in virtual experiments to calibrate and to validate various simulation models (Fig. 20.4). Modeling can be performed empirically and/or on a theoretical basis.

For comprehensive assessments, representative scenarios are needed. From the set of relevant scenarios, individual situations are “sampled,” i.e., created virtually. The values of all stochastic variables are drawn from appropriate distributions.

Sampling may be repeated or independent. In repeated sampling, each randomly generated scenario is simulated multiple times (e.g., with/without ADAS), whereas in independent sampling, new samples are drawn for each run.

Using simulation-based knowledge synthesis, it is possible to derive comprehensive and representative conclusions regarding effectiveness of ADAS.
20.3.4 Process Description for Assessing Pedestrian Protection

Some of the typical process models used for virtual testing can be illustrated in more detail using the example of a hypothetical camera-based ADAS for preventative pedestrian protection [18]. In an exemplary scenario, a pedestrian crosses a straight one-way street (one lane) between two intersections without occlusion. In this scenario, traffic flows with typical (daytime dependent) urban speeds, and there is no evasive maneuver possible for the vehicle.

If there is a conflict between the vehicle with ADAS and a pedestrian, the pedestrian protection system can issue a warning, reduce brake assist thresholds, or brake automatically.

In addition to models for the pedestrian protection system, process models are required for traffic participants, especially for pedestrians, for directly involved vehicles, and for traffic flow in general. The models describe exposure (i.e., how often certain constellations of variables occur), traffic environment, and dynamics of all road users.

As part of the exposure model for pedestrians, samples for the initial situation can be drawn from (possibly correlated) model distributions, including:

- Context variables: frequency of pedestrian crossings by age and sex depending on time of day and day of week
- Physiological attributes of pedestrians: height and weight as functions of age and sex; fatigue and blood alcohol depending on age, gender, day of week, and time of day
- Cognitive characteristics: attention, visual performance, and reaction skills (related to age, sex, alcohol, fatigue)

After a “virtual pedestrian” has been “drawn” and “generated” with his characteristics, the crossing processes are simulated using these features. Included are decision-making processes, such as gap acceptance, walking vs. running, selected destination and walking direction, initial velocity (which of course depends on the estimate crossing time), etc.
The modeling of a process such as gap acceptance often requires a considerable degree of complexity: high traffic density, for example, can increase the waiting time for a larger time gap; hence, some pedestrians become “impatient,” with the result that they also accept smaller gaps after a certain waiting period. Various causes of cognitive estimation errors can lead to misjudgment of the time gap or the time required for crossing and thus provoke traffic conflicts.

Simulated safety metrics exhibit an influence of observable variables, such as age or impairment. The dependence on these variables can be compared to corresponding dependencies in published studies in order to validate detailed modeling aspects. For the assessment of ADAS, such detailed modeling aspects are at least indirectly relevant, because they determine the composition of the risk factors in conflicts and thus the potential effectiveness of ADAS.

Other stochastic process models describe perception of acute conflict, responses, and actions of the pedestrian (retreat, anticipation, etc.). The corresponding probability distributions of actions can, e.g., depend on the cognitive and physiological requirements and states at the time of reaction.

There are also corresponding process models for the driver. The cognitive state, for example, is essential for the driver’s safety performance. The perception models consider, for example, geometric relations, environmental conditions, and dynamics of the situation. Following initial perception, reaction processes are modeled in accordance with established model paradigms [10, 17, 27, 28], depending on the element of surprise and cognitive status.

The actual simulation run includes dynamic models for pedestrian and vehicle movement, which in turn depend on the other process models. For both pedestrians and drivers, strong individual differences can be observed. Individual differences, such as reaction time of the driver or braking intensity, are especially important for the assessment of traffic safety. As a result, a random number of individual, but in sum representative, situations can be created on the basis of a traffic scenario (Fig. 20.5).

At the end, time series of all virtual state variables can be extracted from every run—including those variables that are difficult to detect in standard lab experiments. An example is the determination of which stimulus has led to the reaction and deceleration by the driver, i.e., his own perception of the critical situation or a system warning. The corresponding cognitive processes may occur in parallel. These findings can be incorporated directly into the design of a warning strategy.

### 20.3.5 Simulation of ADAS Effectiveness

In addition to the simulation of traffic flow, vehicle dynamics, other traffic participants, as well as human factors, such as perception and reaction, the process of virtual evaluation (Fig. 20.6) also requires an appropriate model of the driver assistance system.
The precise technical realization of an ADAS plays in general an indirect role for effectiveness assessment. For example, the detection of a pedestrian by means of a camera and the detection of the distance and the relative speed to a preceding vehicle are highly complex and device-specific technical processes. In order to assess the potential effectiveness of an ADAS, it is often useful to begin with an abstract model: system concepts can be analyzed at an early stage, without being limited by the exact technological realization.
Models of performance, detection characteristics (sensors, object classification, etc.), algorithm logic, and actuator models are in many cases important for simulation and assessment of ADAS.

The effectiveness of ADAS is, among other things, determined by the system boundaries. These can theoretically depend on the traffic context (such as vehicle speed or road class given by a digital map), environmental conditions (light conditions, weather), or on conditions for automatic system activation or deactivation. Furthermore, e.g., system activation by the driver or other human factors can influence effectiveness.

Detection systems in traffic are subject to system limits and to various uncertainties and latencies. As a consequence, in practice, detection systems often show stochastic characteristics. For example, the time for stable object recognition by means of a camera can depend on partial occlusion of the object or complexity of the traffic scene. The algorithms for situation detection and action usually rely on measurements from the detection systems; the derived characteristics (such as estimated “time to collision” for a detected object) are therefore also subject to corresponding latencies and uncertainties that require appropriate, mostly stochastic, modeling. Also, system actions often act indirectly by stimulating a driver reaction (e.g., warnings) or interact with (stochastic) driver actions (e.g., by lowering the threshold for brake assist). Overall, stochastic characteristics have a major impact on the overall safety assessment of ADAS.

With increasing complexity of the systems, an abstract representation of the functionality of an ADAS may require considerable effort. Also, verifying that an abstract system model actually behaves like the real technical system poses a challenge with increasing system complexity. To meet this challenge, alternative approaches are possible.

Instead of abstract models, real components of ADAS can be directly connected to the simulation using an appropriate test facility (e.g., hardware in the loop) [12]. In addition, findings from such test benches can be used for calibration and validation of relevant models in the simulation even without a direct connection. The actual code may be used in the simulation instead of an abstract model of the system logic. This results new challenges for the simulation and the models used due to the technical interfaces used.

20.3.6 Interpretation of ADAS Effectiveness

In the example scenario of “preventative pedestrian protection,” millions of virtual crossings in the reference scenario (without pedestrian protection system) lead to several thousand collisions between pedestrians and vehicles. The attributes of these collisions are “known” from the simulation, i.e., speed, vehicle characteristics, collision constellation, age group, body height and weight of the pedestrian, etc. Using an appropriate model (see above), the distribution of the injury severity in collisions without ADAS can be determined. This distribution is used as reference for comparison.
A lower number of collisions occur in the virtual experiment after the same number of crossings with a pedestrian protection system. The properties of the collisions (such as reduced collision speeds) are changed due to the ADAS, and consequently the distributions of injury severity are modified. In this case the metric of injury severity represents the effectiveness of the preventative pedestrian protection system. The overall safety performance includes also NNT and an assessment of possible secondary risks, such as rear-end collisions in upstream traffic.

20.4 New Challenges for Virtual Assessment of Automated Driving Functions

20.4.1 Impact of Automated Driving Functions on Safety-Related Processes in Traffic

The monitoring function of systems of active safety is carried out in continuous operation; however, warnings and interventions are sporadic events. Apart from theoretically possible feedback, for example, through changes in user behavior, the perceived influence on the driving task usually remains very limited.

In contrast ADAS with regular control operations, such as active cruise control (ACC) or ADF, operate continuously. Safety assessment of these functions requires consideration of their effect on safety-related processes in traffic flow as a whole, not only in certain target scenarios. Overall traffic safety includes both positive effects and potential risks. This assessment requires a variety of additional exposure and process models.

20.4.2 Contributions to Safety Impact in Existing Risk Scenarios

A portion of possible positive contributions of ADF (similar to ADAS) comes from consideration of relevant and potentially hazardous scenarios where the advantages of ADF help avoid potential accidents or mitigate their consequences. “Scenario” in this context refers to all potentially hazardous traffic situations that can lead to a certain type of conflict.

Using virtual experimental design (as for ADAS), an appropriate reference sample of relevant scenarios can be defined and considered. The contribution to effectiveness due to ADF can be quantified using a sample of virtual experiments, once the scenarios and their frequencies are known.
**Example, rear-end scenario on a freeway:** Without ADF (or ADAS), an inattentive driver (failure “a”) or a driver with unadjusted speed (failure “b”) has an increased accident risk in a rear-end scenario. The rear-end scenario could be provoked by very slow or stationary downstream traffic behind a sharp curve or a hill. The driver experiences a surprising, sudden speed drop.

Lack of attention (a) is a typical consequence of insufficient driver activation, for example, after a long and strenuous journey in stop-and-go traffic. Inappropriate speed (b) can have various causes, for example, the latent danger of a freeway curve or a hill (due to the inherent visual restriction) might have been inadequately addressed by traffic signs, or the driver might have failed to recognize this hazard despite warning signs.

It seems likely that an ADF will not exhibit these two failure types and thus can avoid the impending rear-end collision—assuming adequate object detection by ADF. On the one hand, technology is always “attentive” (a). On the other hand, automatic adjustment of speed (b) is likely to be part of ADF and thus also reduces the accident risk in this scenario. In a virtual experiment, the safety effects of ADF in this scenario can be compared to the performance of human drivers in various reference situations, e.g., with and without emergency braking assistance.

### 20.4.3 Expanding the Spectrum of Safety-Relevant Scenarios

In general, automation has not only a potential for selective accident prevention from the perspective of the ADF vehicle, but could, given sufficient penetration, increase overall traffic safety due to collective effects such as harmonization of traffic flow (see, e.g., [32]). For example, traffic literature shows (e.g., [4]) that inappropriate speed is not only limited to individual drivers but may be a collective phenomenon of the traffic stream. A high penetration of traffic with ADF vehicles could thus also avoid accidents for non-equipped vehicles (due to the collective effects of early speed adjustment).

On the other hand, due to the continuous action of ADF in traffic, the spectrum of relevant scenarios is considerably larger than for most ADAS. As a consequence, fundamentally new issues and methodological challenges arise for the virtual safety assessment of ADF.

In general, traffic has a very high complexity—due to the numerous direct interactions between traffic participants and indirect interactions between individuals and the collective traffic flow. Nevertheless, traffic flow has several collective or “macroscopic” characteristics. Examples are “fundamental diagram” (empirical relationship between traffic flow and average speed on a freeway section) or “capacity” (characteristic traffic demand above which traffic flow tends to become unstable). Changes in macroscopic characteristics of traffic flow, in this context, due to ADAS or ADF, may in turn have effects on traffic safety.
Automated vehicles can affect direct interaction between vehicles, interaction between individual vehicles and traffic flow, as well as collective characteristics of traffic flow. Since all of these changes can affect conflict and accident probability, they must be included in traffic safety assessment.

To this end, many processes in normal, non-assisted, and nonautomated traffic will need to be reassessed. On the one hand, particularly relevant traffic processes include those in which potential conflicts are normally avoided by human anticipation, intuition, cooperative driver behavior, strategic defensive driving, etc., and thus do not necessary require emergency action. Moreover, human drivers frequently have a very high context sensitivity, which further increases traffic safety performance. These skills are a typical strength of human drivers (see also [6, 12, 18, 22, 25, 26]).

An open question and subject of the research and development is to what extent automated vehicles will have capabilities comparable to those of human drivers? Some safety-relevant characteristics, such as anticipation and defensive driving, could be even more pronounced with ADF than with human drivers.

ADF could have an impact on cooperative interactions between automated vehicles and individual human drivers. Cooperative behavior between human drivers can be illustrated on a freeway with a “keep-right” rule: consider a vehicle driving on the right lane behind a slow truck. On the adjacent lane, another driver recognizes this situation and strategically creates a gap to “let in” the other vehicle. The driver on the left lane intentionally slows down slightly; the other driver recognizes his intention, signals, and begins his lane change while monitoring the gap. Ideally, he can complete the lane change with minimal impact on steady traffic flow.

The ADF vehicle on the left lane must be able to complete two important tasks in order to fulfill the example of the lane change: first, it must recognize the situation of the vehicle on the right lane, and second, it must have an appropriate cooperative action strategy. Whereas cooperation between human drivers can be influenced by factors such as emotion or time pressure, these factors are hardly relevant for automation.

The rate of success of cooperative interactions between human drivers and their sensitivity to context is still subject of transport research. In addition, any influences due to automation are currently being explored. The modeling of cooperative processes in traffic flow and road safety is a current research topic that presents further challenges for simulation.

It is also important to model indirect interactions with impacts on cooperative driving. For example, in dense freeway traffic, automated vehicles might tend to keep larger time gaps than are currently common and thus might even be recognizable as automated vehicles. These larger time gaps could conceivably encourage individual human drivers to cut in more frequently. If automated vehicles react with corresponding braking decelerations to vehicles cutting in, traffic dynamics will change. The occupants of an automated vehicle may even get the subjective impression of an increase in travel time, which would theoretically affect acceptance and thus indirectly traffic safety. A high penetration of ADF might also lead
to macroscopic consequences (fundamental diagram, capacity, platoon formation, frequency of shocks in traffic flow, etc.).

However, objective increases in travel time of automated vehicles are expected to be imperceptibly small and likely to be accepted. Moreover, the process of cutting in during dense “synchronized” traffic is not very rewarding anyways. In general, the presumption seems natural that the overall effectiveness of automated functions regarding traffic safety will be positive. The objective is a quantification of this hypothesis by comprehensive assessment.

### 20.4.4 Philosophy and Procedural Approaches for Validation and Assessment of Automation

While design of ADAS usually involves balancing desired and unintended system actions with regard to traffic safety, more complex trade-offs between conflicting objectives can be expected in the development and design of ADF.

A proportion of positive contributions to the safety record of automated functions can be estimated using virtual experiments (analog to ADAS) in which existing, potentially hazardous traffic scenarios are investigated. However, due to the continuous operation of ADF, the overall safety performance could be strongly influenced by additional scenarios and constellations, taking into account low-probability events. These scenarios are not known a priori and may be “hidden” in the situation space. Furthermore, their relevance can also depend on context or function design (e.g., partial or high automation).

For example, the reaction of automation to lane keeping of other drivers might depend on local driving characteristics. If an automation function is designed for a traffic context with “very precise” lane keeping, as is common in parts of Central Europe, application to regions with less precise lane keeping could require an adaptation to local driving strategy. Otherwise, undesirable effects, for example, frequent braking in response to “vehicles cutting in,” such as platoon formation, could increase.

The development of ADF presents considerable additional challenges: the situation space will be substantially larger than for current ADAS, and prediction of all relevant situations will be difficult. These issues complicate the assessment and optimization of key system characteristics, such as stability, robustness, and safety impact. A possible approach is an integrated and agile development and assessment process using a comprehensive tool chain.

Any methodology for continuous safety assessment during the development requires a comprehensive understanding of various existing methods and their specific role and contribution in such a complex process. Figure 20.7 provides an abstract overview of the roles of different testing instances within an integrated process.
The core of the process is virtual simulation, i.e., software in the loop. The virtual test run can be executed for each major functional step in the development phase and tested against the generated set of scenarios. These virtual test runs are held on system level and identify the most critical, relevant, and failed scenarios which require a more detailed evaluation with a specialized testing instance (e.g., driving simulator, hardware in the loop, fleet testing, etc.).

Data and findings from empirical testing methods are used for modeling, as common for ADAS assessment. In particular, data from NDS, FOT, fleet tests, etc., could be used to form a database including an appropriate (possibly country-specific) model of exposure, behavior, and other key aspects.

Referring to Fig. 20.7, the scenario database as well as other models plays a key role for all development and test tools. The scenarios and their frequency (in terms of an exposure model) can be stored in a scenario database and reused for different test instances. In addition to traditional research methods (e.g., theoretical risk assessments, testing on the road, etc.), virtual continuous simulation offers new opportunities for “discovery” of relevant scenarios, especially, when it comes to combinations of factors which are rare and difficult to derive from the theory.

One approach to discovery of “unknown” scenarios is observing longer durations or distances by long-running continuous simulation: safety-critical scenarios are not explicitly generated (e.g., by certain given constraints), but arise spontaneously from a stochastic comprehensive model of traffic flow.
This approach implies challenging requirements for comprehensive modeling. These include reproduction of all relevant processes (also error processes) in traffic flow, with their respective frequencies. Modeling the frequencies of all relevant processes and scenarios essentially constitutes an advanced exposure model.

An initial set of scenarios can be specified using expert knowledge, field operational tests, and virtual test runs. Virtual testing generates many representations of stochastic processes in traffic based on models of traffic contexts, sensors, drivers, vehicles, and traffic dynamics. The objective is to provide a representative sample of the overall situation space taking into account the large number of potential scenarios including low-probability events [38].

In a virtual test operation, these scenarios could be checked automatically representing a kind of safety cycle. The frequency of scenarios could depend on different factors, for example, countries or environmental conditions. Virtual testing would fulfill the requirements of a safety assessment in this construct, as described above.

Virtual testing by simulation of a \textit{single} scenario results in quantification of effectiveness of an automated system in \textit{this} particular scenario. The safety performance of automation then results from the sum of effectiveness in \textit{all} relevant scenarios weighted by their respective frequency (exposure).

A key issue concerns the validation of process models and, by extension, plausibility of simulation results. Validation of models involves utilization of appropriate testing procedures for each particular method in the development chain. Each method, for instance, test driving or a driving simulator, is used for validating the vehicle model or for MMI concepts. A validated model database increases the reliability of virtual testing; the quality of the models is of key importance for the development chain and the validity of the assessment result.

Verification of simulated system actions represents another important element of the validation process, by drawing samples from all simulations and testing these in recognized test institutes. Validation (and development) of process models could be accompanied by impartial scientific organizations.

Consequently, the objective would be an international consensus and implementation of scenarios, models, and the overall assessment approach by all relevant stakeholders in an international context.

### 20.5 Conclusion and Outlook

The task of safety assessment and optimization of automated functions raises new issues. In contrast to ADAS assessment, quality measures of traffic safety are principally related to all traffic scenarios in which a function is active. Since automation may change collective traffic characteristics, safety analysis must go beyond isolated human errors in currently occurring traffic processes and the impact of automation on these. Newly emerging, automation-related, scenarios have to be considered for a comprehensive safety assessment.
Validation and safety assessment of automated functions have to be understood as continuous and iterative tasks during the development, not as singular activities at the end of the development phase. Due to the variety of possible influences, the necessary assessment of automation approaches during the development would be extremely problematic based, for example, solely on fleet testing, since detection of rare effects requires correspondingly long observation periods. In addition, testing would have to be repeated, in principle, after every single change of the function.

The approach of simulation-based virtual experiments can be interpreted as knowledge synthesis. Still, some challenges for the assessment of automated systems arise. Relevant scenarios for automation are a priori unknown and can only partially be identified using existing methods. Due to the generally larger situation space involved in automation, modeling results in considerably more complexity as for the assessment of ADAS.

Quality requirements for traffic simulation are correspondingly higher, especially in terms of process models used. Traffic simulation will need to consider error processes and their resolution in normal traffic in more depth. The challenges include improved modeling of psychological processes, e.g., attention or activation (Yerkes-Dodson) [37]. An important aspect of the safety potential of ADF arises from avoidance of errors resulting from lack of driver activation and resulting attention lapses.

Despite sophisticated technology, systems will still be subject to system limits within the near future. Virtual experiments could make an important contribution to design and optimization of take-over requests to human drivers, in addition to safety assessment.

Critical traffic situations can require a decision among several unfavorable alternatives for action. Here again, virtual assessment can support the development of transparent decision algorithms. A general discussion of such alternatives has already begun in public [30]. Possibly all stakeholders can achieve consensus and agree on guidelines for the prioritization of alternative actions before market introduction.

Many (novel) projects, initiatives, organizations, and research activities are focusing on the effects of ADAS regarding traffic safety. So far, however, an international consensus on methodological issues in the context of an overall safety assessment of ADAS and ADF is still lacking.

Considering the importance and complexity of decisions and challenges, the initiative “Prospective Effectiveness Assessment for Road Safety” (PEARS) has the objective of developing a standardized and harmonized method for the overall effectiveness assessment of new systems, such as ADAS or ADF, which is accepted by all stakeholders [29]. Both benefits and potential risks should be quantified as part of the assessment. The objectives are, among others, a higher degree of legal predictability and adequate and objective consideration of individual and societal interests. This open platform provides important prerequisites for a global harmonization and standardization.
References

1. M. Aeberhard, S. Rauch, M. Bahram, G. Tanzmeister, J. Thomas, Y. Pilat, F. Homm, W. Huber, N. Kaempchen, Experience, results and lessons learned from automated driving on Germany’s highways. IEEE Intell. Transport. Syst. Mag. 7(1), 42–57 (2015)

2. M. Bahram, Z. Ghandeharioun, P. Zahn, M. Baur, W. Huber, F. Busch, Microscopic Traffic Simulation Based Evaluation of Highly Automated Driving on Highways, in 17th International IEEE Conference on Intelligent Transportation Systems - ITSC 2014, ed. by IEEE (IEEE, 2014)

3. R. Bernotat, Anthropotechnik in der Fahrzeugführung. Ergonomics 13(3), 353–377 (1970)

4. T. Connolly, L. Åberg, Some contagion models of speeding. Accident Anal. Prevent. 25(1), 57–66 (1993)

5. R.J. Cook, D.L. Sackett, The number needed to treat: A clinically useful measure of treatment effect. Br. Med. J. 18, 452–454 (1995)

6. C. Domsch, W. Huber, Integrale Sicherheit - ein ganzheitlicher Ansatz für die Fahrzeugsicherheit. 17 (Aachener Kolloquium Fahrzeug- und Motorentechnik, Aachen, 2008)

7. A. Eskandarian (ed.), Handbook of Intelligent Vehicles (Springer, Heidelberg, 2012)

8. F. Fahrenkrog, A. Zlocki, L. Eckstein, Bewertung Aktiver Sicherheit: Vom Test zur Wirksamkeitsanalyse. ATZ 116(11), 34–39 (2014)

9. I. Ferenčz, T. Helmer, P. Wimmer, R. Kates, Effectiveness Analysis and Virtual Design of Integrated Safety Systems Illustrated Using The Example Of Integrated Pedestrian And Occupant Protection, in 12th International Symposium and Exhibition on Sophisticated Car Occupant Safety Systems (Fraunhofer-Institut für Chemische Technologie ICT, 2014)

10. M. Green, “How long does it take to stop?” Methodological analysis of driver perception-brake times. Transport. Hum. Factors 2(3), 195–216 (2000)

11. T. Helmer, Development of a Methodology for the Evaluation of Active Safety using the Example of Preventive Pedestrian Protection, Thesis (Springer, Heidelberg, 2015), ISBN 978-3-319-12888-7

12. T. Helmer, T. Kühbeck, C. Gruber, R. Kates. Development of an Integrated Test Bed and Virtual Laboratory for Safety Performance Prediction in Active Safety Systems (F2012-F05-005), in FISITA 2012 World Automotive Congress - Proceedings and Abstracts, 2012 (ISBN 978-7-5640-6987-2)

13. T. Helmer, M. Neubauer, S. Rauscher, C. Gruber, K. Kompass, R. Kates, Requirements and Methods to Ensure a Representative Analysis of Active Safety Systems, in 11th International Symposium and Exhibition on Sophisticated Car Occupant Safety Systems (Fraunhofer-Institut für Chemische Technologie ICT, Pfinztal, 2012), pp. 6.1–6.18, ISSN 0722-4087

14. T. Helmer, R.R. Samaha, P. Scullion, A. Ebner, R. Kates, Injury risk to specific body regions of pedestrians in frontal car crashes modeled by empirical, in-depth accident data, in Proceedings of the 54th Stapp Car Crash Conference, 2010

15. T. Helmer, R.R. Samaha, P. Scullion, A. Ebner, R. Kates. Kinematical, Physiological, and Vehicle-Related Influences on Pedestrian Injury Severity in Frontal Car Crashes: Multivariate Analysis and Cross-validation, in Proceedings of the International Research Council on Biomechanics of Injury (IRCOBI) (IRCOBI, Hannover, 2010), pp. 181–198

16. T. Helmer, P. Scullion, R.R. Samaha, A. Ebner, R. Kates, Predicting the injury severity of pedestrians in frontal vehicle crashes based on empirical, in-depth accident data. Int. J. Intell. Transport. Syst. Res. 9(3), 139–151 (2011)

17. W.E. Hick, On the rate of gain of information. Q. J. Exp. Psychol. 4, 11–26 (1952)

18. W. Huber, J. Steinle, M. Marquardt, Der Fahrer steht im Mittelpunkt - Fahrerassistenz und Aktive Sicherheit bei der BMW Group, in 24. VDI/VW-Gemeinschaftstagung - Integrierte Sicherheit und Fahrerassistenzsysteme, vol. 2048 VDI-Berichte (VDI Verlag, Düsseldorf, 2008), pp. 123–137
19. C. Hydén, The Development of a Method for Traffic Safety Evaluation: The Swedish Traffic Conflicts Technique, Bulletin 70 (Department of Traffic Planning and Engineering, Lund University, Sweden, 1987)

20. R. Kates, O. Jung, T. Helmer, A. Ebner, C. Gruber, K. Kompass, Stochastic Simulation of Critical Traffic Situations for the Evaluation of Preventive Pedestrian Protection Systems (VDI-Tagung Erprobung und Simulation in der Fahrzeugentwicklung, Baden-Baden, 2010)

21. D.G. Kleinbaum, M. Klein, Logistic Regression. A Self Learning Text. Statistics for Biology and Health (Springer, Heidelberg, 2010)

22. K. Kompass, C. Gruber, C. Domsch, Der Beitrag von Fahrerassistenzsystemen zur Aktiven und Passiven Sicherheit - die Integrale Sicherheit als Antwort auf die wachsenden Anforderungen an die Fahrzeugsicherheit, in 4. Tagung Sicherheit durch Fahrerassistenz, 2010.

23. K. Kompass, T. Helmer, C. Blaschke, R. Kates, Ganzheitliche und integrale Fahrzeugsicherheit. ATZ - Automobiltechnische Zeitschrift 116, 10–15 (2014)

24. K. Kompass, T. Helmer, L. Wang, R. Kates, Gesamthafte Bewertung der Sicherheitsveränderung durch FAS/HAF im Verkehrssystem: Der Beitrag von Simulation, in Fahrerassistenz und Aktive Sicherheit. Wirksamkeit – Beherrschbarkeit – Absicherung (expert Verlag, Rennlingen, 2015), pp. 45–66

25. K. Kompass, W. Huber, Integrale Sicherheit – Effektive Wertsteigerung in der Fahrzeugsicherheit (VDA, Technischer Kongress, Berlin, 2009)

26. K. Kompass, W. Huber, T. Helmer, Safety and Comfort Systems: Introduction and Overview, in Handbook of Intelligent Vehicles, ed. by A. Eskandarian (Springer, Heidelberg, 2012)

27. N. Lerner, Brake Perception–Reaction Times of Older and Younger Drivers, in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 37 (1993), pp. 206–210

28. P.L. Olson, M. Sivak, Perception-response time to unexpected roadway hazards. Hum. Factors 28, 91–96 (1986)

29. Y. Page, F. Fahrenkrog, A. Fiorentino, J. Gwehenberger, T. Helmer, M. Lindman, O. op den Camp, L. van Rooji, S. Puch, M. Fränzle, U. Sander, P. Wimmer, A Comprehensive and Harmonized Method for Assessing the Effectiveness of Advanced Driver Assistance Systems by Virtual Simulation: The P.E.A.R.S. Initiative, in 24th International Technical Conference on the Enhanced Safety of Vehicles (ESV 2015), number 15-0370, 2015

30. H. Prantl, Error. Süddeutsche Zeitung 37, 13 (2015)

31. E. Schechtman, Odds ratio, relative risk, absolute risk reduction, and the number needed to treat – which of these should we use? Value Health 5(5), 431–436 (2002)

32. C. Steinhoff, R. Kates, H. Keller, B. Färber, B. Färber, Problematik präventiver Schaltungen von Streckenbeeinflussungsanlagen (BMVBW Forschung Straßenbau 853; Bundesanstalt für Straßenwesen, 2001)

33. A. Weitzel, H. Winner, C. Peng, S. Geyer, F. Lotz, L. Sefati, Absicherungsstrategien für Fahrerassistenzsysteme mit Umfeldwahrnehmung (BAST-Bericht F98, Bundesanstalt für Straßenwesen, 2014)

34. R. Wertheimer et al., Ko-PER – Fahrerassistenz und präventive Sicherheit mittels kooperativer Perzeption: Partnerübergreifender Schlussbericht, Schlussbericht (Bundesministerium für Wirtschaft und Technologie (BMWi), 2014)

35. P. Wimmer, A. Rieser, W. Sinz. A Simulation Method for Virtual Design and Evaluation of Integrated Safety Systems, in NAFEMS World Congress, 2013

36. H. Winner, S. Hakuli, G. Wolf (eds.), Handbuch Fahrerassistenzsysteme (Vieweg+Teubner Verlag/GWV Fachverlage GmbH, Wiesbaden, 2009)

37. R.M. Yerkes, J.D. Dodson, The relation of strength of stimulus to rapidity of habit-formation. J. Comp. Neurol. Psychol. 18, 459–482 (1908)

38. A. Zlocki, L. Eckstein, F. Fahrenkrog, Evaluation and sign-off methodology for automated vehicle systems based on relevant driving situations, in Transportation Research Board (TRB), 94th Annual Meeting, Washington DC, USA, 11–15 Jan 2015