Verband der Automobilindustrie

Study

Standardization-Project

Safety requirements for Steer-by-Wire-Steering Systems

The VDA-based DIN Working Group NA 052-00-33-18 AK „Safety requirements for a steer-by-wire steering system“ was tasking fka with the development of the attached scientific study to be used as the basis for a future DIN standard.

The following companies and organizations have contributed to the financing of the preparation of this study:

AUDI AG

BMW AG

Bosch
Automotive Steering GmbH

Ford-Werke GmbH

HELLA GmbH & Co. KGaA

JTEKT Europe

HL Mando Corporation
Europe GmbH

Mercedes-Benz AG

Nexteer Automotive Germany GmbH

Schaeffler Technologies AG & Co. KG

ThyssenKrupp Presta AG

Volkswagen AG
PKW

Volkswagen Group Components

ZF Automotive
Germany GmbH

Conducting subject studies for the DIN working group Steer-by-Wire

Report 217750

fka GmbH - Chassis & NVH

ika, RWTH Aachen University - Traffic Psychology & Acceptance

Final report

Conducting subject studies for the DIN working group Steer-by-Wire

Project number

217750

Client

DIN NA 052-00-33-18 Working Group “Steer-by-Wire”, hosted by VDA

Project manager:

Dr.-Ing. Daniel Wegener

Dr. phil Stefan Ladwig

Project engineer:

Lotte Saupp M.Sc.

Julia Pelzer M.Sc.

Dipl.-Ing. Jörn Lützow

Dipl.-Ing. Tobias Sandmann

Anna-Lena Köhler M.Sc.

Dr.-Ing. Daniel Wegener
Head of Chassis & NVH

Univ.-Prof. Dr.-Ing. Lutz Eckstein
Chairman of the Advisory Board

Dr.-Ing. Jens Kotte
Managing Director

Aachen, November 2022

Contents

1 Introduction and objectives

Steer-by-wire technology opens up new possibilities for redesigning the interaction between driver and vehicle. This enables the development of new driving functions in the context of semi-automated to fully automated driving. In the future, there will be more alternation between (partially) automated and manual driving sections. Therefore, innovative steering concepts will become increasingly important as they provide the driver with more functionality during (partially) automated driving while ensuring intuitive handling.

Compared to conventional mechanical steering, a steer-by-wire system offers novel safety-relevant aspects. Such a system has to be designed in a way that *fail-operational* execution is guaranteed. Essentially, this involves the occurrence of specified failure patterns that have to be addressed by suitable fallback strategies. A distinction shall be made between failure at component level and effects at system as well as overall vehicle level, which have to be controlled by the driver (controllability). For example, a failure at component level could lead to changes in wheel steering angle at system level. The result would be an overall vehicle reaction depending on the driver´s compensatory as well as the vehicle´s behavior. A safety criterion regarding the overall vehicle reaction could be, e.g. the lane deviation in a driving situation.

However, the number of possible influencing factors that may affect a variety of driving scenarios makes it impossible to run holistic controllability studies with normal drivers. Therefore, the system design is developed by experts, traditionally, using the multiple-eye principle, by approximating the aspect of controllability in a practice-oriented manner. As a result, they do finally define a corresponding Automotive Safety Integrity Level (ASIL, ISO 26262). In accordance with ISO 26262, controllability is characterized by the parameter C, in levels from C0 - C3. Whilst C0 stands for "Controllable in general", C1 for "Simply controllable", C2 for "Normally controllable" and C3 for "Difficult to control or uncontrollable". Furthermore, C0 is being defined by the addition "Distracting". The corresponding definitions of all levels are shown in Table 1.

Table 1
Level of controllability according to IS0 26262

Class	C0	C1	C2	C3
Description	Controllable in general	Simply controllable	Normally controllable	Difficult to control or uncontrollable
Definition	Distracting	More than 99% of average drivers or other traffic participants are usually able to control the damage.	More than 85% of average drivers or other traffic participants are usually able to control the damage.	The average driver or other traffic participant is usually unable, or barely able, to control the damage.

The classification into four levels from C0 - C3 according to ASIL largely coincides with a complementary form of classification: The Response 3 - Code of Practice (Brockmann, 2009). This widely accepted framework deals with the assessment of the likelihood that drivers [...] are unable to control a challenging event and are therefore unable to avoid injury or damage. Brockmann (2009) also do represent C0 to be the lowest level with the label " Controllable in general".

The expert-based approach also includes highly dynamic driving situations that adequately represent the effects of a possible system failure situation. The assessment of controllability by normal drivers in the corresponding situations is carried out by experts using the levels described above. From a scientific point of view, this raises the question of the extent to which a system failure effect, parameterized by expert judgements, could gain empirical validity regarding transferability to normal drivers.

In addition to the parameterization of the failure intensity in anticipation of the controllability of average drivers, which must be proven and determined by experts, human factors such as acceptance, trust and subjectively perceived and objectively observed controllability are also of particular importance regarding the transition from a potential system failure to a degraded system state. To our knowledge, there are no reliable, publicly accessible studies that could provide precise information in this regard. However, literature indicates that the psychological constructs of acceptance and trust vary systematically with subjectively perceived system reliability (Numan, 1998; Chavaillaz et. al., 2016). Furthermore, the relationship between acceptance and trust on the one hand and system behavior on the other is considered to be very well studied in various areas of the automotive sector. Especially, when it comes to empirical studies in the field of automated driving (see for example Numan, 1998; Pavlou, 2003; Ghazizadeh et. al. 2012b; Hegener et. al. 2019). Hegener et al. (2019) investigated, amongst other things, the relationship between trust in technology and the intention to hand over control to automated vehicles with confidence. The results indicate a high correlation of trust in technology and any concerns about handing over control to an automated vehicle. Since a possible perceptible failure in the steering system of a vehicle has the potential to representing an objective aspect of the (short-term) loss of control, a corresponding correlation with trust and acceptance shall be expected.

Thus, scientifically speaking another question addresses potential connections between system failure characteristics parameterized by expert judgments and subjective constructs of acceptance and trust of normal drivers.

Against this background, the present project is dedicated to conducting *N* = 6 individual studies to investigate the controllability and trust/acceptance of a steer-by-wire system, parameterized by experts, regarding specified failure patterns and the respective fallback strategies of the system. The *N* = 6 individual studies are based on the following common hypothesis:

> "A C0 level defined by experts with regard to a failure characteristic in a steer-by-wire steering system does not lead to driver behavior in predefined driving maneuvers for normal drivers that involves leaving a driving corridor specified by the expert within a defined period of time from failure activation."

Testing this hypothesis is the basis of the present series of tests. The working group determined the selection of vehicles, maneuvers and failure patterns and communicated these to the authors of this report. A corresponding study design was then developed in close cooperation with the Steer-by-Wire working group. The methodology will be presented in detail below.

2 Methodological approach

The procedure for examining the information specified in section 1 is focus of the following section. Overall, the research project is divided into six individual studies, which were conducted on the vehicle dynamics area of the Aldenhoven Testing Center between March 28th, 2022, and May 20th, 2022. For each individual study, one vehicle, provided by one of the participating companies, was examined with two different types of failures. The hypothesis test was carried out individually for each vehicle-failure combination, so that the test was carried out in a total of N = 12 cases. The derivation of the selected methodology as well as the detailed structure and procedure of the studies are explained below.

2.1 Study design - individual studies

The aim of the research project was to investigate the controllability and the driver's reaction to various failures in a steer-by-wire steering system. With the help of the data collected, the aim was to test whether a C0-level set by experts does not result in driver behavior during predefined driving maneuvers that involves leaving a driving corridor specified by the expert within a defined period of time after failure activation. As can be seen from the hypothesis both, the failure patterns to be investigated and the associated driving maneuvers are specified by the experts on the working group. This specification comprised a total of N = 6 failure patterns, three failures of the *feedback actuator (FBA)* and three failures of the *roadwheel actuator (RWA)*, the controllability of which was to be investigated as part of the study. The working group selected one of three possible driving maneuvers (straight-ahead, circular drive, slalom) for each failure pattern. The individual failure patterns and their combination with the selected driving maneuvers are described in sections 2.2 and 2.3 in detail.

The systematic investigation of the six failure patterns as independent variables required the development of an experimental design that took several aspects into account. First of all, the presentation of multiple failure events in one test drive per test subject did not appear to make sense. The reason for this is the necessity of the element of surprise for a valid measurement of the driver's initial, natural reaction to a failure event. With the presentation of several failure patterns, it was to be assumed that the test subjects would develop an attitude of expectation in the course of the test and thus possibly react in a biased and different way to a failure activation than they would in an unprepared situation. This is accompanied by a strong sequence effect, which can also influence the results. In order to avoid these confounding factors and still take into account the aspect of test economy, two failure patterns were combined in each case, which were tested in randomized order for each test subject. This combination of failure patterns was also carried out by the experts of the working group. Care was taken to combine one RWA and one FBA failure each, which differ from each other in regard to their perceptibility of the driver and are activated in different driving maneuvers. This resulted in both, within-subject factors for the test design, as one test subject experienced two failures within one test drive, and between-subject factors, as the total of three FBA/RWA failure combinations were divided between different groups of test subjects.

In addition to the factor of the different failure patterns, the adjustment of the parameterization to the affected vehicle also plays a decisive role in the noticeability of the failure activation. In order to take this factor into account in the study, a total of six different vehicles from the participating companies in the working group were included in the test design and combined with the failure pairs. This combination was carried out in such a way that each failure pair was examined in two vehicles of different vehicle classes (SUV, compact class, sedan). In the following these are referred to as SUV 1, SUV 2, compact class 1, compact class 2, compact class 3 and sedan. The parameterization of the failures was carried out in advance of the study by the experts of the respective company. The combination of the aforementioned factors - failure patterns with associated driving maneuvers and selected vehicles - was combined to form the overarching test design, which is shown in Illustration 1.

Illustration 1

Superordinate experimental design

<table>
<tr><th>Vehicle</th><th>Failure pattern A
– Maneuver –
Blocked FBA
Slalom</th><th>Failure pattern B
– Maneuver –
RWA square-wave oscillation
Straight ahead driving</th><th>Failure pattern A
– Maneuver –
FBA step
Slalom</th><th>Failure pattern B
– Maneuver –
Uncontrolled RWA
Circular driving</th><th>Failure pattern A
– Maneuver –
FBA Selfsteer + Loss of feedback
Circular driving</th><th>Failure pattern B
– Maneuver –
Blocked RWA
Slalom</th></tr>
<tr><td>SUV 1</td><td></td><td></td><td></td><td></td><td colspan="2">n = 28</td></tr>
<tr><td>Compact 1</td><td></td><td></td><td></td><td></td><td colspan="2">n = 28</td></tr>
<tr><td>Compact 2</td><td></td><td></td><td colspan="2">n = 28</td><td></td><td></td></tr>
<tr><td>Sedan</td><td></td><td></td><td colspan="2">n = 28</td><td></td><td></td></tr>
<tr><td>Compact 3</td><td colspan="2">n = 28</td><td></td><td></td><td></td><td></td></tr>
<tr><td>SUV 2</td><td colspan="2">n = 28</td><td></td><td></td><td></td><td></td></tr>
</table>

How illustration 1 shows the test design involves six individual studies, each of which examines two combinations of vehicle, failure pattern and driving maneuver. The established research hypothesis is tested individually for each of these twelve combinations in the course of the data analysis. The criterion for deciding whether to accept the hypothesis was chosen based on Brockmann (2009). According to Brockmann, a hypothesis is retained if 100% of the test subjects were able to continue to control the vehicle and not leave the lane specified by the maneuver after the failure activation. This 100% must include at least n = 20 test subjects, whose data sets are analyzed according to a priori defined criteria (see section 2.7) are considered valid. As soon as a test subject leaves the lane as a result of the failure, the hypothesis is rejected. In order to ensure that the minimum number of 20 valid data sets is achieved in each individual study and that any losses due to dropouts, no-shows and criterion-based exclusion of data sets can be compensated for, n = 28 test subjects were invited to the tests in each case. The total size of the planned sample of all six individual studies therefore amounted to N = 168 test subjects.

2.2 Failure types

The steer-by-wire failure patterns used in the studies, which were examined with regard to their controllability and the driver's reaction, were defined in type and parameterization by the experts in the working group. A total of six different failure types, including three *feedback actuator* failures (FBA failures) and three *road-wheel actuator* failures (RWA failures), were included in the study. In addition, prior to testing in the test subject study, the parameterization of the failure patterns for the specific vehicle was jointly coordinated by several experts according to the multiple-eye principle. The individual failure patterns for the FBA and RWA failures are shown in Table 2 and Table 3 are described.

Table 2
FBA failure types

Failure pattern (driving maneuver)	Vehicle	Description of failure pattern
Blocked FBA (slalom 60 km/h)	Compact Class 3	▪ Blocking the FBA in the current angular position ▪ Trigger point: -15° FBA steering angle (steering to the right) ▪ Failure duration: 100 ms
	SUV 2	▪ Blocking the FBA in the current angular position (holding the steering angle with 25 Nm) ▪ Trigger point: 5° FBA steering angle (steering to the left) ▪ Failure duration: 200 ms
FBA step (slalom 60 km/h)	Compact Class 2	▪ Torque step on the FBA (reduction of the current steering torque on the FBA by 1.5 Nm) ▪ Trigger point: 0° FBA steering angle (when deflecting from right to left, second pylon) ▪ Failure duration: until vehicle comes to a standstill (manual trigger point)
	Sedan	▪ Torque step on the FBA (increase of the current steering torque on the FBA by 2.2 Nm) ▪ Trigger point: third pylon (left turn) ▪ Failure duration: 10 s
FBA Selfsteer + Loss of feedback (circular drive 50 km/h)	SUV 1	▪ Torque step at the FBA (23.5 Nm for 10 ms) and subsequent damping ▪ Trigger point: manual at a predefined pylon ▪ Failure duration: torque step 10 ms, damping until standstill
	Compact Class 1	▪ Torque step at the FBA (6-8 Nm for 20 ms) and subsequent damping (0.01 Nm/deg/s) ▪ Trigger point: manual at a predefined pylon ▪ Failure duration: Torque step 20 ms, damping until standstill

Table 3
RWA failure types

Failure pattern (driving maneuver)	Vehicle	Description of failure pattern
RWA square-wave oscillation (straight-ahead 80 km/h)	Compact Class 3	▪ Self-steering of the RWA (square-wave sine, 3 periods of 700 ms each) ▪ Trigger point: manually during straight-ahead driving ▪ Failure duration: 2100 ms
	SUV 2	▪ Self-steering of the RWA (square-wave sine, 3 periods of 700 ms each) ▪ Trigger point: manually during straight-ahead driving ▪ Failure duration: 2100 ms
Uncontrolled RWA (circular drive 50 km/h)	Compact Class 2	▪ Torque step to 0 Nm RWA torque (RWA steering angle free) ▪ Trigger point: manual with defined pylon ▪ Failure duration: 150 ms
	Sedan	▪ Torque step to 0 Nm RWA torque (RWA steering angle free) ▪ Trigger point: manual with defined pylon ▪ Failure duration: 150 ms
Blocked RWA (slalom 60 km/h)	SUV 1	▪ Blocking the RWA in the current angular position ▪ Trigger point: 20° FBA steering angle (steering to the left) ▪ Failure duration: 200 ms
	Compact Class 1	▪ Blocking the RWA in the current angular position ▪ Trigger point: 15° FBA steering angle (steering to the left) ▪ Failure duration: 150 ms

2.3 Maneuvers, driving task and course

The driving maneuvers in which the selected failure patterns were activated were defined by the working group in advance of the study. Three driving maneuvers relevant to the respective failure patterns were selected:

- a straight-ahead driving at 80 km/h;
- a steady circular drive at 50 km/h and
- a 36 m slalom at 60 km/h.

The maneuvers were marked out with the help of pylons on a cordoned-off test track area (see Illustration 3). In addition, each maneuver was assigned a specific target speed, which was to be kept constant over the entire length of the maneuver. The guiding criterion here was to select a target speed that could be achieved by everyday drivers in the respective maneuvers with reasonable effort. Each of the six maneuvers described in section 2.2 was combined by the working group with one of these three maneuvers, during the passage of which the failure was to be activated. The resulting combinations (failure, target speed, maneuver) were selected in such a way that the occurrence of a failure triggers a noticeable effect for the drivers.

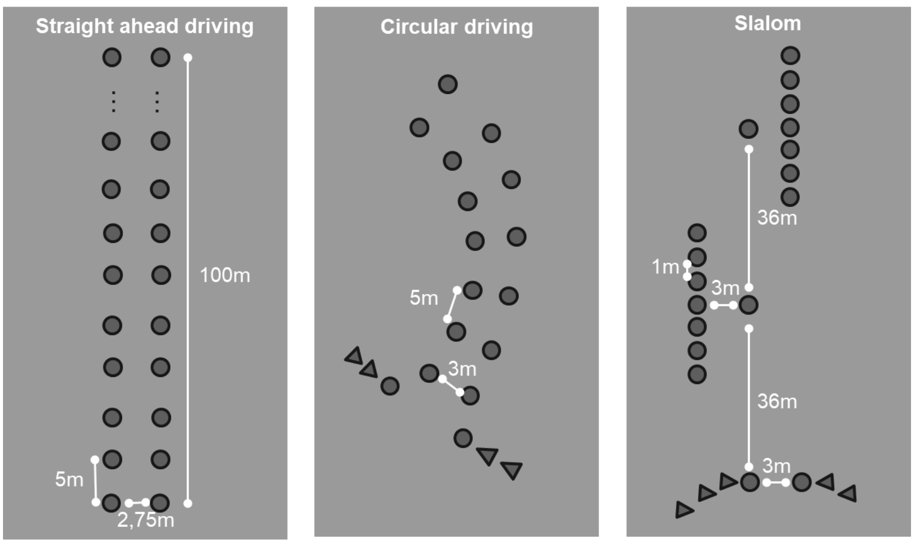
The straight-ahead driving was selected for the activation of the *RWA square-wave oscillation* failure. The maneuver was implemented using a straight lane with a total length of 100 m. The lane width of 2.75m was limited on both sides by pylons (between the two inner edges of the pylons' base). In the longitudinal direction, these pylons were placed at a distance of 5 m from each other (pylon center to pylon center). Illustration 2 shows a schematic representation of the structure of the straight-ahead drive. The recommended speed to be maintained when driving straight-ahead was 80 km/h.

The circular drive was defined as a maneuver for activating the failures *Uncontrolled RWA* and *FBA Selfsteer + Loss of Feedback.* A pylon lane with a radius of 60m and a lane width of 3m (inner edges) was installed here. In the longitudinal direction, the distance between the pylons was 5m (pylon center, inside of the circular drive), analogous to the circular drive described above. The resulting total length of this maneuver was 180m. The entrance to the circular drive was marked with horizontal pylons for better orientation for the test subjects. Illustration 2 shows a schematic drawing of the setup of the circular drive. The recommended speed to be maintained *was 50 km/h.*

In the slalom maneuver, the failures blocked FBA, FBA step and blocked RWA were activated. The slalom consisted of a total of 5 gates, which were to be driven through in a sinusoidal pattern. The entrance to the first gate was marked by horizontal pylons. The gates were marked out at a distance of 36 m (inside edges). A gate consisted of a pylon on the inside and seven pylons, each one meter apart (pylon center) in the longitudinal direction, which bordered the outside of the gate. The total length of the staked-out slalom maneuver was 160 m. A schematic representation of the individual maneuver structures is shown in Illustration 2. The recommended speed was 60 km/h.

Illustration 2

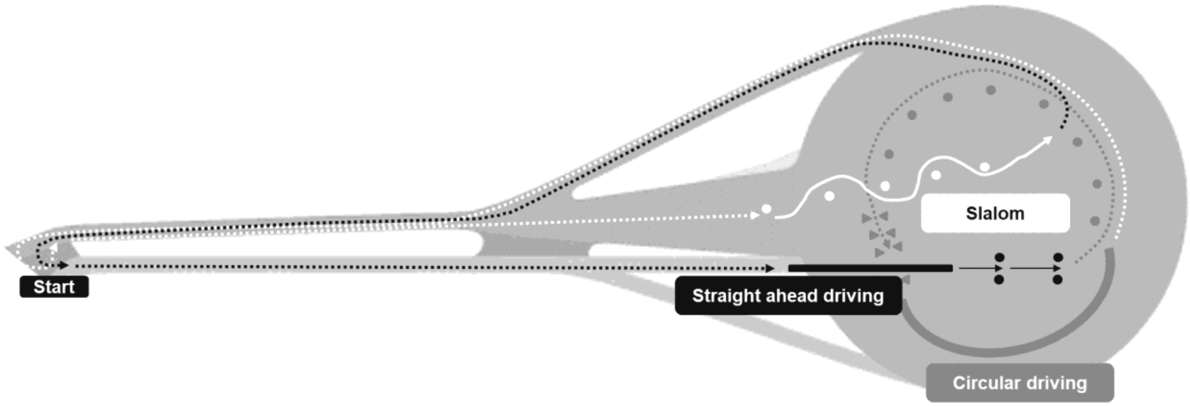
Schematic representation of the structure of the three driving maneuvers



Since the aim of the study was to subjectively and objectively record the initial reaction of normal drivers to an unexpected failure event, the time of the failure activation should not be predictable for the test subjects. For this reason, the maneuvers on the test track were combined to form a course. By completing several maneuvers, the timing of the failure activation was to be masked for the test subject. This served the purpose of maintaining an element of surprise, analogous to any occurrence in everyday use. It was also possible to ensure that the driving task remained constant for all test subjects. A bird's eye view of the layout of the course is provided by Illustration 3.

Illustration 3

Bird's eye view of the course



In the layout of the course on the test track, driving straight-ahead was the first maneuver to be completed. From the starting point (see Illustration 3 on the left), the aim was to accelerate to the required target speed of 80 km/h on an approx. 350 m long approach section. After driving straight-ahead at the required 80 km/h, two pylon gates indicated the braking points. The test subjects were asked to slow down when they reached the first pylon gate (approx. 33 m after driving straight-ahead) so that they could turn left after approx. 70 m at the second pylon gate with a residual speed of 20-30 km/h. In preparation for driving in a double circle, the test subjects were able to orient themselves to pylons that had already been set up in a semicircle with a radius of 60 meters. The test subjects were asked to accelerate to the recommended speed of 50 km/h to be maintained during the circular drive. The test subjects entered the second maneuver, the steady state cornering, through the entrance gate. Following this, the course led via a return (upper part of the Illustration 3) back to the starting point. For the final run through the slalom, a U-turn was performed and the middle access road was selected as the approach route (see Illustration 3, white dashed marking). The test subjects were able to reach the required speed limit of 60 km/h over a distance of 320 m before entering the first slalom gate. After driving through the maneuver, the course led back to the starting point via the return route. The total length of the course through all three maneuvers was approx. 2910 m.

2.4 Measured variables

Three types of measured variables were included in the study in order to assess the drivers' reaction to the activation of the various failures in the driving maneuvers. In order to evaluate the controllability of the failure situation, the lane deviation was operationalized first. In addition, objective vehicle parameters and subjective data were collected by interviewing the test subjects. The individual measured variables are described in more detail below.

2.4.1 Track deviation

The controllability of the failure situation for the drivers was operationalized via the deviation from the specified lane. Such a deviation occurred when test subjects touched (tire touched the base of a pylon) or drove over the pylons that limit the driving corridor of a maneuver. This criterion was measured in two ways. Firstly, the number of cones touched or driven over during the test was recorded by the test supervisor. Secondly, the logging was checked post-hoc using video recordings of the drive (see Illustration 4).

Illustration 4
Camera settings for video recording

Two cameras were mounted on both sides of the vehicle and aimed backwards towards the vehicle's wheels. A third camera filmed the drive from inside the vehicle looking forwards in the direction of driving. For data protection reasons, the camera was aligned so that only the hands of the test subjects could be seen (Illustration 9). The video recording was active for the entire duration of the drive so that all drives could be retraced from different perspectives in the subsequent evaluation if required. Illustration 4 shows the different camera settings.

2.4.2 Objective measured variables

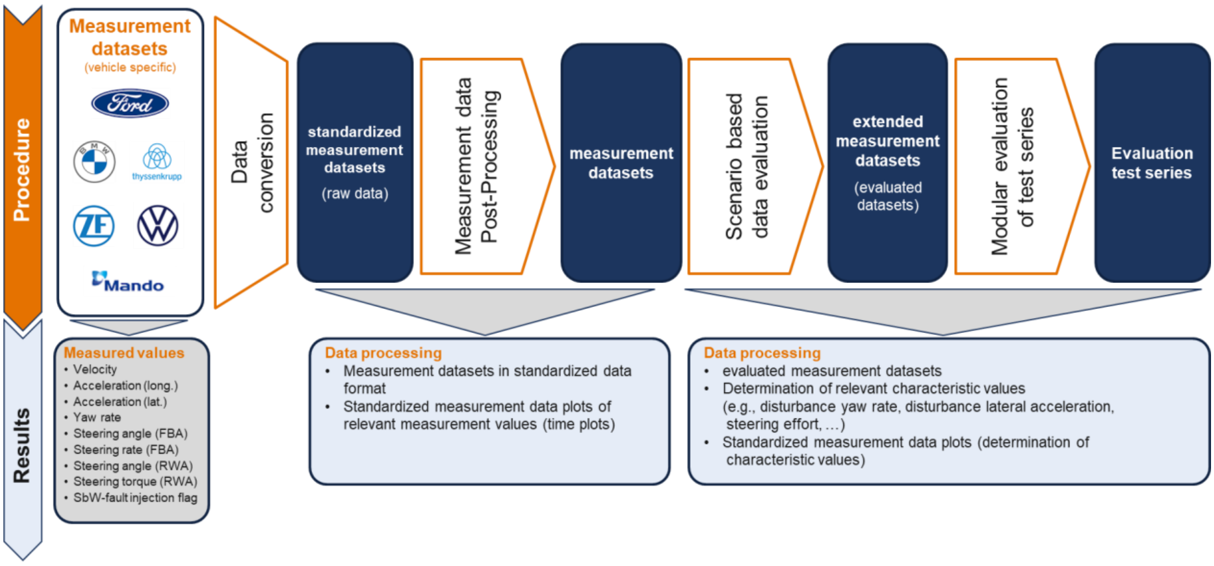
In order to objectively record the influences of the failure circuits on the driver inputs and the resulting driving dynamics of the respective test vehicle, selected vehicle measurement variables were recorded as part of the tests. The measurement data was recorded using the measurement technology integrated in the vehicles of the respective companies. The recorded measured variables are as follows:

- the vehicle speed;
- the longitudinal and lateral acceleration of the vehicle;
- the yaw rate;
- the control values of the *feedback actuator (FBA)* (steering angle, steering torque);
- the control values of the *roadwheel actuator (RWA)* (steering angle, steering torque);
- a failure trigger (recording the time of the failure activation).

The recording and subsequent provision of the vehicle measurement data was carried out during the tests of the respective study by the vehicle supervisors of the respective company. Depending on the company, the data was recorded in different measurement data formats and with different recording parameters (e.g. signal sampling rate, naming convention with regard to the measured variables). With the aim of harmonizing the recorded vehicle measurement data from all tests, downstream measurement data processing was carried out and the vehicle-specific measurement data sets were converted into a uniform measurement data format. The processed measurement data sets were then fed into a driving maneuver-dependent data evaluation. In the course of the data evaluation, relevant characteristic values were calculated and standardized evaluation plots of the relevant vehicle measurement variables were created. The sequence of the individual process steps of measurement data preparation and evaluation is shown in Illustration 5.

Illustration 5

Measurement data preparation and evaluation process for objective vehicle measurement data



The conversion of the measurement data sets into a standardized measurement data format involved the following conversion and processing steps:

- General signal conversions and resampling of the measurement data to a uniform sampling rate of 500 Hz;
- General calculations (e.g. calculation of lateral jerk and yaw acceleration);
- Signal filtering (Butterworth low-pass filter, 10th order, 6.5 Hz cut-off frequency) for selected measured variables (e.g. lateral acceleration, yaw rate, steering rate);
- Trimming of the measurement files to the relevant time range (5 s before/after the failure activation);
- Creation of standardized measurement data plots of relevant measured variables;
- Storage of the measurement data records in a standardized data structure and a standardized data format.

After data preparation, all measurement data sets of the tests carried out were available in a uniform data format (Matlab data format) and can be used for maneuver-dependent data evaluation.

As part of the measurement data evaluation, relevant characteristic values were determined or calculated using the available vehicle measurement data to objectively assess the influences of the failure activation on the driver inputs and the resulting driving dynamics. In addition, individual characteristic values (e.g. the determined driving speed at the time of the failure activation) serve as a criterion for the definition of valid test data sets (see section 2.7).

The evaluation of the measurement data sets comprises the following evaluation steps:

- Determination of general characteristic values
  - Driving speed at the time of failure activation;
  - Steering wheel angle and rate at the time of failure activation.
- Calculation of test-dependent characteristic values (according to working group specifications)
  - Disturbance steering wheel angle;
  - max. steering rate after failure activation ("failure steering rate");
  - "Steering effort" before/after failure activation
    (only for maneuvers in circular drive and straight-ahead);
  - Disturbance lateral acceleration;
  - Lateral jerk;
  - Disturbance yaw rate;
  - Disturbance acceleration.
- Creation of standardized evaluation plot (calculation of characteristic values).

The calculation of the scenario-dependent characteristic values was specified by the working group. The following calculation rules were used to calculate the disturbance influence on the steering wheel angle, the disturbance lateral acceleration, the disturbance lateral jerk, and the disturbance yaw rate (see Illustration 6).

Illustration 6

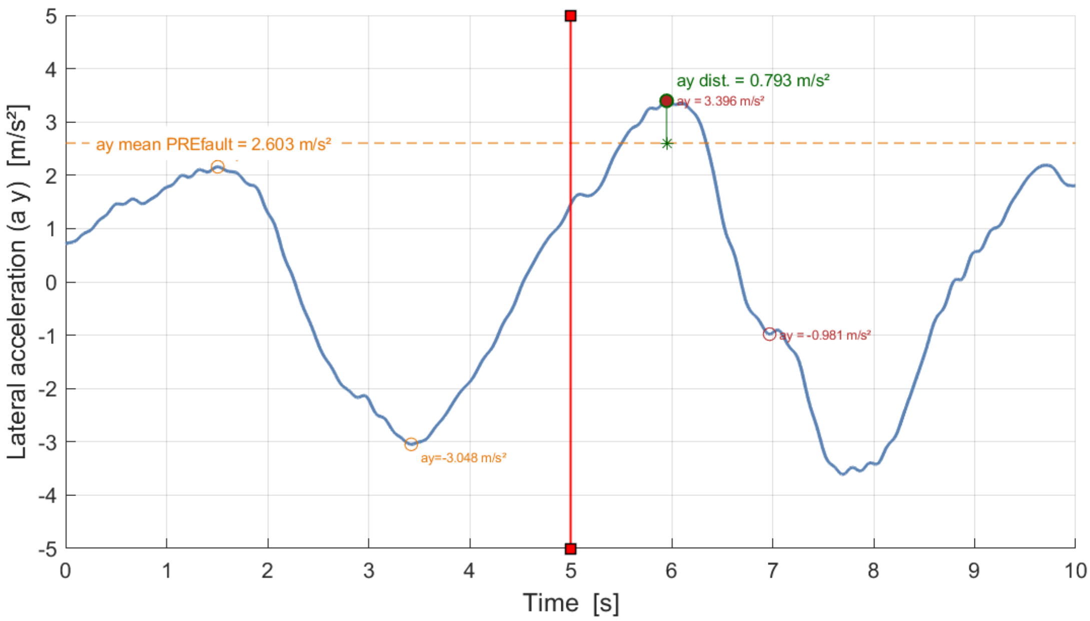
Calculation rules Objective data

<table>
<tr><td colspan="2">Evaluation variables:
▪ Disturbance steering angle: Steering angle
▪ Disturbance lateral acceleration: Lateral acceleration
Lateral jerk (time derivative of lateral acceleration)
▪ Disturbance lateral jerk: Yaw rate
▪ Disturbance yaw rate: Yaw acceleration
▪ Disturbance acceleration:</td></tr>
<tr><td>Failure types:
▪ Blocked FBA
▪ FBA Step
▪ Blocked RWA</td><td>Failure types:
▪ RWA square-wave
▪ Uncontrolled RWA
▪ FBA Selfsteer + Loss of feedback</td></tr>
<tr><td>Driving maneuvers:
▪ Slalom</td><td>Driving maneuvers:
▪ Cornering
▪ Straight-ahead</td></tr>
<tr><td><u>Evaluation before failure activation</u>
▪ Evaluation range 5 s before failure activation
▪ Formation of the mean value of the absolute values of the two distinct peaks of the evaluation variable
<u>Evaluation after failure activation</u>
▪ Evaluation range after failure activation
- Blocked FBA: 5 s
- FBA step: 2 s
- Blocked RWA: 5 s
▪ Identification of the two distinct peaks of the evaluation variable
<u>Calculation of the disturbance variable</u>
▪ Difference between the mean value before failure activation and the absolute value of the largest peak after failure deactivation corresponds to the disturbance variable</td><td><u>Evaluation before failure activation</u>
▪ Evaluation range 5 s before failure activation
▪ Formation of the mean value of the evaluation variable
<u>Evaluation after failure activation</u>
▪ Evaluation area after failure activation
▪ - RWA square-wave: 3 s
- Uncontrolled RWA: 2 s
- FBA Selfsteer + Loss of feedback: 2 s
▪ Identification of the largest peak of the evaluation variable in terms of amount
<u>Calculation of the disturbance variable</u>
▪ Difference between the mean value of the evaluation variable before the failure trigger point and the absolute value of the largest peak after the failure trigger point corresponds to the disturbance variable</td></tr>
</table>

Illustration 7 shows an example of the relevant measuring points for determining the disturbance lateral acceleration for a measurement data set (FBA step during a slalom). The failure is applied at time t = 5 s. The reference value *ay mean PREfailure = 2.603 m/s²* results from the averaging of the absolute values of the two identified peaks of lateral acceleration in the time range 5 s before failure activation. The magnitude of the largest peak in the time range 2 s after failure activation is *ay = 3.396 m/s²*. The lateral acceleration resulting from the difference between these two values is *ay dist = 0.793 m/s²*.

Illustration 7

Determination of the disturbance lateral acceleration for a FBA step situation (slalom)



The maximum steering rates after failure activation are determined for the time range from 0.1 s after failure activation. The "steering effort" is only calculated for the approximately stationary driving maneuvers straight-ahead and cornering. The calculation is performed by time integration of the steering angle signal over a fixed time interval of 5 s before and after the failure activation. The determined values of the steering effort before and after the failure activation are then considered in relation to each other in order to be able to evaluate the increase or decrease in steering effort as a result of the failure activation.

All characteristic values of the respective tests determined and calculated as part of the evaluation of the road test data were summarized and made available to the working group in a results file after the study was completed.

2.4.3 Subjective measured variables

Various subjective measures were used to record the subjective reaction to the failure activation. The instruments are described below. A complete description of the survey concept, including the formulation of the individual items used, can be found in the appendix 6.1 to 6.4.

Three items based on the NASA-TLX (Hart, 2006) were included in the survey concept in order to be able to assess how demanding the specified driving maneuvers were perceived to be by normal drivers, regardless of the failure activation. The test subjects were asked to indicate on a twenty-point scale how high the mental and physical demands were and how well they rated their own performance in each maneuver. In addition to the difficulty of the individual maneuvers, the test subjects were also asked to assess the difficulty of the course as a whole with the help of another item ("How challenging do you find the course overall?"). The extent of trust in the vehicle was also operationalized with a single item based on the Trust in Automation Questionnaire according to Körber (2018) ("How much do you agree with the following statement: I trust the vehicle?"). Both individual items were each answered on a seven-point scale (difficulty of the overall course: (1) "Not at all challenging" to (7) "Very challenging"; trust: (1) "Don't agree at all" to (7) "Fully agree").

The reaction of the test subjects to the failure activation was recorded both qualitatively and quantitatively. On the qualitative level, the experimenter observed the initial reaction of the drivers. In addition, the subjectively perceived reaction was measured by questioning the test subjects on an affective ("What did you feel during the event?"), cognitive ("What did you think during the event?") and behavioral level ("What did you do as a result of the event?"). With regard to the behavioral level, respondents were also asked whether and, if so, why they would act differently if such an event were to happen again. The Neukum scale (Neukum & Krüger, 2003) and subscales of the CTAM (Osswald et al., 2012) were used to assess the quantitative reaction to the failure activation. The Neukum scale for recording the perceived criticality of a driving situation consists of an eleven-point scale, whereby the scale points are assigned to increasing categories. While scale point 0 is assigned to the category "Nothing noticed", points 1 to 3 express "Noticeability" and points 4 to 7 "Disturbance of driving". Ratings 7 to 9 stand for "dangerousness" and the value 10 can be used to express that the "vehicle was no longer controllable". Based on the subscales "Perceived Safety" (5 items) and "Anxiety" (6 items) of the CTAM according to Osswald et al. (2012), the constructs perceived safety ("I believe that such an event would be dangerous.") and anxiety ("The occurrence of this event would be frightening for me.") were recorded in the failure situation. For this purpose, the items suitable for the application context were selected, translated into German and adapted in their wording so that they relate to the failure event. In terms of reliability, the adapted "Perceived Safety" subscale showed a satisfactory Crobach's alpha of $\alpha = .620$ to $\alpha = .949$ across all six individual studies. To improve reliability ($\alpha = .740$ to $\alpha = .926$), one item of the "Anxiety" subscale ("I would be afraid that I would not understand such an event") was excluded from the data analysis.

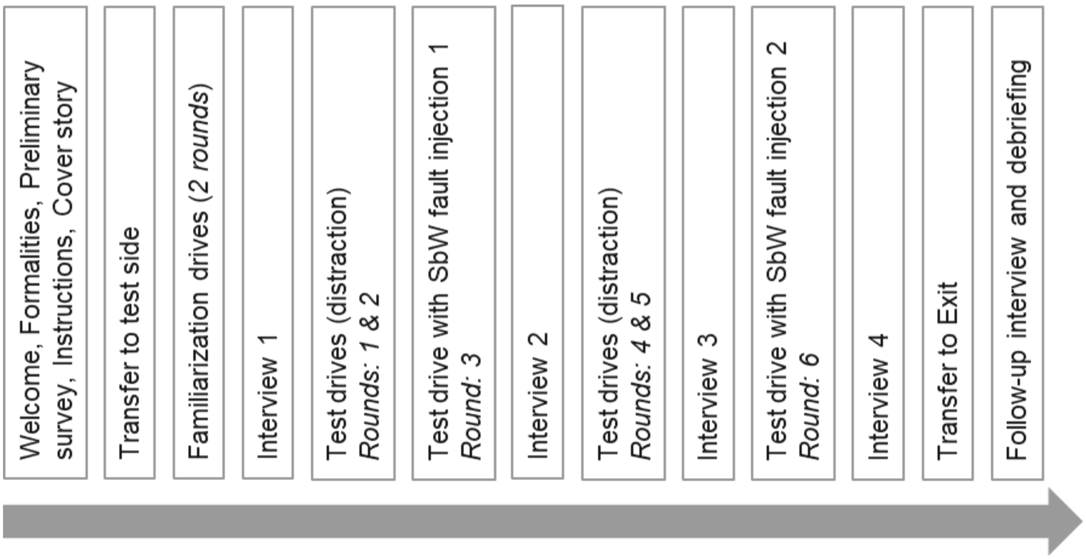
The follow-up survey was used to compare the FBA and RWA failures experienced in each case. Three items were collected for each failure, in which it was possible to indicate how likely it was that a vehicle with such a failure would be used or that a workshop would be visited after experiencing the failure in question. In addition, the failures were compared directly in terms of their criticality ("Which failure do you consider to be more critical?"). Finally, a qualitative, open-ended item was used to justify this comparison ("Please give reasons for your decision").

2.5 Study program

The study procedure is described in detail in the following section. Illustration 8 shows an overview of the procedure.

Illustration 8

Study procedure

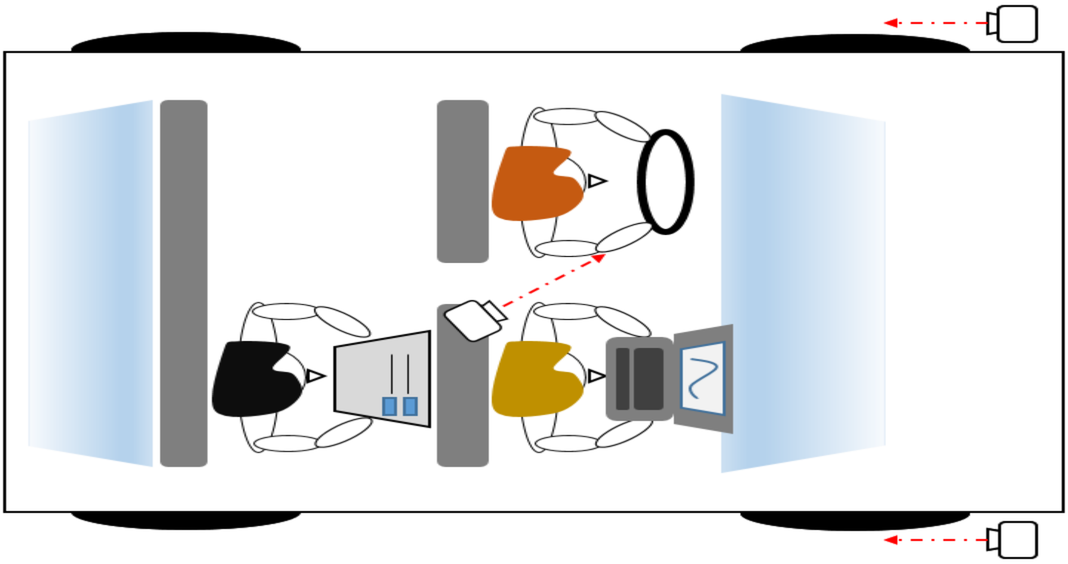


At the beginning, the test subjects were welcomed by the test management and led to an interview station outside the actual test track area. On site, the 2G+ proof was first checked in accordance with the hygiene concept (triple vaccination protection or double vaccination protection with an additional daily negative rapid test) and hands were disinfected. The test subjects then completed the formal documents and a short preliminary survey on socio-demographic data (age, gender), driving habits (driving license acquisition, annual mileage) and experience with driver assistance systems (see Appendix 6.1).

The test management then presented the test subjects with a cover story in order to initially disguise the true background of the study and thus divert the focus away from a possible "troubleshooting" by the test subjects. The test subjects were told that the study was about the further development of automated driving. This would require data from human drivers in order to adapt the automated driving function to human driving style and thus increase subsequent user acceptance. In addition to the cover story, the test subjects were given basic information about the test drive procedure. Following the instructions, the test management drove the test subjects to the driving dynamics area of the test site. On the driving dynamics area, they were transferred to the actual test vehicle. The trained vehicle supervisor responsible for triggering the failures was waiting in the passenger seat. In line with the cover story, this person was introduced to the test subjects as the person responsible for recording the driving data. Before getting into the vehicle, the driver's seat and all contact surfaces on the outside of the vehicle were disinfected by the test supervisor in line with the hygiene concept. The test subjects then got in on the driver's side, while the test supervisor sat in the rear right-hand seat diagonally to the test subjects (see Illustration 9).

Illustration 9

Test setup in the vehicle



In the vehicle, the test subjects first had the opportunity to adjust the seat and mirrors correctly. This was followed by a briefing on the vehicle's controls and further instructions from the test supervisor. The test subjects were shown the three maneuvers and their sequence on the course was explained. As described in section 2.3), the course started with driving straight-ahead, followed by a circular drive and the slalom. One lap through the course therefore corresponded to completing all three maneuvers in the specified order. The test subjects were instructed to maintain a constant speed of 80 km/h (straight-ahead), 50 km/h (circular drive) and 60 km/h (slalom) in the individual maneuvers, to drive through the lanes marked out with cones as centrally and precisely as possible and, if possible, not to drive over any of the cones. A standardized hand position on the steering wheel ("quarter to three position", in which both hands are placed on either side of the steering wheel at 9 and 3 o'clock) was specified by the instruction. After clarifying any questions of understanding, two familiarization drives through the course were completed. These were used exclusively for familiarization with the vehicle and learning the course and were not relevant for the subsequent data analysis. In the first familiarization drive, the focus was particularly on the correct completion of the maneuvers at a self-selected speed, while in the second drive the test subjects were already asked to adhere to the correct recommended speeds.

After the familiarization phase, the test subjects were interviewed for the first time by the test administrator. For better understanding, the test subjects were able to read along with the questionnaire, while their verbal responses were recorded directly by the experimenter. In the course of the first interview, the test subjects were asked about their subjective perception of the difficulty of the individual maneuvers as well as their perception of the overall course and their confidence in the vehicle (see Appendix 6.2). The test subjects then started the actual experimental drives. Along the cover story, the test subjects were informed that from this point onwards the measurement data recording of their human driving style would start and that they would have to drive a total of six laps of the course. They were also informed that the test supervisor would ask them to stop the vehicle at various times in order to conduct another short survey on their subjective perception of the course. The test subjects had the opportunity to inquire about the correct speed limits at any time and received support from the test supervisor in the event of orientation difficulties. The first two rounds were used for distraction, so that no failures were recorded in these rounds. The task of the test supervisor in these and all subsequent rounds was to record any pylons that had been driven over and to reposition them if necessary. In the third round, the first failure was recorded in the corresponding maneuver, following the respective randomization logic of the individual trials. The direct reaction of the test subjects was recorded by the experimenter. After driving through the maneuver in which the failure was activated, the test subjects were asked to stop and questioned for the second time (see Appendix 6.3). The scope of the second interview varied depending on the failure detection. The test subjects were first asked whether they had noticed anything special while driving through the maneuvers. If this was not the case, the difficulty of the last maneuver driven, the perceived difficulty of the overall course and the current level of confidence in the vehicle were recorded in the same way as at interview time 1. The criticality of the failure on the Neukum scale was also recorded by the test supervisor with a zero ("Nothing noticed"). If it was clear from the subjects' description that they had noticed the failure, the subjects were informed that this was an event initiated by the experimenter. This event was deliberately not explained in more detail in order to avoid influencing the subsequent subjective judgments with the negatively connoted word "failure". In addition to the items already mentioned, the test subjects were then asked about their affective, cognitive and behavioral reactions to this initiated event. They were also asked to assess criticality using the categories of the Neukum scale and to answer the items relating to the constructs "perceived safety" and "anxiety", imagining that such an event would occur in a real driving context.

After the second interview, the test subjects completed the interrupted third round and drove through the next two distractor rounds (rounds 4 and 5) without making any further mistakes. Round 5 was followed by survey time 3, which was similar to time 1 (subjective difficulty of the maneuvers, difficulty of the overall course, confidence in the vehicle; see Appendix 6.2). In order to maintain the cover story, the test subjects were informed that only one more round was needed to record sufficient data. In this last round, the second failure was recorded in the associated maneuver. The test subjects were asked about this failure at the fourth interview time in the same way as described for interview time 2 (see Appendix 6.3). After completing the survey, the test subjects were asked to end the round and stop at the vehicle for driving them back. The test manager and test subjects changed into the vehicle there and returned to the interview station outside the test track. Once there, the test subjects were asked to complete the follow-up questionnaire. If only one failure was noticed by the test subjects, only the items associated with this failure were processed. If no failure was noticed, the follow-up survey was not carried out. The test subjects were then informed about the true background of the study, incentivized and bid farewell.

2.6 Selection criteria for the sample

A professional service provider was used to recruit a total of N = 168 test subjects. In order to draw as broad a sample as possible for each individual study, various target variables were specified for recruitment. These included the following socio-demographic characteristics: In addition to a balanced gender distribution, attention was paid in the recruitment process to a distribution of age across the groups 20 to 35 years, 35 to 50 years and 50 to 65 years. The prerequisites for participation were, on the one hand, possession of a class B driving license and an annual mileage of at least 2500 kilometers. The distribution of the total sample among the individual studies results in a separate socio-demographic distribution in each case. For this reason, the individual sample descriptions for each individual study are presented in the presentation of results (section 3) are reported individually.

2.7 Definition of valid data records

In order to define which data records are considered valid and can therefore be included in the data evaluation, three different criteria were established. The criteria relate to a) the verifiability of standardized failure activation, b) compliance with the target speed and c) driving performance in the area of track compliance without failure activation. The three criteria are described in detail below.

The first criterion deals with the verifiability of the standardized connection of the failures in the test situation, as described in section 2.2. The aim of this criterion was to remove those data sets from the data where the failure was activated at the wrong time, the effect of the failure was not (correctly) recognizable in the objective driving parameters or the objective measurement data set was not available or only available with failures.

While the first criterion is aimed in particular at the technical accuracy of the data, criteria 2 and 3 relate to the driving performance of the test subjects. With the help of these criteria, test subjects should be excluded who, due to their driving ability, were not able to reliably complete the various maneuvers with the associated instructions correctly, regardless of a failure. In the course of this, criterion 2 refers to compliance with the correct, specified reference speed in the individual maneuvers in order to ensure that the failure patterns in the intended dynamics have a reliable effect on the test subjects and can be reproduced as accurately as possible throughout the study. Criterion 2 stipulated that test subjects should be excluded if they were outside a tolerance range of the reference speed at the time of the failure activation (see section 2.3) +/- 5 km/h at the time of the failure activation. This criterion, which was defined a priori, proved to be useful for the two less demanding, stationary maneuvers of driving straight-ahead and circular drive. For the slalom maneuver, the criterion was extended post-hoc, as maintaining the target speed over the entire length of the maneuver was a major challenge for many test subjects. On average, the test subjects reached a speed of 56-57 km/h at the time of the failure activation in the slalom. As this reflects the driving performance of the entire sample and not just individual outliers, it was decided to extend the tolerance range for the slalom downwards and only exclude test subjects who drove more than two standard deviations slower than the mean speed. In this way, it was still possible to exclude those data sets in which the speed was clearly too low in comparison and therefore the dynamic effects of the failure patterns were not comparable.

The third criterion deals with the driving performance of the test subjects in relation to lane keeping. It can be assumed that leaving the lane by driving over the pylons after the failure activation can be attributed both to the effects of the failure and to the driving skills of the test subject. In order to minimize the probability of accidental driving errors in the sense of investigating the effects of the failure pattern, the frequency with which the specified driving corridor was left over the entire test without failure activation was examined. In the course of this, test subjects were excluded if they left the lane more than two standard deviations more frequently than the mean value of the sample per individual trial in a single maneuver. For each of the six rounds, the number of times a specific maneuver was not performed correctly was evaluated. A maneuver was considered to have been performed incorrectly as soon as at least one pylon was touched or driven over during the maneuver. The resulting distribution was used to identify the outliers who did not reliably complete a certain maneuver without the influence of a failure and were therefore excluded from the evaluation.

Illustration 10

Overview of the exclusion of test subjects on the basis of criterion 3

Vehicle	**Maneuver**	**M** *Number of deficient maneuvers*	**SD**	**Cut-Off** *(M deficit maneuvers + 2 SD)*	**Quantity invalid data records**	**Deficient maneuvers per invalid data record**
SUV 1	Straight-ahead	0,04	0,19	0,41	1	1
	Circular drive	0,18	0,55	1,27	2	2
	Slalom	0,68	1,25	3,18	1	6
Compact Class 2	Straight-ahead	0,00	0,00	0,00	0	/
	Circular drive	0,04	0,19	0,41	1	1
	Slalom	0,04	0,19	0,41	1	1
Compact Class 1	Straight-ahead	0,00	0,00	0,00	0	/
	Circular drive	0,12	0,44	1,00	2	1 x 1 1 x 2
	Slalom	0,12	0,33	0,78	3	1
Compact Class 3	Straight-ahead	0,00	0,00	0,00	0	/
	Circular drive	0,18	0,47	1,11	1	2
	Slalom	0,04	0,19	0,41	1	1
Sedan	Straight-ahead	0,04	0,19	0,41	1	1
	Circular drive	0,11	0,42	0,94	2	1 x 1 1 x 2
	Slalom	0,11	0,31	0,74	3	1
SUV 2	Straight-ahead	0,07	0,26	0,59	2	1
	Circular drive	0,04	0,19	0,41	1	1
	Slalom	0,04	0,19	0,41	1	1

In Illustration 10 shows an overview of the exclusion of test subjects based on the third criterion. The respective data is broken down per vehicle and maneuver. The third column lists the mean value of the number of deficient maneuvers per individual sample (theoretical range:

0 - 6). The fourth column shows the standard deviations for each of these mean values, followed by the cut-off value, calculated from the mean value plus two standard deviations. The last two columns show the number of test subjects who were defined as invalid due to exceeding the respective cut-off with regard to the deficient maneuvers, as well as the descriptive number of maneuvers that were not correctly mastered by these subjects. A very homogeneous picture emerged here, with the exception of the combination SUV 1 and slalom. The excluded test subject touched at least one of the slalom pylons in each of the six runs and therefore represents an extreme outlier that distorts the cut-off value for this sample upwards. Nevertheless, all other people in the sample had a maximum of two failing maneuvers.

In summary, data sets were considered valid and included in the data analysis where

- the standardized failure isolation could be verified using the objective parameters (criterion 1),
- in which the test subjects complied with the defined speed tolerance at the time of the failure activation (criterion 2) and
- in which no conspicuously frequent lane departure in the specific maneuver was found during the entire course of the test compared to the total sample (criterion 3).

The data sets were excluded on a maneuver-specific basis. If a test subject did not meet the criteria for one of the two types of failure in their experiment, the data for the other failure could still be used if validation was successful. After excluding the invalid data sets, the following picture emerged with regard to the sample sizes of the individual studies (see Illustration 11).

Illustration 11
Exclusion of invalid data records

Failure pattern	Number of participating test persons	Criterion 1 Correct data recording	Complete data sets	Criterion 2 Velocity & Criterion 3 Track compliance	final sample size
Blocked FBA	*n* = 28	Exclusion: 7	*n* = 21	Exclusion: 1	*n* = 20
	n = 28	Exclusion: 0	*n* = 28	Exclusion: 2	*n* = 26
RWA square-wave oscillation	*n* = 28	Exclusion: 0	*n* = 28	Exclusion: 0	*n* = 28
	n = 28	Exclusion: 0	*n* = 28	Exclusion: 4	*n* = 24
FBA step	*n* = 28	Exclusion: 0	*n* = 28	Exclusion: 2	*n* = 26
	n = 27	Exclusion: 3	*n* = 24	Exclusion: 2	*n* = 22
Uncontrolled RWA	*n* = 28	Exclusion: 1	*n* = 27	Exclusion: 5	*n* = 22
	n = 27	Exclusion: 0	*n* = 27	Exclusion: 2	*n* = 25
FBA Selfsteer + Loss of feedback	*n* = 28	Exclusion: 0	*n* = 28	Exclusion: 4	*n* = 24
	n = 25	Exclusion: 1	*n* = 24	Exclusion: 2	*n* = 22
Blocked RWA	*n* = 28	Exclusion: 1	*n* = 27	Exclusion: 2	*n* = 25
	n = 25	Exclusion: 8	*n* = 17	Exclusion: 2	*n* = 15

The target sample size was achieved for almost all individual studies. Only in the case of the combination of *blocked RWA* in the compact class 1 vehicle was it not possible to achieve the sample size due to poor weather conditions during the test and technical problems with the failure activation.

3 Results

In the following, the results are reported separately for each sub-study in the order in which they were conducted. The analysis was carried out using the statistical software IBM SPSS Statistics (version 27) (IBM, 2020) and Matlab (version R2018b).

3.1 Study 1

The first part of the study was carried out together with the SUV 1 vehicle and took place between March 28 and 31, 2022. The SUV 1 vehicle was used and the selected failure patterns were the FBA failure *Selfsteer + Loss of Feedback*, which was activated in the maneuver Circular drive and the RWA failure *Blocked RWA* in the associated maneuver Slalom. The results of the study section are presented in detail below.

3.1.1 Sample

The sample size comprises N = 28 participants, 11 of whom are female. The mean age is M = 36.71 years (SD = 12.94 years), with the youngest subject being 20 years old and the oldest 63 years old. The average annual mileage is M = 21,428.57 km (SD = 14,317.64 km) with a range of 8,000 km to 65,000 km. None of the test subjects stated that they had an uncorrected visual or hearing impairment.

3.1.2 Results

The presentation of the results is divided into two sections: failure-independent and failure-specific results. Firstly, the results are reported which relate to the general difficulty of the selected maneuvers in the sense of testing the selected test design, independently of the failure setups. The failure-specific results are then presented. In this section, the results for testing the controllability of both failures (FBA and RWA) are discussed first. This is followed by the subjective test data regarding the experience of the failure activation and the objective test data regarding the driver and vehicle reaction as a result of the failure activation, first for the FBA failure and then for the RWA failure.

3.1.2.1 Failure-independent results

The generally perceived difficulty of the maneuvers was assessed based on the NASA-TLX instrument (Hart, 2006) using three items relating to the mental demands, the physical demands and the subjective assessment of one's own performance on a 20-point scale (1 = very low to 20 = very high). The questionnaire was used at the beginning of the test (after the familiarization ride) and before the end of the test (after ride 5). For inferential statistical analysis of the learning curve, the requirements at the beginning and end of the ride were compared using a paired t-test. All test subjects with valid data sets for both failure cases were included in the analysis.

At both survey times, the three maneuvers of driving straight-ahead, circular drive and slalom were classified in a low to medium range on the 20-point scale in terms of mental demands.

The slalom maneuver appears to be associated with the comparatively highest mental demands and driving straight-ahead with the lowest. For a presentation of the characteristic values, see Table 4. Compared to the beginning of the test, all three maneuvers were perceived as mentally significantly less demanding before the last ride (after lap 5) (straight driving: $t(21) = 5.1$, $p < .001$, $d = 1.09$; circular drive: $t(21) = 4.83$, $p < .001$, $d = 1.03$; slalom: $t(21) = 4.7$, $p < .001$, $d = 1.01$).

Table 4

Mental demands of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high demands) based on the NASA TLX instrument (Hart, 2006).

Mental requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	22	5,50	3,29	1	12
	before the last round	22	3,14	1,86	1	9
Circular drive	after familiarization	22	9,36	3,81	2	15
	before the last round	22	6,50	3,56	2	15
Slalom	after familiarization	22	11,18	3,94	2	16
	before the last round	22	8,14	4,32	2	19

In terms of physical demands, all three maneuvers were also perceived as low to medium demanding on the 20-point scale. Here too, the slalom maneuver is rated as the most demanding, followed by circular drive and driving straight-ahead. For a presentation of the characteristic values, see Table 5. Before the last drive, all three maneuvers were perceived as significantly less physically demanding than at the beginning of the test (straight driving: $t(21) = 2.1$, $p = .024$, $d = 0.45$, circular drive: $t(21) = 2.97$, $p = .007$, $d = 0.63$; slalom: $t(21) = 3.24$, $p = .004$, $d = 0.69$).

Table 5

Physical demands of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high demands) based on the NASA TLX instrument (Hart, 2006).

Physical requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	22	3,45	2,67	1	9
	before the last round	22	2,77	1,63	1	8
Circular drive	after familiarization	22	6,95	3,17	2	13
	before the last round	22	5,64	3,20	2	12
Slalom	after familiarization	22	10,00	4,41	3	17
	before the last round	22	8,00	4,13	2	17

The test subjects rated their own performance over the course of the test and for all maneuvers in a high range on the 20-point scale, with performance being rated highest for the straight-ahead maneuver and comparatively lowest for the slalom maneuver. For a presentation of the characteristic values, see Table 6. At the end of the test, performance was rated significantly better in each maneuver compared to the beginning of the test (straight driving: $t(21) = -4.37$, $p < .001$, $d = -0.93$; circular drive: $t(21) = -3.78$, $p = .001$, $d = -0.81$; slalom: $t(21) = -4.37$, $p < .001$, $d = -0.93$).

Table 6

Assessment of own performance in the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high performance) based on the NASA TLX instrument (Hart, 2006).

Assessment of own performance						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	22	17,14	2,70	10	20
	before the last round	22	18,50	1,82	13	20
Circular drive	after familiarization	22	14,95	3,18	8	20
	before the last round	22	17,32	2,15	13	20
Slalom	after familiarization	22	14,18	3,66	8	20
	before the last round	22	16,41	2,46	12	20

3.1.2.2 Failure-specific results

In this section, the results for testing the controllability of the two failure patterns *FBA Selfsteer + Loss of Feedback* and *Blocked RWA are* discussed first. This is followed by a separate presentation of the subjective results for both failures with regard to the experience of the failure activation and then the objective test data with regard to the driver and vehicle reaction as a result of the failure events.

3.1.2.2.1 Controllability - hypothesis testing

For the first sub-study, the following applies to both the FBA failure *Selfsteer + Loss of Feedback* activated in the maneuver Circular Drive and the RWA failure *Blocked RWA* activated in the slalom: 100% of the test subjects with valid data records did not leave the lane after the failure was activated. 100% of the data records comprise at least $N = 20$ data records. The hypothesis can therefore be maintained for the combination of vehicle and failures selected in this part of the study. Within the scope of the study, both failures were controllable for all test subjects at the C0 level defined by experts.

3.1.2.2.2 Failure type: FBA Selfsteer + Loss of Feedback

In relation to the failure pattern *FBA Selfsteer + Loss of Feedback,* the results are reported below with regard to the objective vehicle measurement data to describe the vehicle and driver reaction as a result of the failure and the subjective perception of the connection.

3.1.2.2.3 Results of objective vehicle measurement data

The statistical distributions of the characteristic objective values for the *FBA Selfsteer + Loss of Feedback* failure pattern in the SUV 1 determined on the basis of the recorded vehicle measurement variables are shown in Illustration 12 shown. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern investigated are also shown in Table 7 summarized. For a qualitative comparison of the objective parameters determined, the following are shown in Illustration 12 also shows the statistical evaluations for the second vehicle, which was examined with the same failure pattern as part of the overall study. However, due to vehicle and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.

In the present case, the SUV 1 compared to the compact class 1 tends to show slightly less Disturbance with regard to the *FBA Selfsteer + Loss of Feedback* failure pattern in terms of the driver's steering response and the resulting vehicle dynamics. However, it should be noted that at 10 ms, the activation time of the torque step in the SUV 1 is only half as long as the activation time of the torque step in the compact class 1.

Illustration 12

Statistical evaluation of the calculated objective characteristic values for Disturbance steering angle, Disturbance steering rate, Disturbance lateral acceleration and Disturbance yaw rate for the failure pattern FBA Selfsteer + Loss of Feedback
(gray: SUV 1; white: compact class 1)

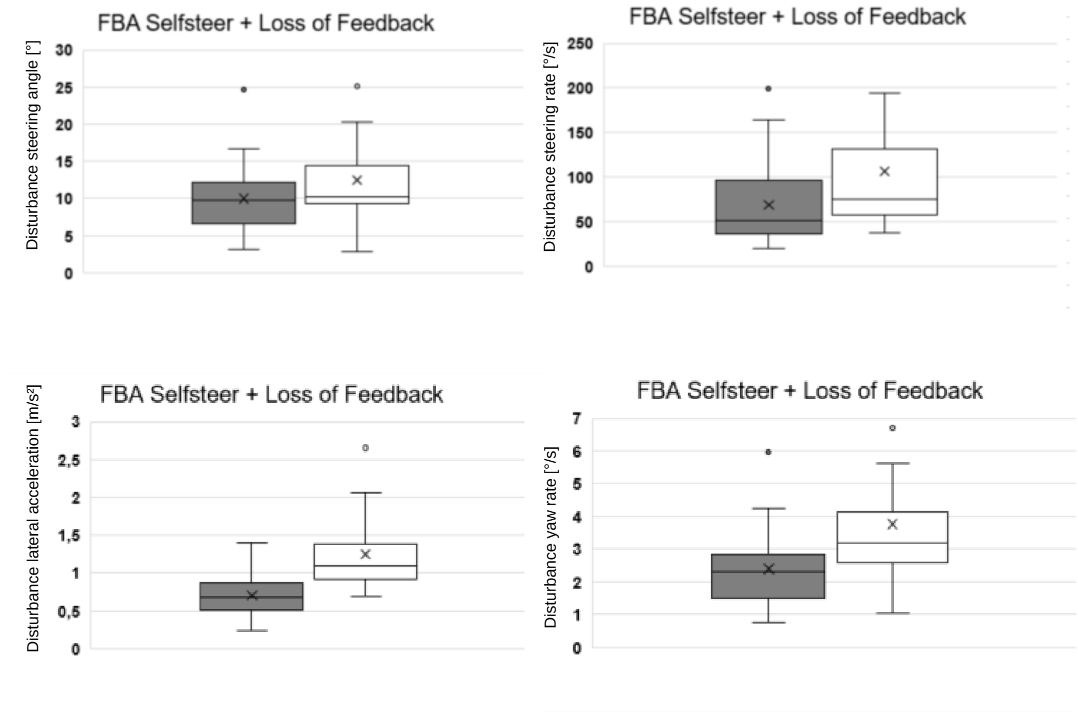


Table 7

Descriptive statistical characteristics FBA Selfsteer + Loss of Feedback (SUV 1)

Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
SUV 1	**Disturbance steering angle [°]**	24	10,01	4,70	9,77	3,13	24,64
	Disturbance steering rate[°/s]	24	68,47	46,76	50,63	19,44	198,88
	Dist. lateral acceleration [m/s²]	24	0,70	0,27	0,67	0,24	1,40
	Disturbance yaw rate [°/s]	24	2,41	1,16	2,32	0,78	5,98

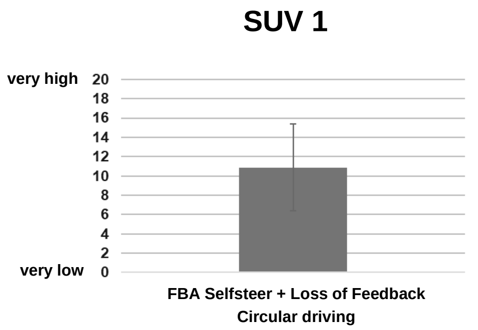
3.1.2.2.4 Results of subjective measures

In the following, the perceived difficulty of the maneuver of circular drive in the context of the activation of the failure is discussed first. The survey was conducted using the three items based on the NASA-TLX instrument (Hart, 2006) with regard to the mental and physical demands experienced as well as the subjective assessment of one's own performance using a 20-point response scale (1 = very low to 20 = very high). Due to the fact that the same type of failure was also made in sub-study 3 (for details see 4.3) with the compact class 1 vehicle using their test vehicle, the results of this sub-study are already presented here below with regard to the perceived difficulty of the maneuver in the event of a failure as a basis for comparison. However, it should be noted at this point that potential differences in the results of the two sub-studies can only be presented but not explained. An explanation is not possible due to the simultaneous manipulation of vehicle and failure parameterization. All valid data sets were included in the analysis.

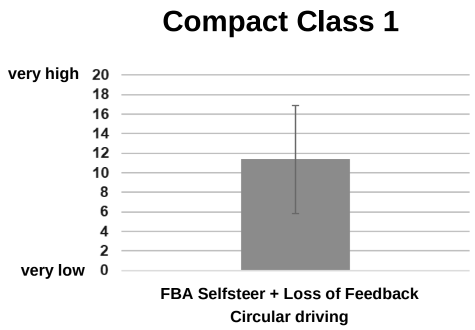
In the sub-study with the vehicle SUV 1 and the associated car used as the test vehicle, following the activation of the *FBA Selfsteer + Loss of Feedback* failure in the maneuver of circular drive, the mental requirements were classified on average in the middle range of the 20-point scale (*M* = 10.88, *SD* = 4.48). Similarly, the mental demands in the sub-study with the compact class 1 vehicle and the associated test vehicle were classified on average in the middle range of the response scale (*M* = 11.36, *SD* = 5.05). For a presentation of the results, see Illustration 13.

Illustration 13

Mental requirements of the maneuver of circular drive with activation of the FBA Selfsteer + Loss of Feedback failure were assessed on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
24	10,88	4,48	2	20



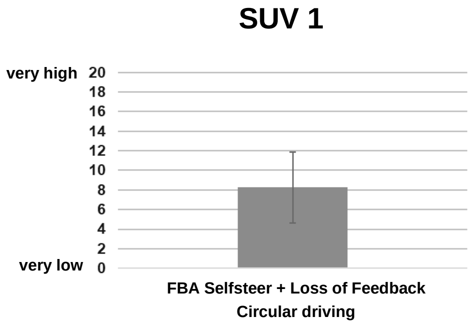
N	M	SD	Min	Max
22	11,36	5,05	4	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

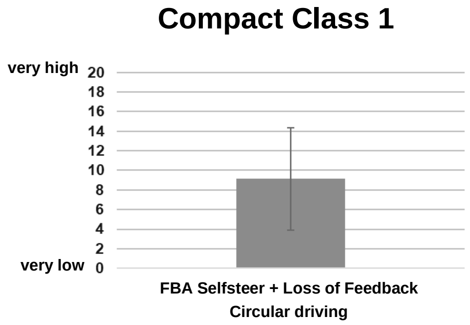
The physical requirements for completing the maneuver of circular drive were assessed in the middle of the 20-point response scale (*M* = 8.25, *SD* = 3.60) in the sub-study with SUV 1. This was also the case in the sub-study with the compact class 1 (*M* = 9.14, *SD* = 5.22). For a presentation of the results, see Illustration 14.

Illustration 14

Physical demands of maneuvering in a circular drive with activation of the FBA Selfsteer + Loss of Feedback failure were assessed on a 20-point response scale (1 = very low to 20 = very high demands) based on the NASA-TLX instrument (Hart, 2016).



N	M	SD	Min	Max
24	8,25	3,60	3	15



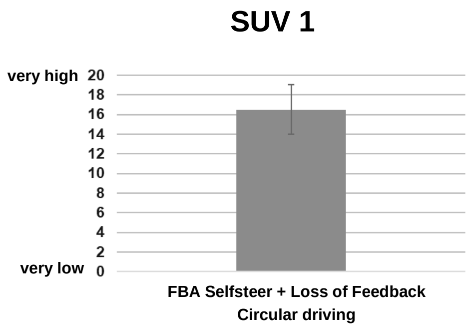
N	M	SD	Min	Max
22	9,14	5,22	1	18

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

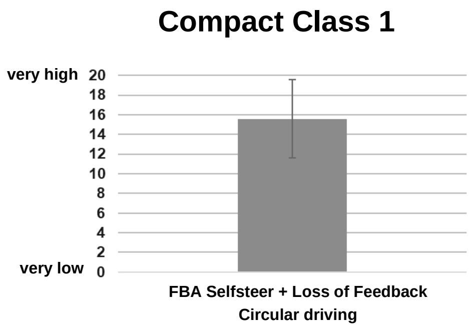
The subjective performance in completing the maneuver of circular drive with failure activation was rated in the medium to high range on the 20-point response scale in both sub-studies (SUV 1: M = 16.50, SD = 2.52; compact class 1: M = 15.59, SD = 4.00). For a presentation of the results, see Illustration 15.

Illustration 15

Subjective performance in completing the maneuver of circular drive with activation of the FBA Selfsteer + Loss of Feedback failure, measured on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
24	16,50	2,52	12	20



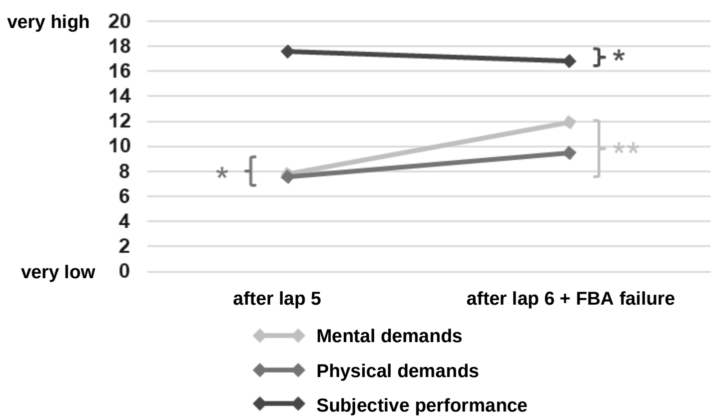
N	M	SD	Min	Max
22	15,59	4,00	4	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

With the aim of finding out the influence of the failure on the perceived difficulty of the maneuver, the perceived mental and physical demands as well as the subjective performance assessment after round 5 without failure activation were compared with those after failure activation in round 6. It was assumed that the learning curve of the test subjects with regard to completing the maneuvers had flattened out at this point in the test and that differences between the assessments were therefore attributable to the failure activation. It should be noted here that due to the randomized activation of the two failures (RWA and FBA failures) in rounds 3 and 6, only half of the valid data sets generated (in which the respective failure was activated in round 6) could be included in the analysis. The inferential statistical analysis of the comparison of the two survey times was carried out using a paired *t-test*. In sub-study 1, it became clear that both the mental and the physical demands in the last round with the activation of the FBA failure *Selfsteer + Loss of Feedback* in the associated manoeuver Circular drive increased significantly compared to the completion of the maneuver in the previous round 5, although they did not exceed the mean value range (mental demands: $t(10) = 3.56$, $p = .003$, $d = 1.08$; physical demands: $t(10) = 2.06$, $p = .034$, $d = 0.62$). In addition, the subjectively perceived performance decreased significantly, but was still perceived in a high range ($t(10) = -2.19$, $p = .027$, $d = -0.66$). For a graphical representation of the course and the associated parameters, see Illustration 16.

Illustration 16

*Subjective perceived difficulty assessed using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands and the subjectively perceived performance in the maneuver of circular drives in lap 5 without failure activation compared to lap 6 after activation of the FBA failure Selfsteer + Loss of Feedback. (Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	11	7,82	3,66	3	15
	after lap 6 + failure	11	11,91	4,87	2	20
Physical demands	after lap 5	11	7,55	2,98	3	12
	after lap 6 + failure	11	9,45	3,24	3	14
Subjective performance	After lap 5	11	17,55	1,92	15	20
	after lap 6 + failure	11	16,82	2,48	12	20

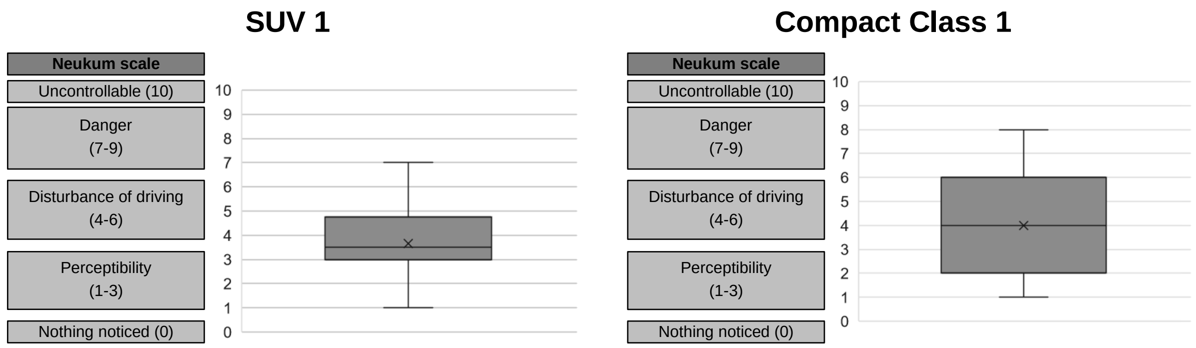
When the failure was activated, the test administrator observed the behavior of the test subjects with regard to an initial driver reaction to the event. Here, 75% of the test subjects showed a reaction noticeable to the test administrator in the immediate period of the FBA failure occurring, which in the vast majority of cases took the form of a verbal statement (n = 14) and less frequently a perceived physical reaction, such as "flinching" (n = 2). In order to find out whether the test subjects had noticed the failure activation, they were asked whether they had noticed anything in particular after completing the maneuver of circular drive in which the failure activation took place. This question was answered in the affirmative by 100% of the test subjects. When describing the event, an intervention or a bump on the steering was mentioned most frequently (n = 18). A cracking noise in the steering was also mentioned repeatedly (n = 7) and the assumption of having driven over a cone or something else was also expressed several times (n = 5).

The initial description of the experienced event was followed by an examination of the reaction to the failure activation. This was divided into the three aspects of affective, cognitive and behavioral reaction. To record the affective reaction, the test subjects were asked to describe which feelings were associated with the failure activation. n = 15 people mentioned a feeling of fright or surprise in this context. n = 7 people stated that they had not felt any fear, but had rather perceived the event with a neutral attitude. In each case, n = 3 people stated that they had either felt concern / fear or confusion about what had just happened. The cognitive reaction was recorded by asking people what they were thinking at the time the failure occurred. The most frequently mentioned answer was the assumption that they had knocked over a cone or something else (n = 11). n = 5 times the thought was expressed that something was different / not right. In each case, n = 3 times it was noted that the question "What was that?" or mental driving instructions (e.g. "Stay in control") had arisen when the incident was noticed. To record the behavioral reaction, the test subjects were asked to describe what they did as a result of the event. The most common response (n = 11) was to drive on normally / calmly. The second most common response (n = 8) was to reduce speed and the third most common response (n = 5) was to hold / control the steering wheel more firmly. If the event were to happen again, 66.67% (n = 16) of the test subjects would react exactly as described and 33.33% (n = 8) would react differently. Of the people who would react differently, the most common response (n = 4) was that they would be less frightened and react more calmly. It was mentioned n = 2 times each that either there would be no reduction in speed or there would be a check as to why the event had happened.

The Neukum scale (Neukum & Krüger, 2003) was used to assess the subjectively perceived criticality of the failure activation. Here, too, the results are compared with the results obtained for the same type of failure in sub-study 3 (for details see 3.3) with the test vehicle from the compact class 1 vehicle. Again, it must be taken into account that potential differences in the results of the two sub-studies can only be presented but cannot be explained due to the simultaneous manipulation of vehicle and failure parameterization. All valid data sets were taken into account for the analysis. The mean subjectively perceived criticality when experiencing the failure activation in the SUV 1 was classified on the 11-point scale (0 = nothing noticed - 10 = vehicle not controllable) in the range from a high level in the noticeability category to a low level in the disturbance of driving category (*M* = 3.67, *SD* = 1.27). In the sub-study carried out with the compact class 1 vehicle, the failure activation event was also classified on average at a low level in the driving disturbance category (*M* = 4, *SD* = 1.92). The dispersion of the values around the mean value appears somewhat wider here. For a representation of both distributions, see Illustration 17.

Illustration 17

Subjectively perceived criticality of the failure types FBA Selfsteer + Loss of Feedback in the two sub-studies with the SUV 1 and compact class 1 vehicles surveyed using the Neukum scale (Neukum & Krüger, 2003).



N	M	SD	Median	Percentile		
				25	50	75
24	3,67	1,27	3,50	3,00	3,50	4,75

N	M	SD	Median	Percentile		
				25	50	75
20	4	1,92	4,00	2,00	4,00	6,00

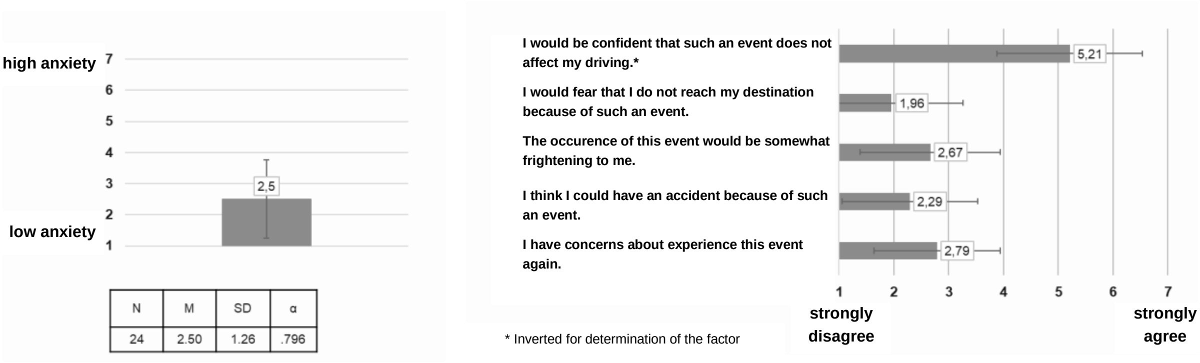
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The experience of the situation was assessed using adapted rating scales for the *Anxiety* and *Perceived Safety* factors of the *Car Technology Acceptance Model* (Osswald et al., 2012). This data was collected from all test subjects who stated that they had noticed the failure.

On average, the subjects classified their anxiety when the failure event occurred in a low range of the 7-point scale (1 = low anxiety - 7 = high anxiety) (*M* = 2.50; *SD* = 1.26). For a presentation of the results of the *anxiety* factor and the individual items used for the calculation, see Illustration 18.

Illustration 18

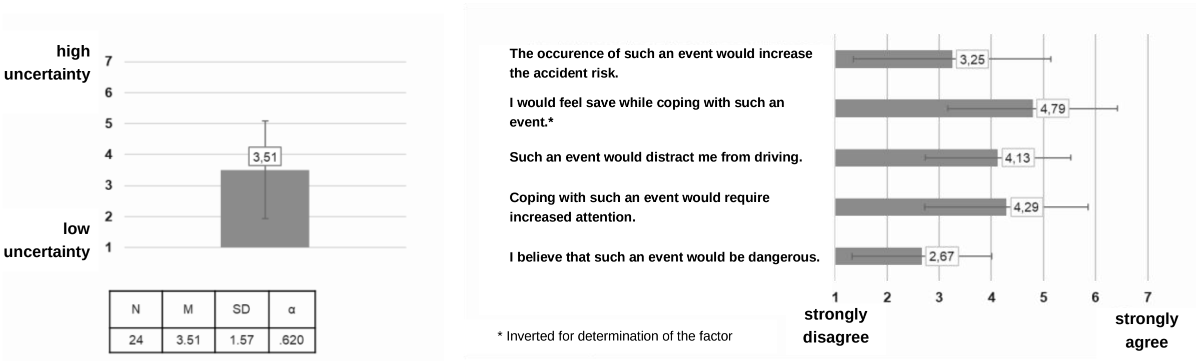
Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the failure event FBA Selfsteer + Loss of Feedback.



The perceived uncertainty with the occurrence of the failure was classified in a medium range of the 7-point scale (1 = low uncertainty - 7 = high uncertainty) (*M* = 3.51; *SD* = 1.57). For a presentation of the results of the *Perceived Safety* factor and the individual items used for the calculation, see Illustration 19.

Illustration 19

Results of the factor Perceived Safety (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the failure event FBA Selfsteer + Loss of Feedback.



3.1.2.2.5 Failure type: *Blocked* RWA

The results regarding the subjective perception of the failure activation and the objective results describing the vehicle and driver reaction as a result of the activation of the *blocked RWA* failure in the slalom maneuver are described below.

3.1.2.2.6 Results of objective vehicle measurement data

The statistical distributions of the objective characteristics determined on the basis of the recorded vehicle measured variables for the blocked RWA failure pattern in the SUV 1 are shown in Illustration 20. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern investigated are also shown in Table 8 summarized. For a qualitative comparison of the objective parameters determined, the following Illustration 20 also shows the statistical evaluations for the second vehicle, which was examined as part of the overall study with the same failure pattern. However, due to vehicle and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.

In the present case, the SUV 1 tends to show a slightly lower disturbance influence with regard to the steering angle compared to the compact class 1 with regard to the *blocked RWA* failure pattern. In contrast, the disturbance steering rates measured as a result of the failure activation are somewhat higher than in the compact class 1. With regard to the driving dynamics disturbance variables of lateral acceleration and yaw rate, there are no major differences between the two vehicles.

Illustration 20

Statistical evaluation of the calculated objective parameters of Disturbance steering angle, Disturbance steering rate, Disturbance lateral acceleration and Disturbance yaw rate for the blocked RWA failure pattern

(gray: SUV 1; white: compact class 1)

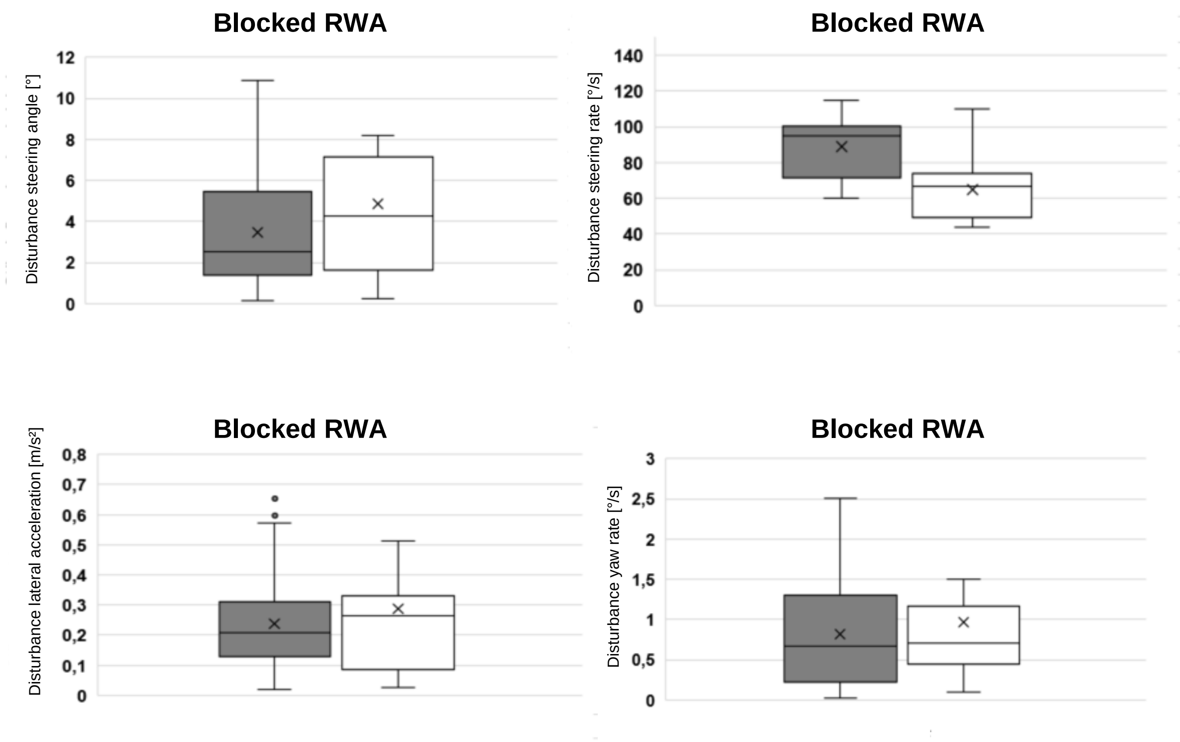


Table 8

Descriptive statistical values Failure pattern blocked RWA (SUV 1)

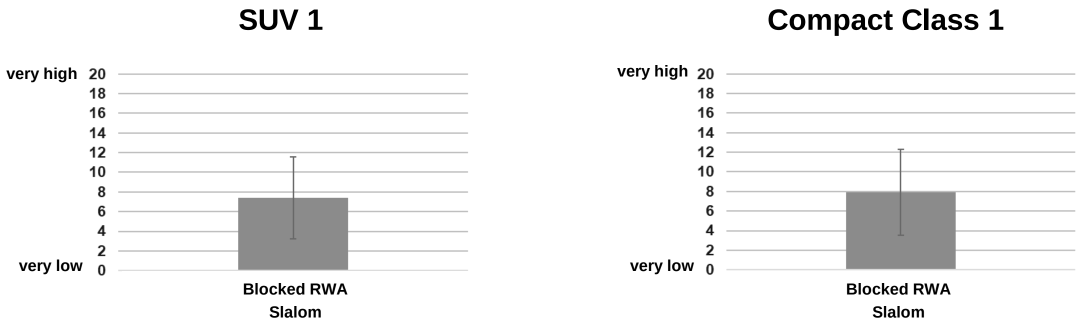
Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
SUV 1	Disturbance steering angle [°]	25	3,48	2,98	2,53	0,15	10,87
	Disturbance steering rate[°/s]	25	88,77	15,87	94,93	60,20	114,97
	Dist. lateral acceleration [m/s²]	25	0,24	0,16	0,21	0,02	0,65
	Disturbance yaw rate [°/s]	25	0,81	0,69	0,67	0,03	2,50

3.1.2.2.7 Results of subjective measures

First of all, the perceived difficulty of the slalom maneuver with failure activation is described, which was assessed using the three items based on the NASA-TLX instrument (Hart, 2006) with regard to the mental and physical demands experienced, as well as the subjective assessment of one's own performance on a 20-point response scale (1 = very low to 20 = very high). Again, in addition to the presentation of the results for the part of the study with the SUV 1 vehicle, the results of sub-study 3 (for details see 3.3) with the compact class 1 vehicle for the same type of failure. An explanation of potential differences in the results is not possible due to the simultaneous variation of failure parameterization and test vehicle. All valid data sets were included in the analysis. In both sub-studies, the mental requirements for completing the slalom maneuver were classified in the middle range of the 20-point scale when the *blocked RWA* failure was activated (SUV 1: M = 7.40, SD = 4.19; compact class 1: M = 7.93, SD = 4.38). For a presentation of the results, see Illustration 21.

Illustration 21

Mental requirements of the slalom maneuver with activation of the blocked RWA failure were assessed on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
25	7,40	4,19	1	15

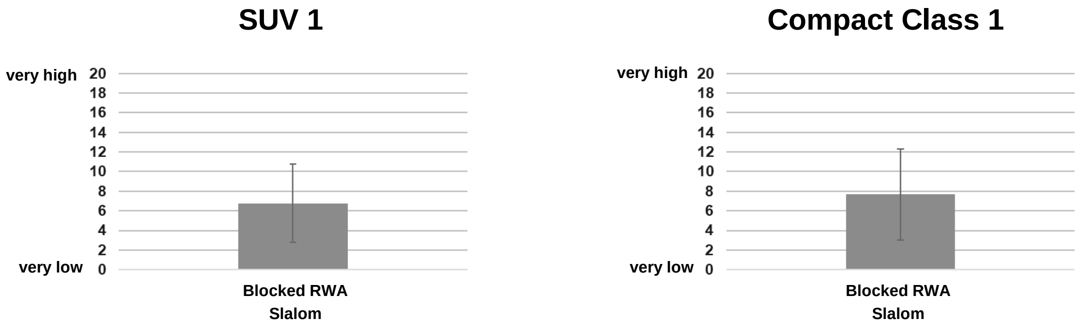
N	M	SD	Min	Max
15	7,93	4,38	3	15

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The physical requirements for completing the slalom maneuver with failure activation were rated in a low to medium range of the 20-point scale in both sub-studies (SUV 1: M = 6.76, SD = 4.00; compact class 1: M = 7.67, SD = 4.64). For a presentation of the results, see Illustration 22.

Illustration 22

Physical requirements of the slalom maneuver with activation of the blocked RWA failure were assessed on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
25	6,76	4,00	2	17

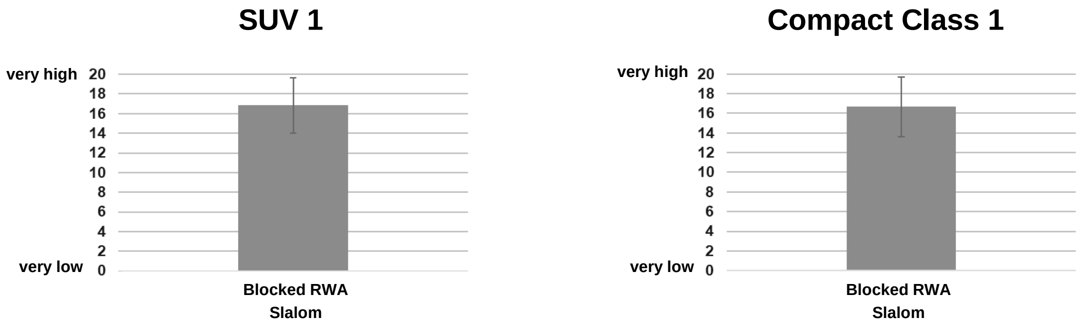
N	M	SD	Min	Max
15	7,67	4,64	2	18

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The subjective performance in completing the slalom maneuver with failure activation was rated in the medium to high range on the 20-point response scale in both sub-studies (SUV 1: *M* = 16.84, *SD* = 2.84; compact class 1: *M* = 16.67, *SD* = 3.02). For a presentation of the results, see Illustration 23.

Illustration 23

Subjective performance in completing the slalom maneuver with activation of the blocked RWA failure measured on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
25	16,84	2,84	8	20

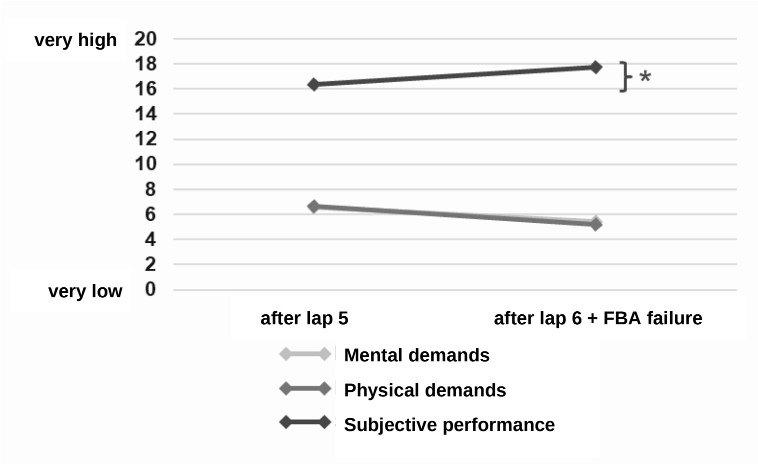
N	M	SD	Min	Max
15	16,67	3,02	10	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

In order to identify the influence of the failure activation on the perceived difficulty of the slalom maneuver, the perceived mental and physical demands as well as the subjective performance assessment after round 5 without failure activation were compared with those after failure activation in round 6. The inferential statistical analysis of the comparison of the two survey times was carried out using a paired *t-test*. It became clear that both the mental and the physical requirements remained constant in the comparison of the penultimate and the last round with the activation of the failure *"blocked RWA"* (mental requirements: $t(11) = 1.08$, $p = .15$; physical requirements: $t(11) = 1.53$, $p = .08$). At the same time, the subjectively perceived performance increases significantly (t(11) = = -2.24, $p = .024$, $d = -0.65$). For a graphical representation of the course and the associated characteristic values, see Illustration 24.

Illustration 24

Subjectively perceived difficulty assessed using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands as well as the subjectively perceived performance in the slalom maneuver in round 5 without failure activation compared to round 6 after activation of the blocked RWA failure.
*(Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	12	6,58	3,55	2	15
	after lap 6 + failure	12	5,42	3,68	1	13
Physical demands	after lap 5	12	6,67	3,68	2	15
	after lap 6 + failure	12	5,17	2,95	2	11
Subjective performance	After lap 5	12	16,33	2,46	13	20
	after lap 6 + failure	12	17,75	2,09	13	20

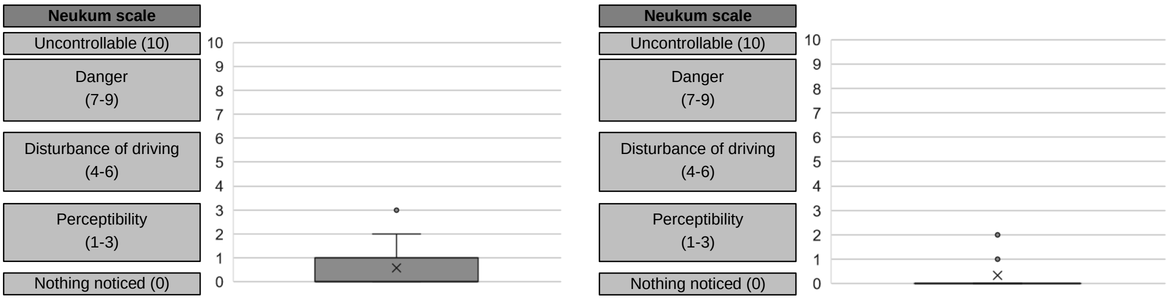
The observation of the behavior of the test subjects at the time of the failure activation by the test supervisor revealed that none of the test subjects showed a noticeable reaction to the activation of the *blocked RWA* failure. When asked whether the test subjects noticed anything special during the relevant slalom maneuver, 68% of the test subjects (n = 17) answered that they did not notice anything. 28% (n = 7) replied that they had noticed something and the most common associated description was an intervention in the steering / correction (n = 4), followed by the assumption that they had run over a cone or something else (n = 2). 4% of the test subjects (n = 1) were unsure whether they had noticed anything in particular.

The test subjects who stated that they had noticed something special were then asked about their affective, cognitive and behavioral reaction to the failure activation. The most frequently mentioned description of the affective reaction was neutral / nothing (n = 5), and two further individual opinions (*n* = 1 each) were no irritation and no unpleasant feeling, thus pointing in a similar direction. Other individual opinions (*n* = 1 each) were slight shock, astonishment and brief uncertainty. When asked about the cognitive reaction, the most frequently mentioned answer was that nothing special was thought of (*n* = 3). It was also mentioned several times (*n* = 2) that they thought they had driven over a cone or something else. Thoughts that were mentioned by individuals (*n* = 1) were "Not bad", "I left the lane", "What was that?" and "Something is different". When asked about the behavioral reaction, most of the test subjects who noticed the failure indicated that they had continued driving unchanged (*n* = 7). One person (*n* = 1) stated that they had held the steering wheel more tightly. When asked whether the test subjects would react differently if the event were to happen again, 100% of the test subjects (*n* = 8) answered no.

The evaluation of the subjectively perceived criticality of the event using the 11-point Neukum scale (0 = not noticed - 10 = vehicle not controllable) (Neukum & Krüger, 2003) revealed that the *blocked RWA* failure in the sub-study conducted with the SUV 1 vehicle and the vehicle was either not noticed or classified at a low level in the noticeability category (*M* = 0.60, *SD* = 0.91). All valid data sets were considered for the analysis. For comparison, the evaluation of subjective criticality when experiencing the same type of failure in sub-study 3 (for details see 3.3) with the compact class 1 vehicle using the associated test vehicle. As already noted above, it should also be mentioned at this point that potential differences in the results of the sub-studies can only be presented, but cannot be explained due to the simultaneous manipulation of vehicle and failure parameterization. In the sub-study of the compact class 1 vehicle, the *blocked RWA* failure was also either not noticed or classified at a low level in the noticeability category (*M* = 0.33, *SD* = 0.72). For a representation of both distributions, see Illustration 25.

Illustration 25

Subjectively perceived criticality of the blocked RWA failure type in the two sub-studies with the SUV 1 and compact class 1 vehicles using the Neukum scale (Neukum & Krüger, 2003).



SUV 1

N	M	SD	Median	Percentile		
				25	50	75
24	0,60	0,91	0,00	0,00	0,00	1,00

Compact Class 1

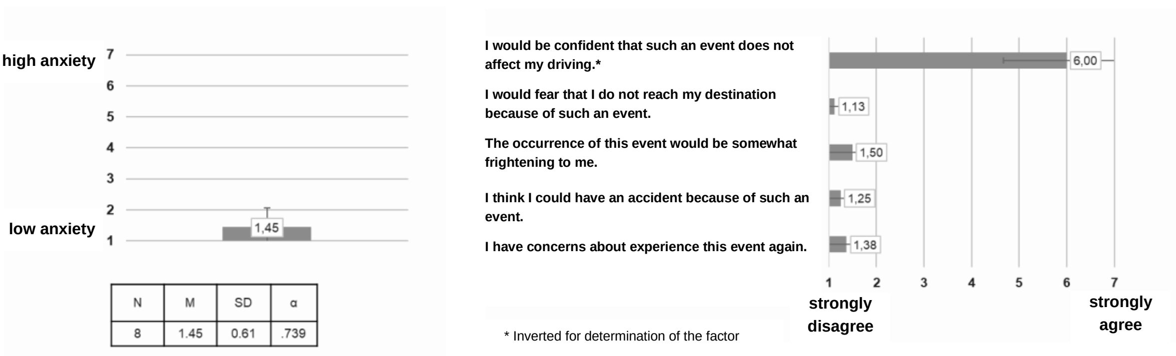
N	M	SD	Median	Percentile		
				25	50	75
15	0,33	0,72	0,00	0,00	0,00	0,50

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

Anxiety when noticing the failure event appears at a low level on the 7-point rating scale (1 = low anxiety - 7 = high anxiety) (M = 1.45, SD = 0.61). For a representation of the factor and the individual items used, see Illustration 26.

Illustration 26

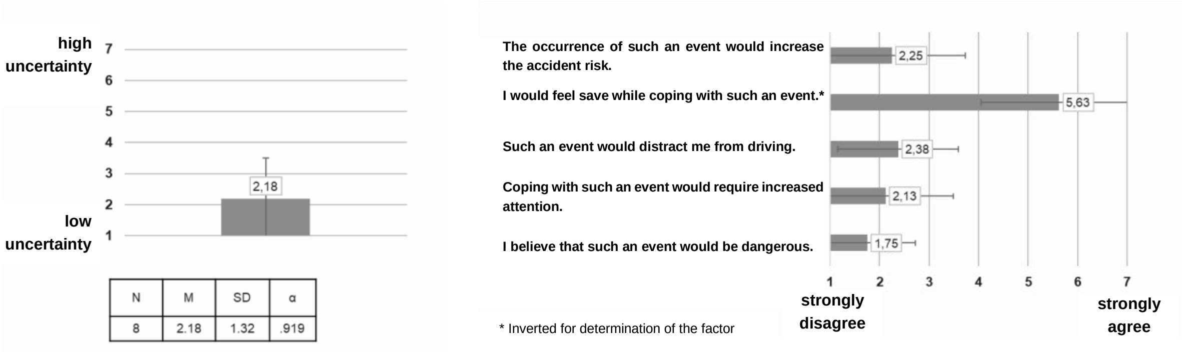
Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the blocked RWA failure event.



When the failure is noticed, the perceived uncertainty is classified in a low range of the 7-point rating scale (1 = low uncertainty - 7 = high uncertainty) (M = 2.18, SD = 1.32). For a presentation of the results of the factor and the underlying individual items, see Illustration 27.

Illustration 27

Results of the Perceived Safety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the blocked RWA failure event.



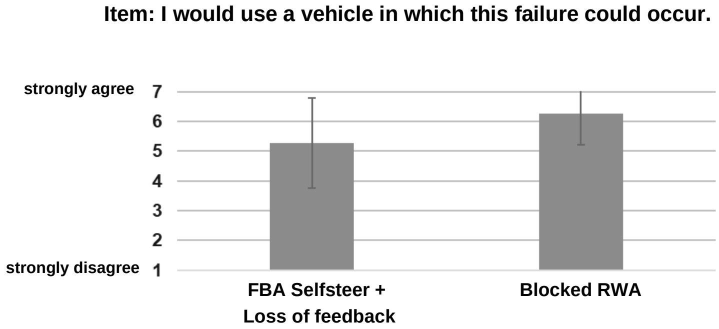
N	M	SD	α
8	2.18	1.32	.919

3.1.2.2.8 Follow-up survey

After the test subjects had been informed about the subject of the study following the test drives, a follow-up survey was conducted. The test subjects only answered the questions if they had noticed the respective failure during the test. First, the test subjects were asked to indicate the extent to which they agreed with the statement "I would use a vehicle in which this failure could occur". On a 7-point scale (1 = strongly disagree - 7 = strongly agree), the 36.36% of test subjects (n = 8) who noticed the *blocked RWA failure stated that* they would tend to use a vehicle in which such a failure could occur (M = 6.25, SD = 1.04). The FBA failure *Selfsteer + Loss of Feedback* was noticed by all test subjects (N = 22) and the intention to use appears to be descriptively lower here (M = 5.27, SD = 1.52). For a presentation of the results, see Illustration 28.

Illustration 28

Intention to use a vehicle that could exhibit the experienced failure FBA Selfsteer + Loss of Feedback or Blocked RWA, surveyed using a 7-point Likert scale.



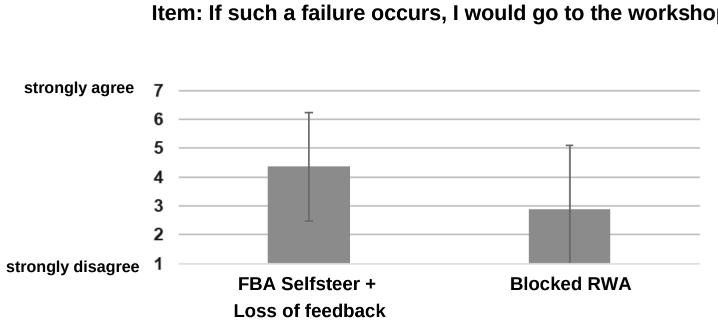
Failure	N	M	SD	Min	Max
FBA Selfsteer + Loss of feedback	22	5,27	1,52	2	7
Blocked **RWA**	8	6,25	1,04	5	7

Note. The item was only answered by test subjects who noticed the respective failure during the test run.

In addition, the test subjects were asked to indicate how much they agreed with the statement "If such a failure occurs, I would drive to the garage". The test subjects who noticed the RWA failure (*n* = 8) tended to disagree with this statement (*M* = 2.88, *SD* = 2.23). For the FBA failure, the subjects on average seemed to neither agree nor disagree with this statement (*M* = 4.36, *SD* = 1.87). For a presentation of the results, see Illustration 29.

Illustration 29

Intention to visit a workshop if the experienced failures FBA Selfsteer + Loss of Feedback or Blocked RWA occur in own vehicle, measured using a 7-point Likert scale.



Failure	N	M	SD	Min	Max
FBA Selfsteer & Loss of feedback	22	4,36	1,87	1	7
Blocked **RWA**	8	2,88	2,23	1	7

Note. The item was only answered by test subjects who noticed the respective failure during the test run.

3.1.3 Summary of Study 1

The following is an overview of the main results of sub-study 1 (vehicle: SUV 1; failure: *FBA Selfsteer + Loss of Feedback* / *Blocked RWA*).

- The design of the course appears appropriate, the test subjects are neither under- nor overchallenged.
- A learning curve of the test subjects over the course of the test becomes clear.
- The research hypothesis was retained for both failure patterns.
- The *FBA Selfsteer + Loss of Feedback* failure
  - was noticed by all test subjects;
  - is associated with a significant increase in the perceived mental and physical demands and a perceived deterioration in one's own performance in the driving maneuver in question;
  - is described by the test subjects as an intervention or impact on the steering;
  - is mostly associated with shock and surprise on an affective level;
  - is classified in terms of criticality at a high level of "noticeability" up to a slight level of the category "disturbance of driving".
- The failure *Blocked RWA*
  - was not noticed by the majority of test subjects (68%);
  - does not cause a significant increase in the perceived mental and physical demands or a significant deterioration in the subjectively perceived performance in the driving maneuver in question;
  - did not lead to a specific affective or behavioral reaction in the test subjects who detected the failure;
  - is classified at a low level of "noticeability" in terms of criticality if it was noticed by test subjects.

The content of the results of the sub-studies is discussed in Section 4.

3.2 Study 2

The second part of the study was carried out together with the compact class 2 vehicle in the period from April 1 to 6, 2022. The associated test vehicle was used and the selected failure patterns were the *FBA step,* which was activated in the slalom maneuver, and the *uncontrolled RWA* in the associated maneuver circular drive. The results of the study section are presented in detail below.

3.2.1 Study 2 sample

The sample size is *N* = 28, 18 of whom are male. The mean age is *M* = 42.71 years (*SD* = 14.20), with the youngest subject being 21 years old and the oldest 63 years old. The average annual mileage is *M* = 18,857.14 km (*SD* = 10,372.56 km) with a range of 3,000 km to 40,000 km. None of the test subjects stated that they had an uncorrected visual or hearing impairment at the time of the study.

3.2.2 Results of study 2

The presentation of the results is divided into two sections: failure-independent and failure-specific results. Firstly, the results are reported, which relate to the general difficulty of the selected maneuvers in terms of testing the selected test design, independently of the failure setups. The failure-specific results are then presented. In this section, the results for testing the controllability of both failures (FBA and RWA) are discussed first. This is followed by the subjective test data regarding the experience of the failure activation and the objective test data regarding the driver and vehicle reaction as a result of the failure activation, first for the FBA failure and then for the RWA failure.

3.2.2.1 Failure-independent results

The generally perceived difficulty of the maneuvers was assessed based on the NASA-TLX instrument (Hart, 2006) using three items relating to the mental demands, the physical demands and the subjective assessment of one's own performance on a 20-point scale (1 = very low to 20 = very high) at the beginning of the test (following the familiarization ride) and before the end of the test (following ride 5). For inferential statistical analysis of the learning curve, the requirements at the beginning and end of the ride were compared using a paired t-test. All test subjects with valid data sets for both failure cases were included in the analysis.

At both survey times, the three maneuvers of driving straight-ahead, circular drive and slalom were classified in a low to medium range on the 20-point scale in terms of mental demands. For a presentation of the results, see Table 9. Compared to the beginning of the test, all three maneuvers were perceived as mentally significantly less demanding before the last ride (after lap 5) (straight-ahead driving: $t(20) = 2.19$, $p = .02$, $d = 0.48$; circular drive: $t(20) = 1.97$, $p = .035$, $d = 0.43$; slalom: $t(20) = 4.33$, $p < .001$, $d = 0.94$). Table 9

Mental demands of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high demands) based on the NASA TLX instrument (Hart, 2006).

Mental requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	After familiarization	21	5,76	4,00	1	15
	Before the last round	21	4,52	3,25	1	11
Circular drive	After familiarization	21	9,19	4,78	1	17
	Before the last round	21	7,76	4,66	1	15
Slalom	After familiarization	21	10,52	4,79	1	19
	Before the last round	21	7,52	4,70	1	16

The three maneuvers are also classified in a low to medium range of the 20-point scale in terms of physical demands over the course of the test. For a presentation of the results, see Table 10. The physical demands of the maneuvers driving straight-ahead and circular drive remain constant over both survey times (driving straight-ahead: $t(20) = -0.86$, $p = .20$; circular drive: $t(20) = 1.32$, $p = .10$), in the slalom they decrease significantly over the course of the test ($t(20) = 4.3$, $p < .001$, $d = 0.94$).

Table 10

Physical demands of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high demands) based on the NASA TLX instrument (Hart, 2006).

Physical requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	21	3,52	2,52	1	9
	before the last round	21	3,81	2,52	1	9
Circular drive	after familiarization	21	7,52	4,34	1	15
	before the last round	21	6,95	4,66	1	18
Slalom	after familiarization	21	9,33	5,17	1	19
	before the last round	21	7,24	4,70	1	16

For all maneuvers over the course of the test, personal performance was assessed at a high level on the 20-point scale (see Table 11). At the end of the test, performance in the straight-ahead driving and slalom maneuvers was rated as significantly better compared to the beginning of the test (straight-ahead driving: $t(20) = -1.91$, $p = .036$, $d = -0.42$; slalom: $t(20) = -1.99$, $p = .03$, $d = -0.43$), whereas performance in the circular drive maneuver was rated as constant over the course of the test ($t(20) = -1.53$, $p = .07$).

Table 11

Assessment of own performance in the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high performance) based on the NASA TLX instrument (Hart, 2006).

Assessment of own performance						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	21	16,10	3,40	10	20
	before the last round	21	17,14	2,59	13	20
Circular drive	after familiarization	21	15,29	3,39	8	20
	before the last round	21	15,90	2,88	12	20
Slalom	after familiarization	21	14,71	4,31	6	20
	before the last round	21	16,00	2,85	10	20

3.2.2.2 Failure-specific results

In this section, the results for testing the controllability of the two failure patterns *FBA step* and *uncontrolled RWA are* discussed first. This is followed by a separate presentation of the subjective results for both failures with regard to the experience of the failure activation and then the objective test data with regard to the driver and vehicle reaction as a result of the failure events.

3.2.2.3 Controllability - hypothesis testing

For the second sub-study, 100% of the test subjects with valid data records did not leave the lane with failure activation, both for the *FBA step* failure activated in the slalom maneuver and for the *uncontrolled RWA* failure activated in the circling maneuver. 100 % of the data records comprise at least N = 20 data records. The hypothesis can therefore be maintained for the combination of vehicle and failures selected in this part of the study. Within the scope of the study, both failures were controllable for all test subjects at the C0 level defined by experts.

3.2.2.4 Failure type: *FBA step*

In relation to the *FBA step* failure pattern, the objective results describing the vehicle and driver reaction as a result of the failure are reported below, as well as the results regarding the subjective perception of the connection.

3.2.2.4.1 Results of objective vehicle measurement data

The statistical distributions of the objective characteristics determined on the basis of the recorded vehicle measurement variables for the FBA step failure pattern in the compact class 2 are shown in Illustration 30. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern investigated are also shown in Table 12. For a qualitative comparison of the objective parameters determined, the following Illustration 30 also shows the statistical evaluations for the second vehicle, which was examined as part of the overall study with the same failure pattern. However, due to vehicle and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.In the present case, the compact class 2 compared to the sedan tends to show slightly less Disturbance with regard to the *FBA step* failure pattern in terms of the driver's steering response and the resulting vehicle dynamics. However, the torque step of 1.5 Nm in the compact class 2 is also lower than in the sedan (2.2 Nm).

Illustration 30

Statistical evaluation of the calculated objective parameters of Disturbance steering angle, Disturbance steering rate, Disturbance lateral acceleration and Disturbance yaw rate for the FBA step failure pattern
(gray: compact class 2; white: sedan)

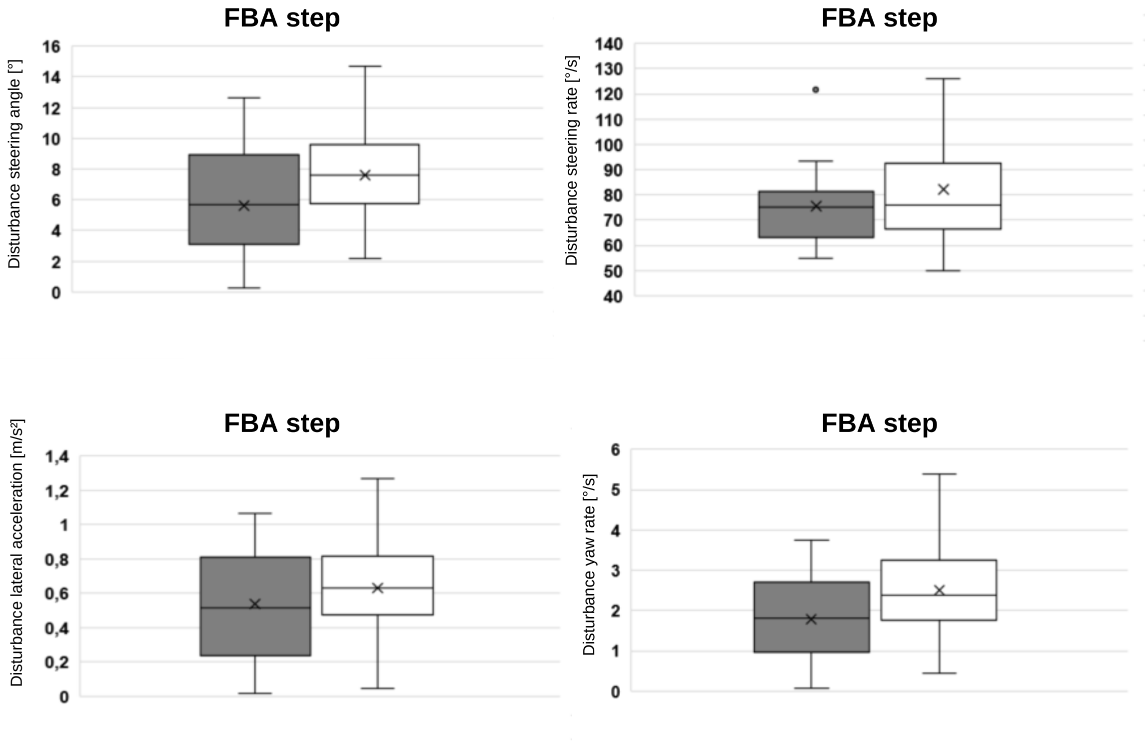


Table 12

Descriptive statistical characteristics Failure pattern FBA step *(compact class 2)*

Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
Compact Class 2	**Disturbance steering angle [°]**	26	5,63	3,44	5,67	0,25	12,63
	Disturbance steering rate[°/s]	26	75,42	14,47	74,97	55,00	121,52
	Dist. lateral acceleration [m/s²]	26	0,54	0,32	0,52	0,02	1,07
	Disturbance yaw rate [°/s]	26	1,78	1,08	1,81	0,07	3,74

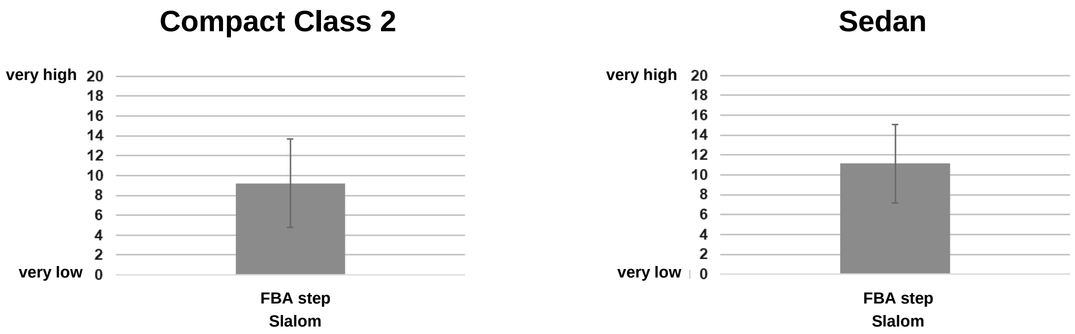
3.2.2.4.2 Results of subjective measures

The following describes the perceived difficulty of the slalom maneuver when the failure is activated. The survey was carried out using three questions based on the NASA-TLX instrument (Hart, 2006) regarding the subjective mental and physical demands as well as the assessment of one's own performance based on a 20-point response scale (1 = very low to 20 = very high). As the same type of failure was also used in sub-study 5 (for details see 3.5) with the sedan vehicle using their test vehicle, the results are already compared here. Differences in the results cannot be explained due to the simultaneous variation of vehicle and failure parameterization. All valid data sets were included in the analysis.

The mental requirements for completing the slalom maneuver were classified at a medium level on the 20-point scale in both sub-studies (compact class 2: *M* = 9.23, *SD* = 4.46; sedan: *M* = 11.14, *SD* = 3.96). For a presentation of the results, see Illustration 31.

Illustration 31

Mental requirements of the slalom maneuver with activation of the FBA step failure assessed on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
26	9,23	4,46	2	18

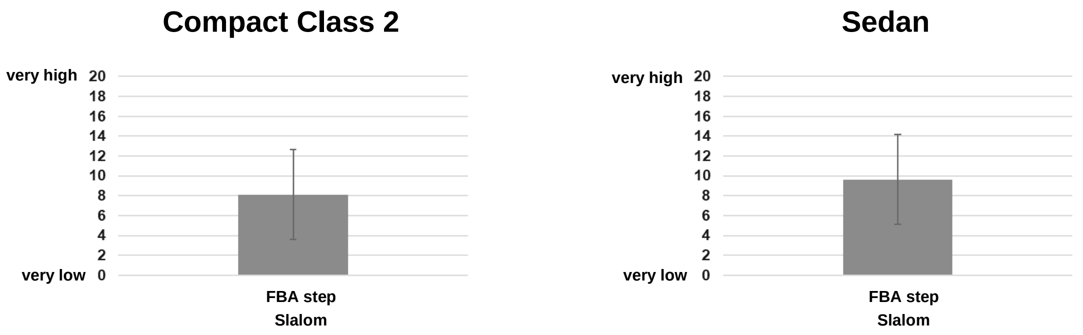
N	M	SD	Min	Max
22	11,14	3,96	4	17

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The physical requirements for completing the slalom maneuver with activation of the *FBA step* failure were also rated on average in the middle range of the 20-point scale in both sub-studies (compact class 2: *M* = 8.12, *SD* = 4.50; sedan: *M* = 9.64, *SD* = 4.50). For a presentation of the results, see Illustration 32.

Illustration 32

Physical demands of the slalom maneuver with activation of the FBA step failure assessed on a 20-point response scale (1 = very low to 20 = very high demands) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
26	8,12	4,50	1	17

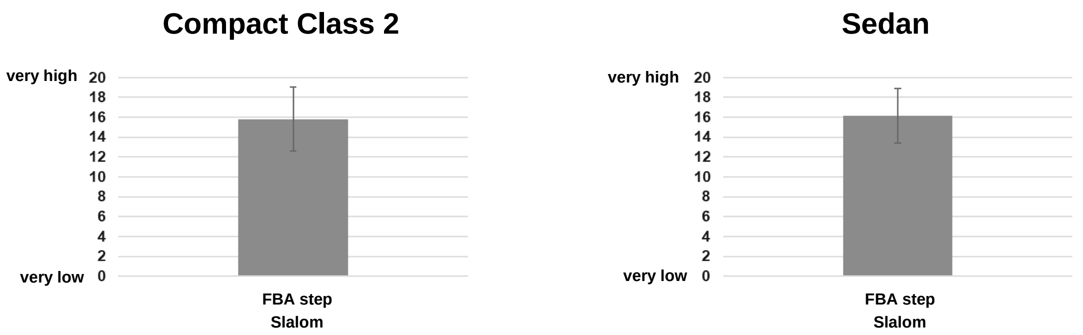
N	M	SD	Min	Max
22	9,64	4,50	1	17

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

In both sub-studies, the participants rated their own performance in completing the maneuver with the failure circuit at a medium to high level (compact class 2: M = 15.81, SD = 3.24; sedan: M = 16.14; SD = 2.73). For a presentation of the results, see Illustration 33.

Illustration 33

Subjective performance in completing the slalom maneuver with activation of the FBA step failure on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
26	15,81	3,24	10	20

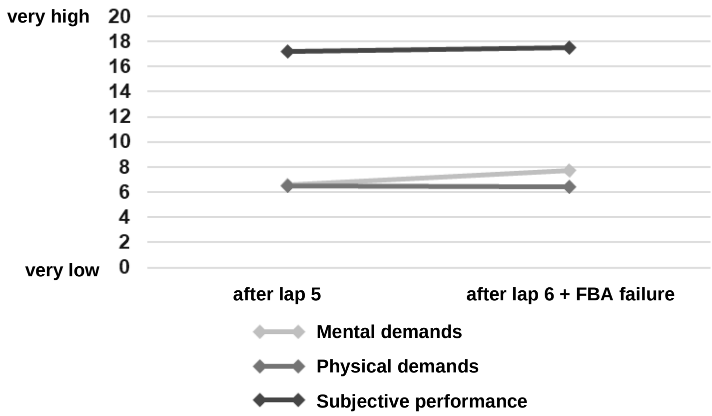
N	M	SD	Min	Max
22	16,14	2,73	9	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

To identify the influence of the failure activation on the perceived difficulty of the slalom maneuver, the perceived mental and physical demands as well as the subjective performance assessment after round 5 without failure activation were compared with those after failure activation in round 6. The basis of the comparison is the assumption that the learning curve of the test subjects when completing the maneuvers is flattened at this point in the test and that differences between the rounds can be attributed to the activation of the failure. It should be noted here that due to the randomized activation of RWA and FBA failures either in round 3 or in round 6, only half of the valid data sets generated (in which the respective failure was activated in round 6) could be included in the analysis. The inferential statistical analysis of the comparison of both survey times was carried out using a paired *t-test*. For the data sets in which the *FBA step* failure was activated in round 6, there were no statistically significant differences between rounds 5 and 6 in terms of mental and physical demands, or in the assessment of personal performance (mental demands: $t(11) = -1.65$, $p = .06$; physical demands: $t(11) = 0.19$, $p = .43$; subjective performance: $t(11) = -0.6$, $p = .28$). For a graphical representation of the course and the associated characteristic values, see Illustration 34.

Illustration 34

Subjectively perceived difficulty assessed using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands and the subjectively perceived performance in the slalom maneuver in round 5 without failure activation compared to round 6 after activation of the FBA step failure.
*(Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	12	6,58	3,48	2	12
	after lap 6 + failure	12	7,75	4,62	2	15
Physical demands	after lap 5	12	6,50	4,34	2	15
	after lap 6 + failure	12	6,42	3,73	2	12
Subjective performance	after lap 5	12	17,17	3,10	10	20
	after lap 6 + failure	12	17,50	2,54	12	20

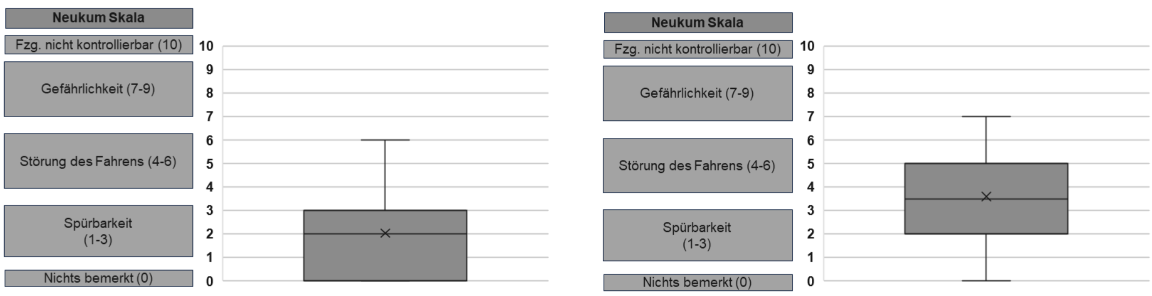
At the time of the failure activation, the behavior of the test subject was observed by the experimenter with regard to the initial reaction to the event. 19.23% of the test subjects ($n = 5$) showed a noticeable reaction, which was made clear in all cases by a verbal statement. In contrast, the vast majority of 80.77% of the test subjects ($n = 21$) did not show any noticeable reaction for the experimenter. After completing the slalom maneuver, the test subjects were asked whether they had noticed anything special during the maneuver. 73.08% of the test subjects (*n* = 19) answered in the affirmative and mentioned an intervention in the steering. 26.92% of the test subjects (*n* = 7) did not notice anything special. The test subjects who had noticed something special were subsequently asked to describe their reaction on an affective, cognitive and behavioral level. To record the affective reaction, the subjects were asked to describe their feelings at the time of the event. The most frequently mentioned statement (*n* = 6) described not having felt fear, but having been neutral towards the event. The second most common response (*n* = 4) was surprise/uncertainty. Fright/surprise, increased attention/tension, an unpleasant feeling and a positive/secure feeling were mentioned *n* = 3 times each. In order to describe the cognitive reaction, the test subjects were asked to report their thoughts at the time of the event. The most frequently mentioned thought (*n* = 11) was "There was something on the steering wheel / vehicle". The second most frequently reported thought was "something is different / wrong". To record the behavioral reaction, the test subjects were asked to describe what they did during the event. Equally frequently (*n* = 8), the test subjects reported either having counter-steered / corrected, or having continued driving normally / not having shown any particular reaction. If the event were to happen again, the majority of the test subjects (89.47% (*n* = 17)) would react in exactly the same way as the first time they experienced it.

The subjectively perceived criticality of the failure was recorded using the Neukum scale (Neukum & Krüger, 2003). The results for the *FBA step* failure type for the sub-study with the compact class 2 are presented below and also compared with the results of sub-study 5 (for details see 3.5) with the sedan, which had the same failure type using their vehicle as the test object. Potential differences between the results can only be reported, but not explained, as the vehicle and failure were manipulated simultaneously. All valid data sets were considered for the analysis. The average subjectively perceived criticality when experiencing the failure activation in the compact class 2 is at a medium level in the noticeability category (*M* = 2.04, *SD* = 1.71) on the 11-point scale (0 = nothing noticed - 10 = vehicle not controllable). In the sub-study with the sedan, the criticality of the failure was classified slightly higher between a high level of the category noticeability and a low level of the category disturbance of driving (*M* = 3.59, *SD* = 1.65). For a representation of both distributions see Illustration 35.

Illustration 35

Subjectively perceived criticality of the failure type FBA step in the two sub-studies with the compact class 2 and sedan vehicles using the Neukum scale (Neukum & Krüger, 2003).

Compact Class 2 **Sedan**



N	M	SD	Median	Percentile		
				25	50	75
26	2,04	1,71	2,00	0,00	2,00	3,00

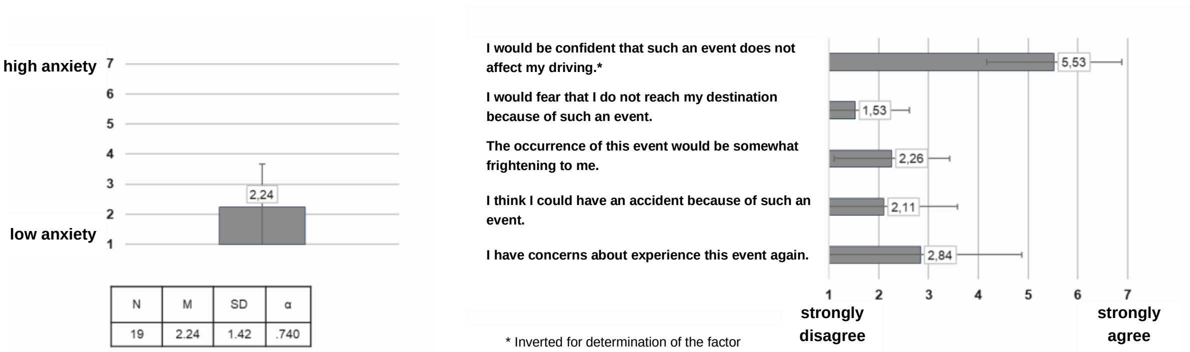
N	M	SD	Median	Percentile		
				25	50	75
22	3,59	1,65	3,50	2,00	3,50	5,00

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

Modified rating scales of the *Anxiety* and *Perceived Safety* factors of the Car Technology Acceptance Model (Osswald et al., 2012) were selected to assess the experience of the situation. This data was collected from all test subjects who stated that they had noticed the failure. The anxiety with experiencing the failure was classified on average at a low level of the 7-point scale (1 = low anxiety - 7 = high anxiety) (*M* = 2.24, *SD* = 1.42). For a presentation of the results of the factor and the individual items used for the calculation, see Illustration 36.

Illustration 36

Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the FBA step failure event.



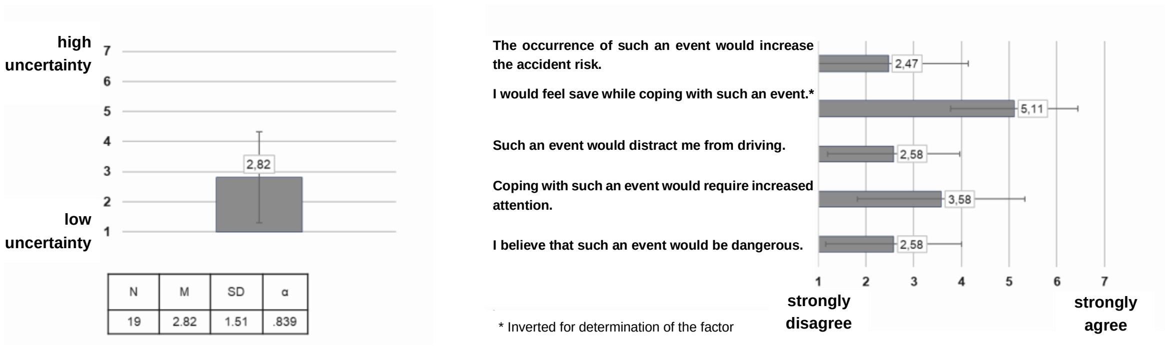
N	M	SD	α
19	2.24	1.42	.740

The perceived uncertainty with failure activation was also classified on average in a lower range of the 7-level scale (1 = low uncertainty - 7 = high uncertainty) (*M* = 2.82, *SD* = 1.51).

For a presentation of the results of the *Perceived Safety* factor and the individual items used for the calculation, see Illustration 37.

Illustration 37

Results of the Perceived Safety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the FBA step failure event.



3.2.2.5 Failure pattern: *Uncontrolled* RWA

In the following, the objective results describing the vehicle and driver reaction as a result of the activation of the failure pattern in the maneuver circular drive are described, as well as the subjective results regarding the perception of the failure activation.

3.2.2.5.1 Results of objective vehicle measurement data

The statistical distributions of the objective characteristics determined on the basis of the recorded vehicle measured variables for the failure pattern of *uncontrolled RWA* in the compact class 2 are shown in Illustration 38. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern investigated are also shown in Table 13 summarized. For a qualitative comparison of the objective parameters determined, the following Illustration 38 also shows the statistical evaluations for the second vehicle, which was examined with the same failure pattern as part of the overall study. However, due to vehicle and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance. In the present case, the compact class 2 compared to the sedan shows significantly lower disturbance influences in relation to the steering reaction of the driver and the resulting vehicle dynamics with regard to the failure pattern of *uncontrolled RWA.*

Illustration 38

Statistical evaluation of the calculated objective parameters of Disturbance steering angle, Disturbance steering rate, Disturbance lateral acceleration and Disturbance yaw rate for the uncontrolled RWA failure pattern
(gray: compact class 2; white: sedan)

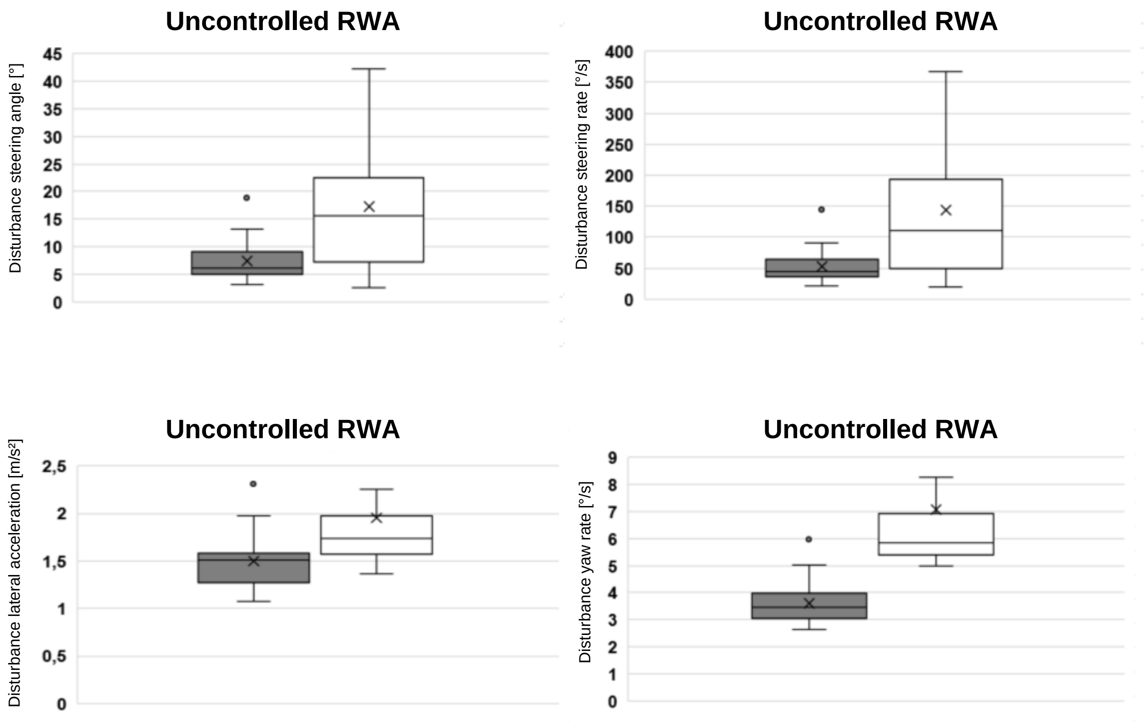


Table 13

Descriptive statistical parameters Failure pattern of uncontrolled RWA (compact class 2)

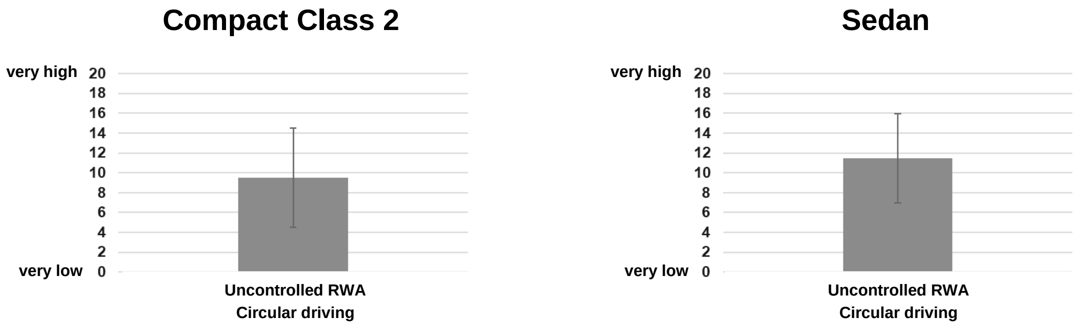
Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
Compact Class 2	**Disturbance steering angle [°]**	22	7,35	3,82	6,20	3,07	18,85
	Disturbance steering rate[°/s]	22	52,56	26,57	43,59	21,54	144,32
	Dist. lateral acceleration [m/s²]	22	1,50	0,27	1,51	1,07	2,31
	Disturbance yaw rate [°/s]	22	3,61	0,78	3,47	2,63	5,96

3.2.2.5.2 Results of subjective measures

The following section first presents the results on the perceived difficulty of the maneuver circular drive with failure activation. The survey was conducted using three items based on the NASA-TLX instrument (Hart, 2006) with regard to the mental and physical demands experienced and the subjective assessment of one's own performance when completing the circular drive maneuver with failure activation on a 20-point response scale (1 = very low to 20 = very high). In addition to the presentation of the results for the sub-study with the compact class 2 vehicle, the results obtained in sub-study 5 (for details see 3.5) with the sedan vehicle for the same failure type with their test vehicle. An explanation of potential differences between the sub-studies beyond the presentation of the results cannot be provided due to the simultaneous manipulation of failure parameterization and test vehicle. All valid data sets were included in the analysis. In both sub-studies, the mental requirements for completing the maneuver, including the failure, were classified in a medium range of the 20-point scale (compact class 2: *M* = 9.50, *SD* = 4.97; sedan: *M* = 11.48, *SD* = 4.47). For a presentation of the results, see Illustration 39.

Illustration 39

Mental requirements of maneuvering in a circular drive with activation of the failure Uncontrolled RWA collected on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
22	9,50	4,97	3	19

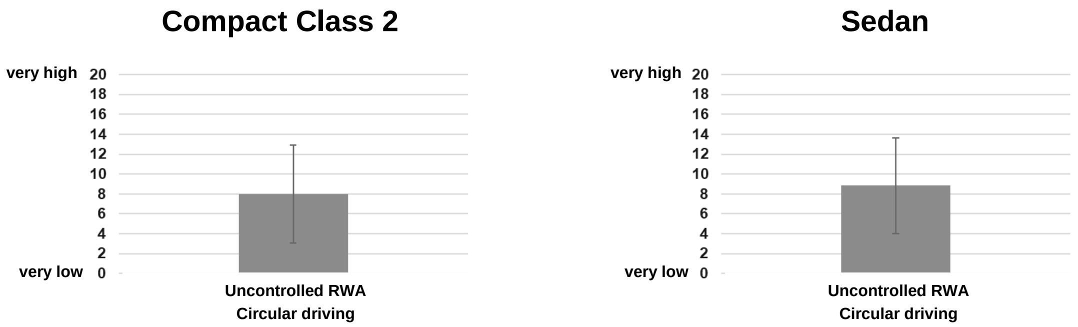
N	M	SD	Min	Max
25	11,48	4,47	3	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

In both sub-studies, the physical requirements for completing the maneuver circular drive with failure activation were assessed in a low to medium range of the scale (compact class 2: *M* = 8.00, *SD* = 4.91; sedan: *M* = 8.84, *SD* = 4.82) (see Illustration 40).

Illustration 40

Physical requirements of the maneuver Circular motion with activation of the failure Uncontrolled RWA collected on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
22	8,00	4,91	1	17

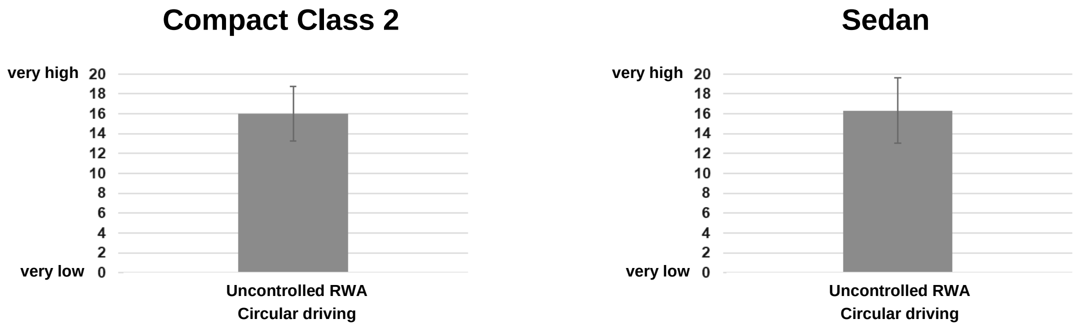
N	M	SD	Min	Max
25	8,84	4,82	1	19

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

In both sub-studies, the assessment of one's own performance when completing the maneuver circular drive with failure activation was at a medium to high level on the 20-point scale (compact class 2: M = 16.00, SD = 2.76; sedan: M = 16.32, SD = 3.28). For a presentation of the results, see Illustration 41.

Illustration 41

Subjective performance in completing the maneuver of circular drive with activation of the failure Uncontrolled RWA, measured on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
22	16,00	2,76	11	20

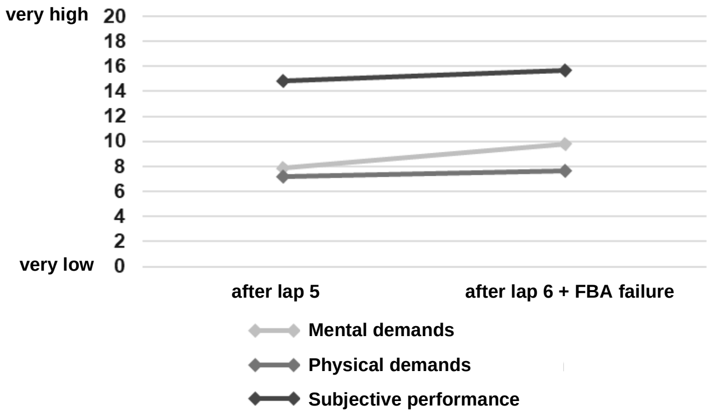
N	M	SD	Min	Max
25	16,32	3,28	8	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

In order to find out the influence of the failure on the perceived difficulty of the slalom maneuver, the perceived mental and physical demands as well as the subjective performance assessment after round 5 without failure activation were compared with those after failure activation in round 6. The comparison is based on the assumption that the learning curve of the test subjects with regard to completing the maneuver is flattened at this point and that differences between rounds 5 and 6 can be attributed to the activation of the failure. Due to the randomization of the activation of both failures to rounds 3 and 6, only half of the generated valid data sets (in which the respective failure was activated in round 6) can be included in the analysis. The inferential statistical analysis of the comparison of both survey times was carried out using a paired *t-test*. It is clear that there is no statistically significant difference between the sixth round with failure activation and the previous round 5 without failure activation, both in terms of mental and physical demands (mental demands: $t(10) = -1.66$, $p = .06$; physical demands: $t(10) = -0.79$, $p = .23$) The subjectively perceived performance also remains constant across both rounds ($t(10) = -1.48$, $p = .085$). For a graphical representation of the course and the associated characteristic values, see Illustration 42.

Illustration 42

Subjectively perceived difficulty measured using three items based on the NASA-TLX instrument (Hart, 2016) in relation to the mental and physical demands as well as the subjectively perceived performance in the maneuver circular drive in round 5 without failure activation compared to round 6 after activation of the failure Uncontrolled RWA.
*(Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



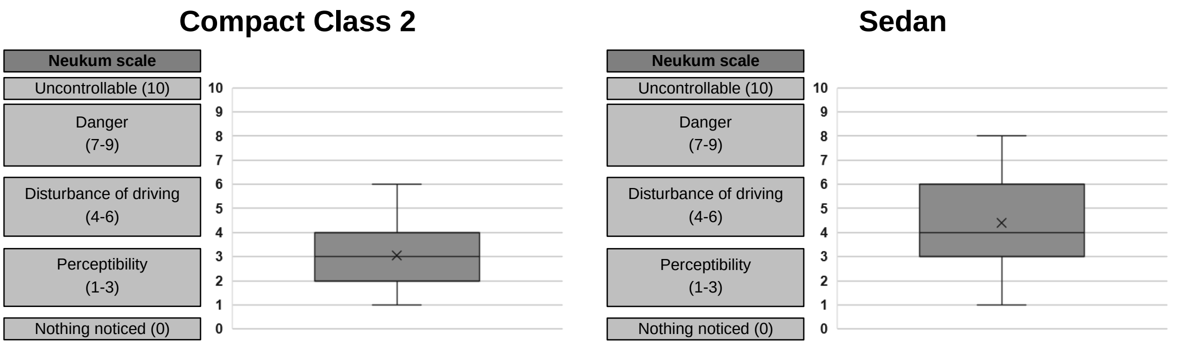
Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	11	7,91	4,51	1	15
	after lap 6 + failure	11	9,82	4,17	5	19
Physical demands	after lap 5	11	7,18	4,00	1	13
	after lap 6 + failure	11	7,64	4,80	1	17
Subjective performance	after lap 5	11	14,82	2,89	12	20
	after lap 6 + failure	11	15,64	2,73	12	20

At the time of the failure activation, the behavior of the test subjects was observed by the test supervisor with the aim of recording the initial reaction. 40.91 % (*n* = 9) of the test subjects showed a noticeable reaction when the *uncontrolled* RWA was triggered on. This was most frequently expressed verbally (*n* = 6) or by looking in the mirror/looking backwards (*n* = 3). In order to find out how many test subjects subjectively noticed the failure activation, they were asked after completing the maneuver with failure activation whether they noticed anything in particular. 100% of the test subjects (*n* = 22) answered this question in the affirmative. When describing the event, the most common response (*n* = 11) was that a cone or something else was run over. A jolt/vibration was described almost as frequently (*n* = 10), as was the vehicle sliding away (*n* = 9). The test subjects were then asked to describe their reaction to the event on an affective, cognitive and behavioral level. On an affective level, feelings of shock/surprise were expressed most frequently (*n* = 14). On a cognitive level, the subjects most frequently reported thinking about having knocked down a pillar/something else (*n* = 11) or thinking that something was different/not okay (e.g. flat tire) (*n* = 7). With regard to the behavioral reaction, it was mentioned most frequently (*n* = 12) that the test subjects had reacted to the failure activation with counter-steering/correcting. The second most common response (*n* = 6) was to continue driving normally/look at the road. When asked whether they would react differently if the event occurred again, some people (*n* = 3) replied that they would stop and look at the reason for the event, or that they would be more relaxed if the event occurred again (*n* = 3).

The 11-point Neukum scale (0 = nothing noticed - 10 = vehicle not controllable; Neukum & Krüger, 2003) was used to survey the subjectively perceived criticality of the failure activation. As the failure type *Uncontrolled RWA was* not only used in the sub-study with the compact class 2, but also in the sub-study with the sedan (for details see 3.5) using their test vehicle, the results of both studies are compared with regard to the assessment of the criticality of the event. However, due to the simultaneous variation of failure parameterization and test vehicle, potential differences can only be pointed out, but not explained. All valid data sets were considered for the analysis. The mean subjectively perceived criticality of the *Uncontrolled RWA* failure was classified at a high level of the noticeability category in the sub-study with the compact class 2 (*M* = 3.05, *SD* = 1.29). In the sub-study with the sedan, the criticality of the failure pattern was reported on average at a low level in the disruption to driving category (*M* = 4.40, *SD* = 1.66). For a representation of both distributions see Illustration 43.

Illustration 43

Subjectively perceived criticality of the failure type uncontrolled SHE in the two sub-studies with the compact class 2 and sedan vehicles, measured using the Neukum scale (Neukum & Krüger, 2003).



N	M	SD	Median	Percentile		
				25	50	75
22	3,05	1,29	3,00	2,00	3,00	4,00

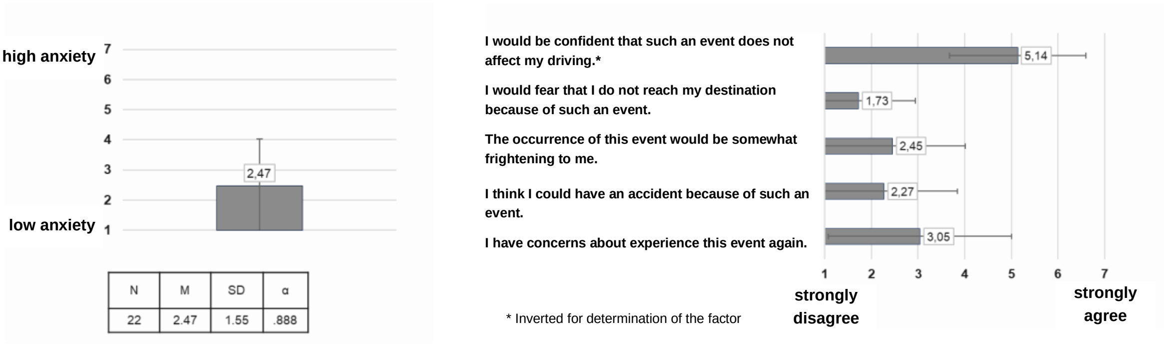
N	M	SD	Median	Percentile		
				25	50	75
25	4,40	1,66	4,00	3,00	4,00	6,00

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

Anxiety with noticing the failure event appears at a low level on the 7-point rating scale (1 = low anxiety - 7 = high anxiety) (*M* = 2.47, *SD* = 1.55). For a presentation of the factor and the individual items used for the calculation, see Illustration 44.

Illustration 44

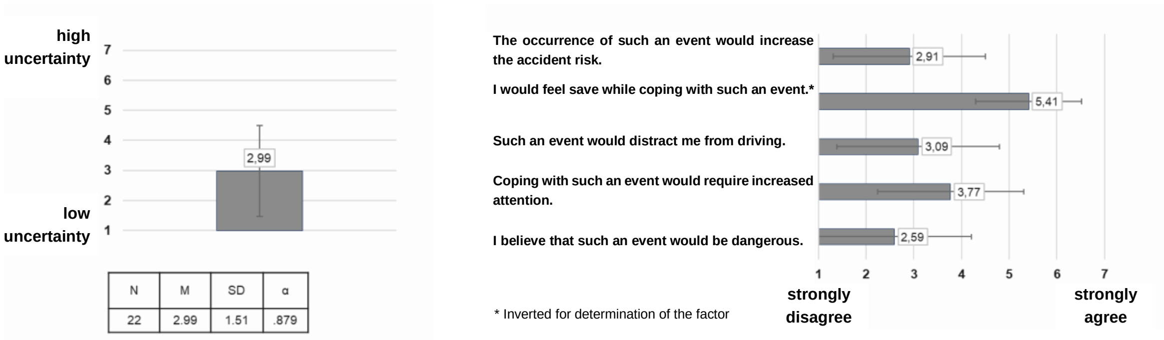
Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the Uncontrolled RWA failure event.



The perceived uncertainty with noticing the failure is classified on a lower to medium range of the 7-point scale (1 = low uncertainty - 7 = high uncertainty) (*M* = 2.99, *SD* = 1.51). For a presentation of the results of the factor and the individual items, see Illustration 45.

Illustration 45

Results of the Perceived Safety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the failure event Uncontrolled RWA.

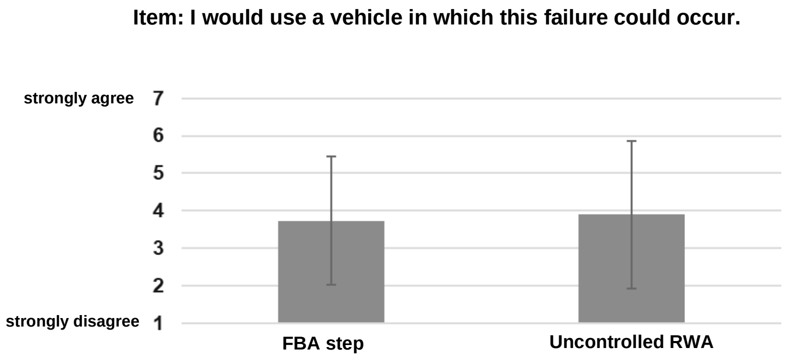


3.2.2.6 Follow-up survey

After the test subjects had been informed about the subject of the study following the test drives, a follow-up survey was conducted. The test subjects only answered the questions if they had noticed the relevant failure during the test. The test subjects were asked to indicate the extent to which they agreed with the statement "I would use a vehicle in which this failure could occur". Of the 71.43% (n = 15) test subjects who had noticed the *FBA step failure,* the intention to use the vehicle was rated in the middle of the 7-point response scale (1 = strongly disagree - 7 = strongly agree) (M = 3.73, SD = 1.71). The failure *Uncontrolled SHE* was assessed by all test subjects (N = 21) and was also rated by them at a medium level of the scale (M = 3.90, SD = 1.97). For a presentation of the results, see Illustration 46.

Illustration 46

Intention to use a vehicle that could exhibit the experienced failure FBA step or uncontrolled RWA, surveyed using a 7-point Likert scale.



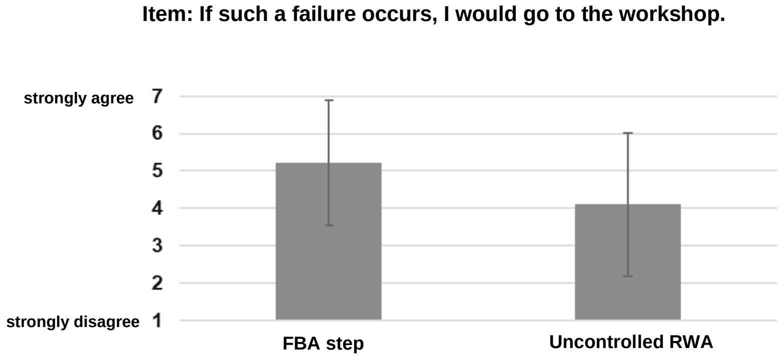
Failure	N	M	SD	Min	Max
FBA step	15	3,73	1,71	1	6
Uncontrolled **RWA**	21	3,90	1,97	1	7

Note. The item was only answered by test subjects who noticed the respective failure during the test run.

Furthermore, the test subjects were asked how much they agreed with the statement "If such a failure occurs, I would drive to the garage". The people who noticed the FBA failure (n = 19) tended to agree with this statement (M = 5.21, SD = 1.67). For the RWA failure, the test subjects appear to neither agree nor disagree with this statement on average (M = 4.10, SD = 1.92). For a presentation of the results, see Illustration 47.

Illustration 47

Intention to visit a workshop if the experienced failures FBA step or uncontrolled SHE in own vehicle occur, measured using a 7-point Likert scale.



Failure	N	M	SD	Min	Max
FBA step	14	5,21	1,67	2	7
Uncontrolled **RWA**	20	4,10	1,92	2	7

3.2.3 Summary of study 2

The following is an overview of the main results of sub-study 2 (vehicle: compact class 2; failure: *FBA step* / *uncontrolled RWA*).

- The design of the course appears appropriate, the test subjects are neither under- nor overchallenged.
- A learning curve of the test subjects over the course of the test becomes clear.
- The research hypothesis was retained for both failure patterns.
- The *FBA step* failure
  - was noticed by 73% of the test subjects;
  - does not cause a significant increase in the perceived mental and physical demands or a significant deterioration in the subjectively perceived performance in the driving maneuver in question;
  - is described by the majority of test subjects as an intervention in the steering;
  - is mostly associated on an affective level with astonishment, uncertainty or no specific affective reaction;
  - leads to counter-steering or corrective behavior in some of the test subjects at the level of behavioral response;
  - is classified at a medium level in the "noticeability" category in terms of criticality.
- The failure *Uncontrolled RWA*
  - was noticed by all test subjects;

- does not cause a significant increase in the perceived mental and physical demands or a significant deterioration in the subjectively perceived performance in the driving maneuver in question;
- is mostly associated on an affective level with feelings of shock and surprise;
- leads to counter-steering or corrective behavior in some of the test subjects at the level of behavioral response;
- is classified at a high level of "tangibility" in terms of criticality.

The content of the results of the sub-studies is discussed for the entire test series in Section 4.

3.3 Study 3

The third sub-study was carried out together with the compact class 1 vehicle in the period from April 7 to 12, 2022, using the associated test vehicle. The selected failure patterns were the *FBA Selfsteer + Loss of Feedback,* which was activated in the maneuver Circular drive, as well as the *Blocked RWA* in the associated maneuver Slalom. The results are shown below.

3.3.1 Sample

N = 25 people took part in the study, 10 of whom were female. The mean age is *M* = 41.80 years (*SD* = 13.40), with the youngest subject being 21 years old and the oldest 62 years old. The average annual mileage is *M* = 16,800.00 km (*SD* = 11,622.54 km) with a range of 3,000 km - 60,000 km. None of the test subjects stated that they had an uncorrected visual or hearing impairment.

3.3.2 Results

The presentation of the results is divided into two sections: failure-independent and failure-specific results. Firstly, the results are reported which relate to the general difficulty of the selected maneuvers in the sense of testing the selected test design, independently of the failure setups. The failure-specific results are then presented. In this section, the results for testing the controllability of both failures (FBA and RWA) are discussed first. This is followed by the subjective test data regarding the experience of the failure activation and the objective test data regarding the driver and vehicle reaction as a result of the failure activation, first for the FBA failure and then for the RWA failure.

3.3.2.1 Failure-independent results

The generally perceived difficulty of the maneuvers was recorded on a 20-point scale (1 = very low to 20 = very high) based on the NASA-TLX instrument (Hart, 2006) using three questions regarding the mental and physical demands as well as the subjective assessment of one's own performance. The surveys were conducted at the beginning (after the familiarization ride) and before the end of the test (following ride 5). For the inferential statistical analysis of the learning curve, the requirements at the beginning and end of the ride were compared using a paired *t-test*. All test subjects with valid data sets for both failure cases were included in the analysis.

At both survey times, the three maneuvers of driving straight-ahead, circular drive and slalom were rated in the low to medium range of the 20-point scale in terms of mental demands. The maneuver driving straight-ahead appears to be associated with the comparatively lowest mental demands. For a presentation of the characteristic values, see Table 14. Before the last ride, all three maneuvers were perceived as significantly less mentally demanding than at the beginning of the test (straight driving: $t(12) = 2.16$, $p = .026$, $d = 0.6$; circular drive: $t(12) = 3.4$, $p = .003$, $d = 0.94$; slalom: $t(12) = 2.12$, $p = .028$, $d = 0.59$).

Table 14

Mental requirements of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high requirements) based on the NASA TLX instrument (Hart, 2006).

Mental requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	13	5,15	3,11	1	10
	before the last round	13	4,08	2,50	1	8
Circular drive	after familiarization	13	10,54	3,13	4	16
	before the last round	13	8,00	4,10	3	15
Slalom	after familiarization	13	11,31	3,25	5	16
	before the last round	13	9,00	4,26	3	16

In terms of physical demands, the three maneuvers are also classified in a low to medium range, with driving straight-ahead appearing to be the least demanding. For a presentation of the characteristic values, see Table 15. Over the course of the test, the physical demands of driving straight-ahead and circular drive remain constant (driving straight-ahead: $t(12) = -1.07$, $p = .15$; circular drive: $t(12) = 1.66$, $p = .06$), while they decrease significantly in the slalom ($t(12) = 2.27$, $p = .027$, $d = 0.63$).

Table 15

Physical demands of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high demands) based on the NASA TLX instrument (Hart, 2006).

Physical requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	13	3,00	2,27	1	7
	before the last round	13	3,54	2,22	1	7
Circular drive	after familiarization	13	7,15	3,51	2	14
	before the last round	13	5,92	3,57	1	13
Slalom	after familiarization	13	9,92	4,39	2	18
	before the last round	13	8,31	4,42	3	18

Over the course of their studies, students rated their own performance in the high range of the 20-point scale. For a presentation of the characteristic values, see Table 16. At the end of the test, performance in the maneuvers straight-ahead and slalom was rated significantly better than at the beginning (straight-ahead: $t(12) = -3.55$, $p = .002$, $d = -0.99$; slalom: $t(12) = -3.15$, $p = .004$, $d = -0.87$), whereas it remained constant in the maneuver circular drive ($t(12) = -1.77$, $p = .05$).

Table 16

Assessment of own performance in the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high performance) based on the NASA TLX instrument (Hart, 2006).

Assessment of own performance						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	13	16,38	2,60	12	20
	before the last round	13	17,92	1,80	15	20
Circular drive	after familiarization	13	14,77	3,66	6	20
	before the last round	13	16,69	2,84	12	20
Slalom	after familiarization	13	14,85	3,18	10	20
	before the last round	13	17,38	2,26	14	20

3.3.2.2 Failure-specific results

In the following, the results for testing the controllability of the two failure patterns *FBA Selfsteer + Loss of Feedback* and *Blocked RWA are* discussed first. This is followed by a separate presentation of the subjective results for both failures with regard to the experience of the failure activation and then the objective test data with regard to the driver and vehicle reaction as a result of the failure events.

3.3.2.2.1 Controllability - hypothesis testing

In the context of the third sub-study, 100% of the test subjects with valid data records did not leave the lane with failure activation for the *FBA Selfsteer + Loss of Feedback* failure activated during the maneuver. 100% of the data records comprise at least N = 20 data records. The hypothesis can therefore be maintained for this selected combination of failure parameterization and vehicle. Within the scope of the study conducted, this failure was controllable for all test subjects at the C0 level defined by experts. Due to an insufficient number of valid data sets (n = 15), it was not possible to test the hypothesis for the *blocked SHE failure.*

3.3.2.2.2 Failure type: *FBA Selfsteer + Loss of* Feedback

The objective results describing the vehicle and driver reaction as a result of the FBA failure *Selfsteer + Loss of Feedback*, as well as the results regarding the subjective perception of the activation are reported below.

3.3.2.2.3 Results of objective vehicle measurement data

The statistical distributions of the characteristic objective values for the FBA Selfsteer+ Loss of Feedback failure pattern in the compact class 1 determined on the basis of the recorded vehicle measurement variables are shown in Illustration 48 shown. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern investigated are also shown in Table 17. For a qualitative comparison of the objective parameters determined, the following Illustration 48 also shows the statistical evaluations for the second vehicle, which was examined with the same failure pattern as part of the overall study. However, due to vehicle- and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.

Illustration 48
Statistical evaluation of the calculated objective parameters of steering angle Disturbance, Disturbance steering rate, Disturbance lateral acceleration and Disturbance yaw rate for the FBA Selfsteer+ Loss of Feedback failure pattern
(white: compact class 1; gray: SUV 1)

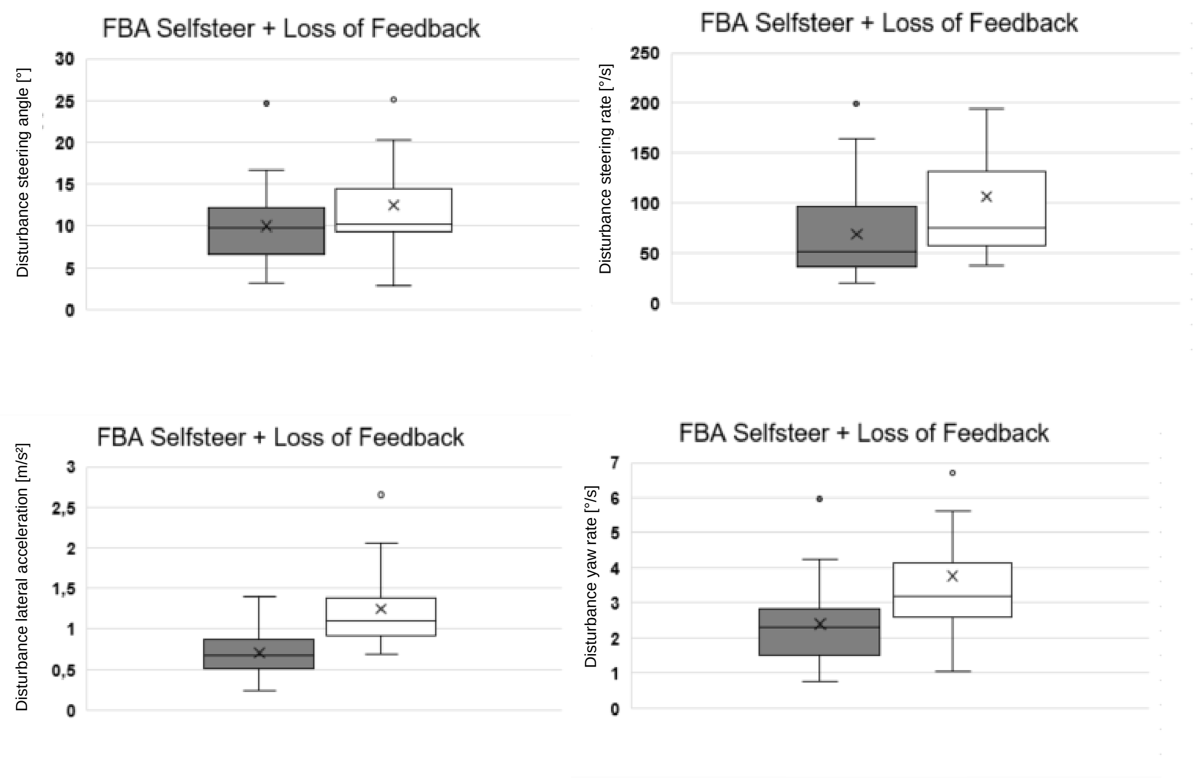


Table 17

Descriptive statistical parameters FBA Selfsteer+ Loss of Feedback *failure pattern (compact class 1)*

Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
Compact Class 1	**Disturbance steering angle [°]**	22	12,48	6,46	10,27	2,84	32,66
	Disturbance steering rate[°/s]	22	106,61	69,75	75,14	36,78	301,70
	Dist. lateral acceleration [m/s²]	22	1,24	0,47	1,10	0,69	2,65
	Disturbance yaw rate [°/s]	22	3,77	1,99	3,20	1,06	9,66

In this case, the compact class 1 compared to the SUV 1 tends to show slightly higher disturbances with regard to the *FBA Selfsteer+ Loss of Feedback* failure pattern in relation to the driver's steering response and the resulting vehicle dynamics. However, it should be noted that at 20 ms, the activation time of the torque step in the compact class 1 is also twice as long as the activation time of the torque step in the SUV 1.

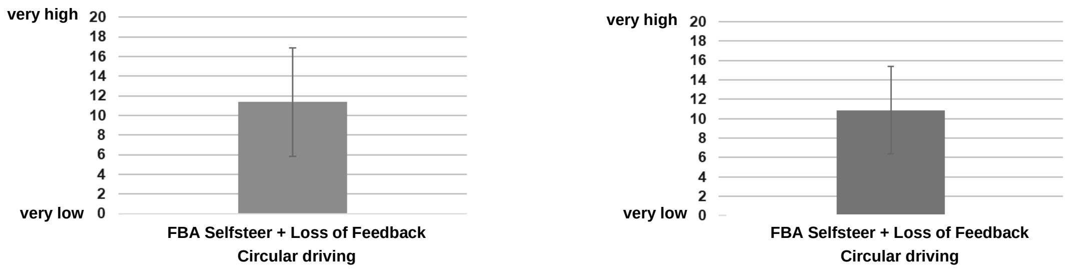
3.3.2.2.4 Results of subjective measures

In the following, the perceived difficulty of the maneuver of circular drive with failure activation is discussed first. This was recorded using three items based on the NASA-TLX instrument with regard to the mental and physical demands experienced and the subjective assessment of one's own performance on a 20-point response scale (1 = very low to 20 = very high). The results are compared with those of sub-study 1 (for details see 3.1) with the SUV 1 vehicle, as the same type of failure was taken into account here. Differences between the sub-studies can be described, but cannot be explained due to the simultaneous manipulation of failure parameterization and test vehicle. All valid data sets were included in the analysis. In the sub-study with compact class 1 vehicle used, the mental requirements for completing the maneuver circular drive with failure activation were reported in the middle range of the 20-point response scale (M = 11.36, SD = 5.05). Similarly, the mental requirements for the sub-study with the SUV 1 vehicle using their test vehicle for the maneuver circular drive with failure activation were classified in the middle range of the scale (M = 10.88, SD = 4.48). For a presentation of the results, see Illustration 49.

Illustration 49

Mental requirements of the maneuver of circular drive with activation of the FBA Selfsteer + Loss of Feedback failure were assessed on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).

Compact Class 1 **SUV 1**



N	M	SD	Min	Max
22	11,36	5,05	4	20

N	M	SD	Min	Max
24	10,88	4,48	2	20

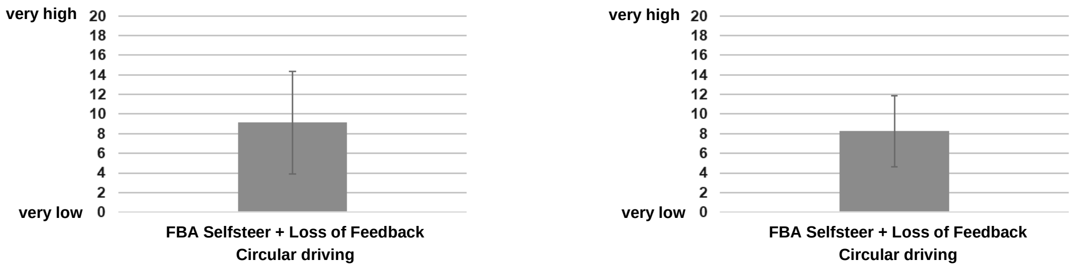
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The physical requirements for completing the maneuver circular drive with failure activation were assessed in both sub-studies in the middle range of the 20-point response scale (compact class 1: *M* = 9.14, *SD* = 5.22; SUV 1: *M* = 8.25, *SD* = 3.60). For a presentation of the results, see Illustration 50.

Illustration 50

Physical demands of maneuvering in a circular drive with activation of the failure FBA Selfsteer + Loss of Feedback assessed on a 20-point response scale (1 = very low to 20 = very high demands) based on the NASA-TLX instrument (Hart, 2006).

Compact Class 1 **SUV 1**



N	M	SD	Min	Max
22	9,14	5,22	1	18

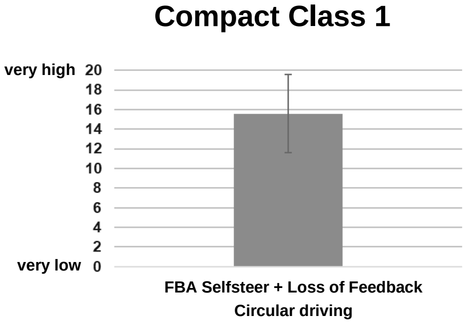
N	M	SD	Min	Max
24	8,25	3,60	3	15

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

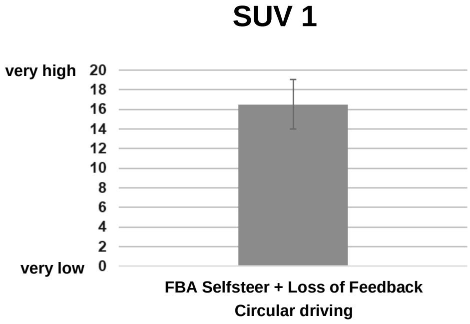
In both sub-studies, the subjective assessment of one's own performance in completing the maneuver circular drive with failure activation is in the medium to high range of the 20-point scale (compact class 1: *M* = 15.59, *SD* = 4.00; SUV 1: *M* = 16.50, *SD* = 2.52). For a presentation of the results, see Illustration 51.

Illustration 51

Subjective performance in completing the maneuver of circular drive with activation of the failure FBA Selfsteer + Loss of Feedback on a 20-point scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
22	15,59	4,00	4	20



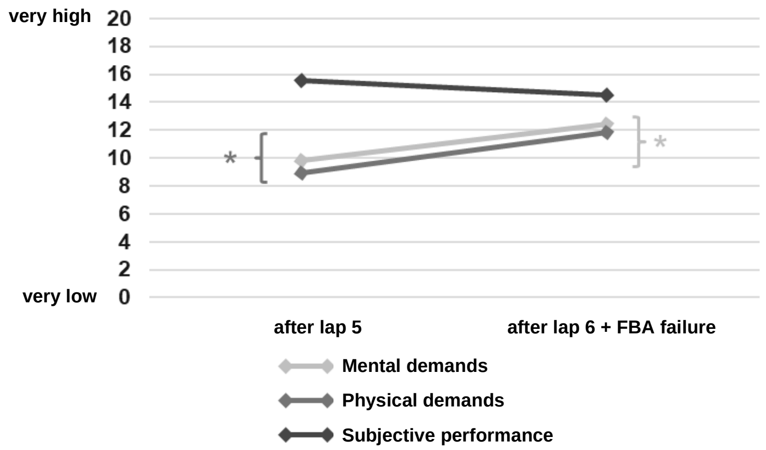
N	M	SD	Min	Max
24	16,50	2,52	12	20

Remark. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization

To identify the influence of the failure event on the perceived difficulty of the maneuver, the perceived mental and physical demands as well as the subjective performance assessment after round 5 without failure activation were compared with those after failure activation in round 6. This approach was based on the assumption that at this point the learning curve of the test subjects regarding the completion of the maneuver is flattened and differences between the assessments are due to the activation of the failure. Due to the randomized activation of both failures of this sub-study (RWA and FBA failures) in rounds 3 and 6, the number of data sets that can be included in the analysis is reduced to about half of the valid data sets (in which the respective failure was activated in round 6). The inferential statistical analysis of the comparison of the two survey times was carried out using a paired *t-test*. In this sub-study, it became clear that both the mental and physical demands increased significantly with the activation of the FBA failure in round 6 compared to the previous round 5, without exceeding the middle range of the scale (mental demands: $t(10) = -2.21$, $p = .03$, $d = -0.64$, physical demands: $t(10) = -2.55$, $p = .015$, $d = -0.77$). The comparison of subjectively perceived personal performance showed no significant difference between the two rounds ($t(10) = 0.91$, $p = .19$). For a graphical representation of the progression and the associated characteristic values, see Illustration 52.

Illustration 52

*Subjectively perceived difficulty measured using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands and the subjectively perceived performance in the maneuver circular drive in round 5 without failure activation compared to round 6 after activation of the failure FBA Selfsteer + Loss of Feedback. (Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



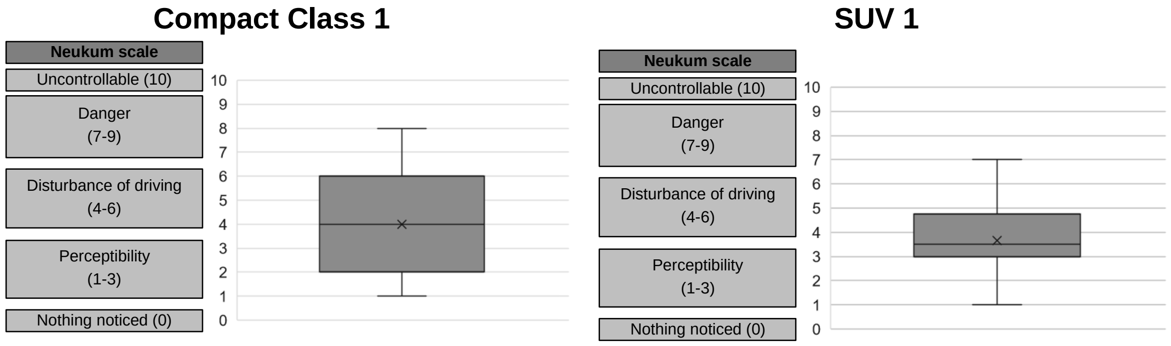
Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	11	9,82	4,14	4	18
	after lap 6 + failure	11	12,45	4,89	6	20
Physical demands	after lap 5	11	8,91	3,39	5	16
	after lap 6 + failure	11	11,82	4,26	5	18
Subjective performance	after lap 5	11	15,55	2,91	11	20
	after lap 6 + failure	11	14,55	4,41	4	20

At the time the failure was activated, the behavior of the test subjects was observed by the experimenter to record the initial reaction to the event. 45.45% (*n* = 10) of the test subjects showed a reaction observable by the experimenter when the FBA failure *Selfsteer* + *Loss of Feedback* occurred, which was most frequently shown by a verbal utterance (*n* = 7) and less frequently by a glance in the mirror (*n* = 2). After completing the maneuver, the test subjects were also asked whether they had noticed anything special during the maneuver. 100 % of the test subjects (*n* = 22) answered yes to this question. When describing the event, steering intervention (e.g. by a lane assistant) was mentioned most frequently (*n* = 19). The assumption of having driven over a cone or something else was mentioned much less frequently (*n* = 4). Following on from this, the test subjects were asked to describe their reaction to the event on an affective, cognitive and behavioral level. To record the affective reaction, the subjects were asked to describe their feelings associated with the perception of the event. The most frequently mentioned description was surprise/fright (*n* = 15). Irritation or an unpleasant feeling were mentioned much less frequently, although they were the second most common, with *n* = 3 mentions each. To record the reaction on a cognitive level, the test subjects were asked to describe their thoughts at the time of the event. The most frequently mentioned question in this context was "What was that?" (*n* = 4). In addition, thoughts were expressed with a frequency of *n* = 3 each of being too far to the right / left in the lane so that the vehicle had to intervene, of having run over a cone or something else, or that it could have been one's own error, an initiated failure or a failure of the vehicle. To describe the behavioral reaction, the test subjects were asked to report what they did when the event occurred. The most common response (*n* = 10) was to counter-steer/correct, followed by the statement that they had continued driving unchanged (*n* = 8).

The Neukum scale (Neukum & Krüger, 2003) was used to survey the subjectively perceived criticality of the failure recording. At this point, the results are again compared with those from sub-study 1 (for details see 3.1) with the vehicle SUV 1, in which the same type of failure was taken into account. It should be noted once again that potential differences in the results of the two sub-studies can be described, but cannot be explained due to the parallel manipulation of failure parameterization and test vehicle. All valid data sets were considered for the analysis. The mean perceived criticality of the activation of the FBA failure in the sub-study with the compact class 1 was in a lower range of the category disturbance of driving on the 11-point response scale (0 = nothing noticed - 10 = vehicle not controllable) (*M* = 4.00, *SD* = 1.92). In the sub-study with the SUV 1 test vehicle, criticality was classified on average at a high level in the noticeability category and a low level in the driving disturbance category (*M* = 3.67, *SD* = 1.27). The spread in the sub-study with the compact class 1 appears wider. For a representation of both distributions, see Illustration 53.

Illustration 53

Subjectively perceived criticality of the failure types FBA Selfsteer + Loss of Feedback in the two sub-studies with the compact class 1 and SUV 1 vehicles using the Neukum Scale (Neukum & Krüger, 2003).



N	M	SD	Median	Percentile		
				25	50	75
20	4	1,92	4,00	2,00	4,00	6,00

N	M	SD	Median	Percentile		
				25	50	75
24	3,67	1,27	3,50	3,00	3,50	4,75

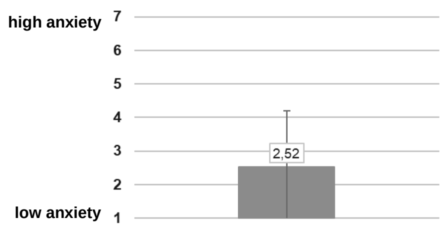
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

Two adapted rating scales of the factors *Anxiety* and *Perceived Safety* of the *Car Technology Acceptance Model* (Osswald et al., 2012) were used to assess the experience of the situation. This data was collected from all test subjects who stated that they had noticed the failure.

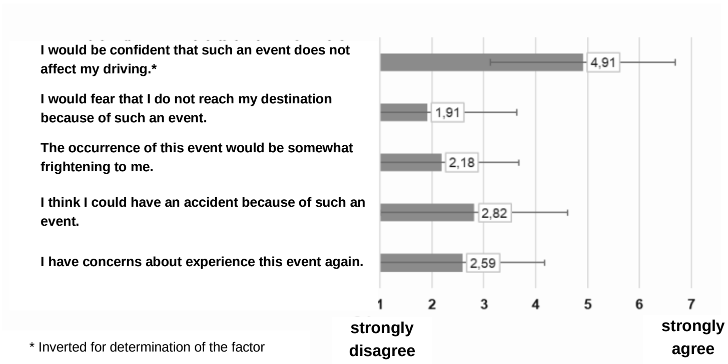
Anxiety with the occurrence of the failure event was rated on average in a low range of the 7-point scale (1 = low anxiety - 7 = high anxiety) (M = 2.52, SD = 1.68). For an illustration of the *anxiety* factor and the individual items used in the calculation, see Illustration 54.

Illustration 54

Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the failure event FBA Selfsteer + Loss of Feedback.



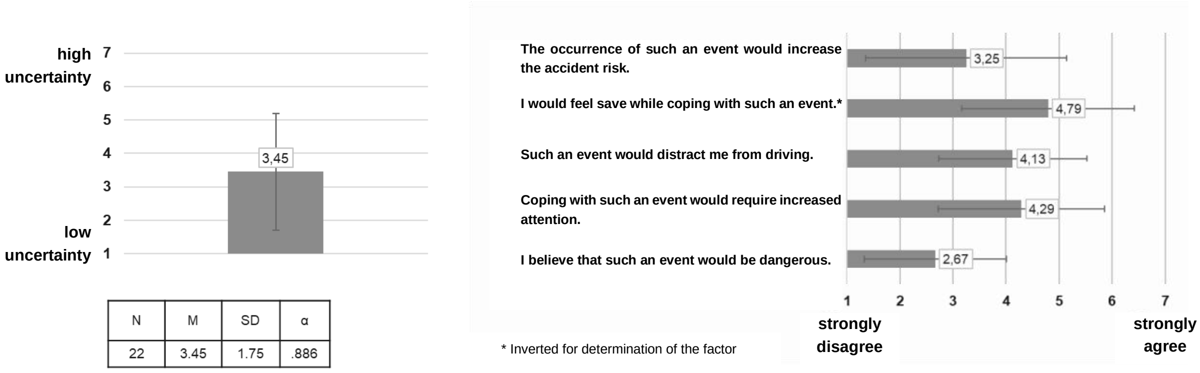
N	M	SD	α
22	2.52	1.68	.871



The perceived uncertainty when the failure event occurs is classified in a medium range of the 7-point scale (1 = low uncertainty - 7 = high uncertainty) (M = 3.45, SD = 1.75). For a presentation of the results of the *Perceived Safety* factor and the individual items, see Illustration 55.

Illustration 55

Results of the factor Perceived Safety (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the failure event FBA Selfsteer + Loss of Feedback.



N	M	SD	α
22	3.45	1.75	.886

3.3.2.2.5 Failure type: *Blocked* RWA

In the following, the objective results regarding the description of the vehicle and driver reaction as a result of the activation of the *blocked RWA* failure in the slalom maneuver are reported, as well as the results regarding the subjective perception of the failure activation.

3.3.2.2.6 Results of objective vehicle measurement data

The statistical distributions of the characteristic objective values for the *blocked RWA* failure pattern in the compact class 1 determined on the basis of the recorded vehicle measurement variables are shown in Illustration 56 is shown. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern investigated are also shown in Table 18. For a qualitative comparison of the objective parameters determined, the following

Illustration 56 also shows the statistical evaluations for the second vehicle, which was examined with the same failure pattern as part of the overall study. However, due to vehicle and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.

In the present case, the SUV 1 tends to show a slightly lower disturbance influence with regard to the steering angle compared to the compact class 1 with regard to the *blocked RWA* failure pattern. In contrast, the disturbance steering rates measured as a result of the failure activation are slightly higher than in the compact class 1. With regard to the driving dynamics disturbance variables of lateral acceleration and yaw rate, no major differences between the two vehicles are discernible.

Illustration 56

Statistical evaluation of the calculated objective parameters of Disturbance steering angle, Disturbance steering rate, Disturbance lateral acceleration and Disturbance yaw rate for the blocked RWA failure pattern

(white: compact class 1; gray: SUV 1)

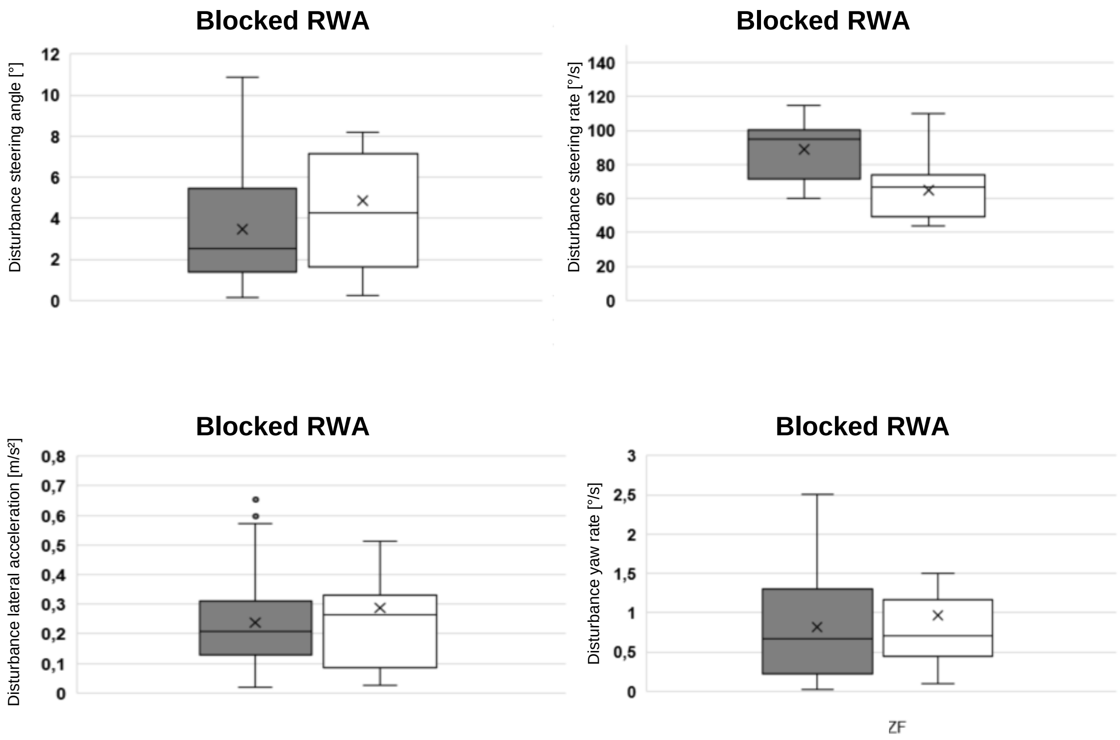


Table 18

Descriptive statistical parameters Failure pattern blocked RWA (compact class 1)

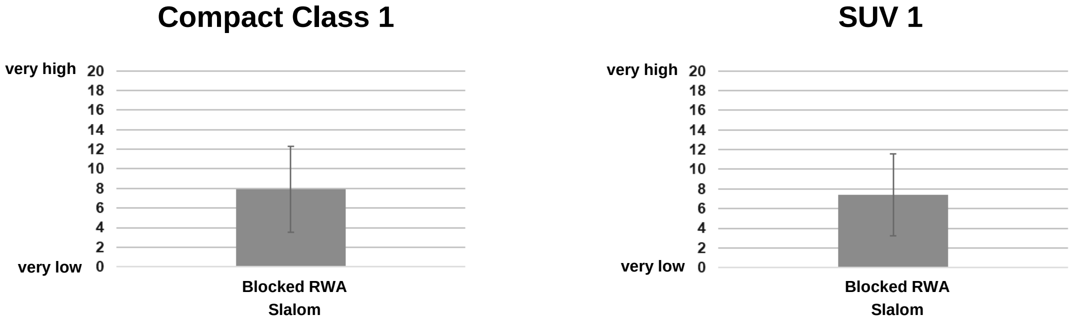
Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
Compact Class 1	**Disturbance steering angle [°]**	15	4,84	3,79	4,26	0,25	16,26
	Disturbance steering rate[°/s]	15	64,70	16,34	66,81	43,65	110,24
	Dist. lateral acceleration [m/s²]	15	0,29	0,28	0,26	0,02	1,17
	Disturbance yaw rate [°/s]	15	0,97	0,87	0,71	0,10	3,83

3.3.2.2.7 Results of subjective measures

The following section deals with the perceived difficulty of the slalom maneuver with failure activation, which was assessed using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands, as well as the subjective assessment of one's own performance on a 20-point scale (1 = very low to 20 = very high). The results are presented by comparing the data with those collected in sub-study 1 (for details see 3.1) with the SUV 1 test vehicle with regard to the same failure type. Potential differences between the two sub-studies can be shown, but cannot be explained due to the simultaneous variation of failure parameterization and test vehicle. All valid data sets were included in the analysis. In both sub-studies, the mental requirements for completing the slalom maneuver with failure activation were assessed in the middle range of the 20-point response scale (compact class 1: *M* = 7.93, *SD* = 4.38; SUV 1: *M* = 7.40, *SD* = 4.19). For a presentation of the results, see Illustration 57.

Illustration 57

Mental requirements of the slalom maneuver with activation of the blocked RWA failure were assessed on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
15	7,93	4,38	3	15

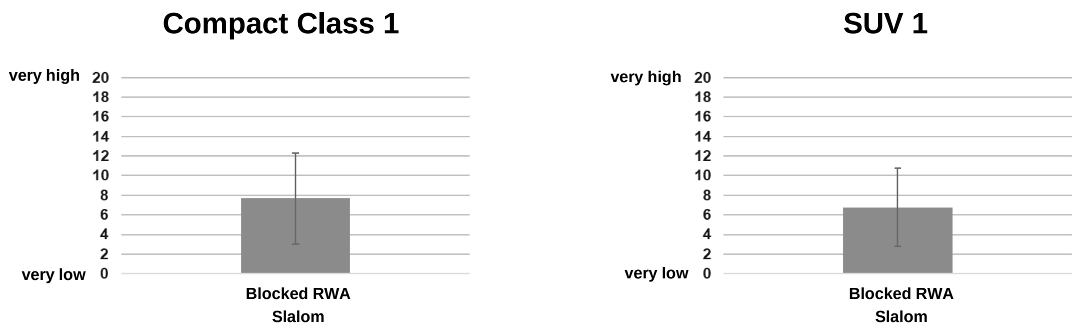
N	M	SD	Min	Max
25	7,40	4,19	1	15

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The physical requirements for completing the slalom maneuver with failure activation were classified in a low to medium range of the 20-point response scale in both sub-studies (compact class 1: *M* = 7.67, *SD* = 4.64; SUV 1: *M* = 6.76, *SD* = 4.00). For a presentation of the results, see Illustration 58.

Illustration 58

Physical requirements of the slalom maneuver with activation of the blocked RWA failure were assessed on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
15	7,67	4,64	2	18

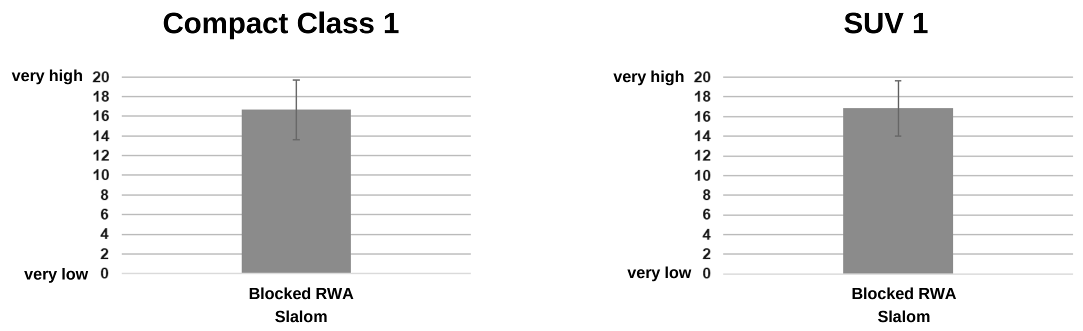
N	M	SD	Min	Max
25	6,76	4,00	2	17

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The subjective performance with regard to completing the slalom maneuver with failure activation was rated in the medium to high range on the 20-point scale in both sub-studies (compact class 1: M = 16.67, SD = 3.02; SUV 1: M = 16.84, SD = 2.84). For a presentation of the results, see Illustration 59.

Illustration 59

Subjective performance in completing the slalom maneuver with activation of the blocked RWA failure measured on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).



N	**M**	**SD**	**Min**	**Max**
15	16,67	3,02	10	20

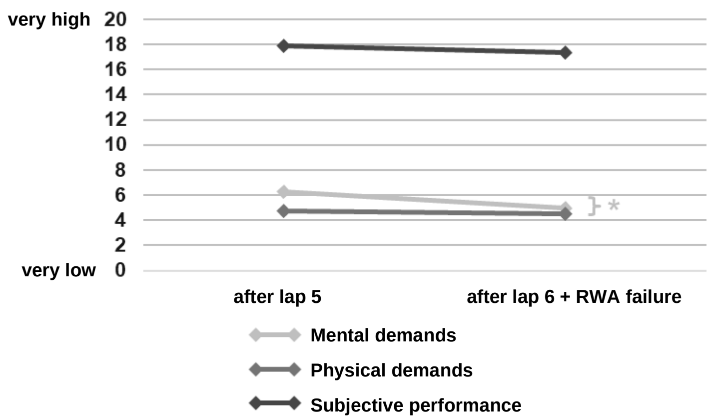
N	**M**	**SD**	**Min**	**Max**
25	16,84	2,84	8	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

In order to find out how great the influence of the failure event is on the perceived difficulty of the slalom maneuver, a comparison was made of the perceived mental and physical demands, as well as the subjective assessment of one's own performance after round 5 without failure activation with those after failure activation in round 6. The inferential statistical analysis of the comparison of the two survey times was carried out using a paired *t-test*. It was found that the mental demands were statistically significantly lower after round 6 with failure activation compared to round 5 without failure activation, without leaving the low to medium range of the scale ($t(7) = 3.04$, $p = .009$, $d = 1.07$). The assessment of physical demands and subjective performance remained constant over both rounds (physical demands: $t(7) = 1.00$, $p = .18$; subjective performance: $t(7) = 1.08$, $p = .16$). For an illustration of the course and the associated characteristic values, see Illustration 60.

Illustration 60

*Subjectively perceived difficulty assessed using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands as well as the subjectively perceived performance in the slalom maneuver in round 5 without failure activation compared to round 6 after activation of the blocked RWA failure. (Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	**8**	6,25	3,45	3	14
	after lap 6 + failure	8	5,00	3,02	3	12
Physical demands	after lap 5	8	4,75	2,44	2	10
	after lap 6 + failure	8	4,50	2,45	2	10
Subjective performance	after lap 5	8	17,88	2,80	13	20
	after lap 6 + failure	8	17,38	4,00	10	20

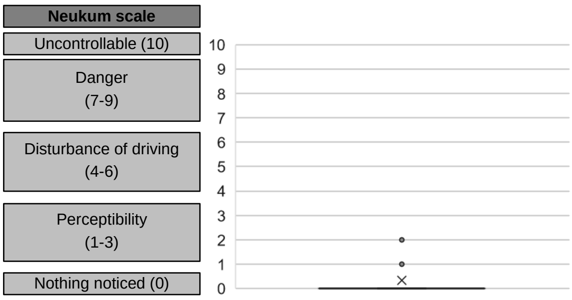
At the time the failure was activated, behavioral observation was carried out by the test supervisor with the aim of recording the initial reaction of the test subjects. When the SHE failure was activated, however, no reaction observable by the test administrator was observed in any of the test subjects. When asked, 20% of the test subjects ($n = 3$) stated that they had noticed something. It was described that the vehicle braked slightly (*n* = 1) or that there was a slight rumble (*n* = 1). The test subjects who had noticed something special while completing the maneuver were then asked to describe their reaction on an affective, cognitive and behavioral level. No specific reaction could be identified on either an affective or cognitive level. With regard to the behavioral reaction to the failure event, participants described having increased their attention (*n* = 2) or having continued their drive unchanged (*n* = 2). If the event were to be repeated, none of the test subjects would react differently than shown in the experiment.

The Neukum scale (Neukum & Krüger, 2003) was used to assess the perceived criticality of the failure event. The results are again compared with those of sub-study 1 (for details see 3.1), which was carried out with the vehicle SUV 1 using their test vehicle. An explanation of potential differences between the sub-studies is not possible due to the parallel manipulation of failure parameterization and test vehicle. All valid data sets were taken into account for the analysis. In both sub-studies, the *blocked RWA* failure was either not noticed or classified at a low level of criticality in the noticeability category on the 11-point response scale (0 = not noticed - 10 = vehicle not controllable) (compact class 1: *M* = 0.33, *SD* = 0.72; SUV 1: *M* = 0.60, *SD* = 0.91). For a representation of both distributions, see Illustration 61.

Illustration 61

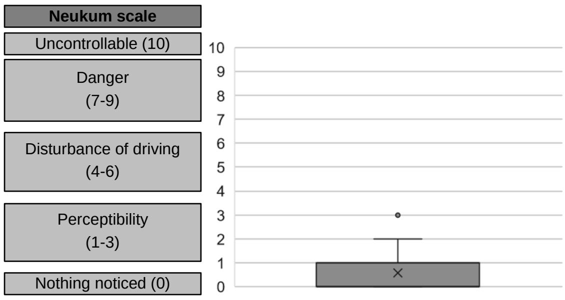
Subjectively perceived criticality of the blocked RWA failure type in the two sub-studies with the compact class 1 and SUV 1 vehicles using the Neukum scale (Neukum & Krüger, 2003).

Compact Class 1



N	M	SD	Median	Percentile		
				25	50	75
15	0,33	0,724	0,00	0,00	0,00	0,00

SUV 1



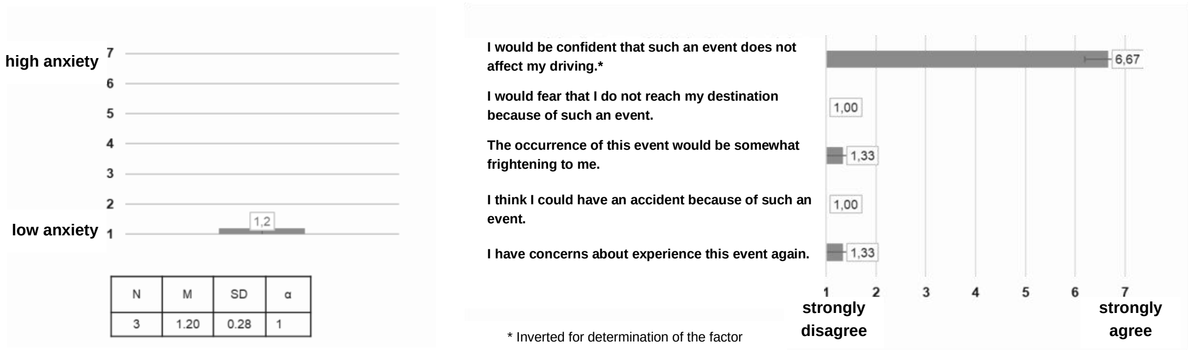
N	M	SD	Median	Percentile		
				25	50	75
24	0,60	0,91	0,00	0,00	0,00	1,00

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

Anxiety when noticing the failure event is classified at a low level on the 7-point scale (1 = low anxiety - 7 = high anxiety) (*M* = 1.20, *SD* = 0.28). For a presentation of the factor and the results of the individual items, see Illustration 62.

Illustration 62

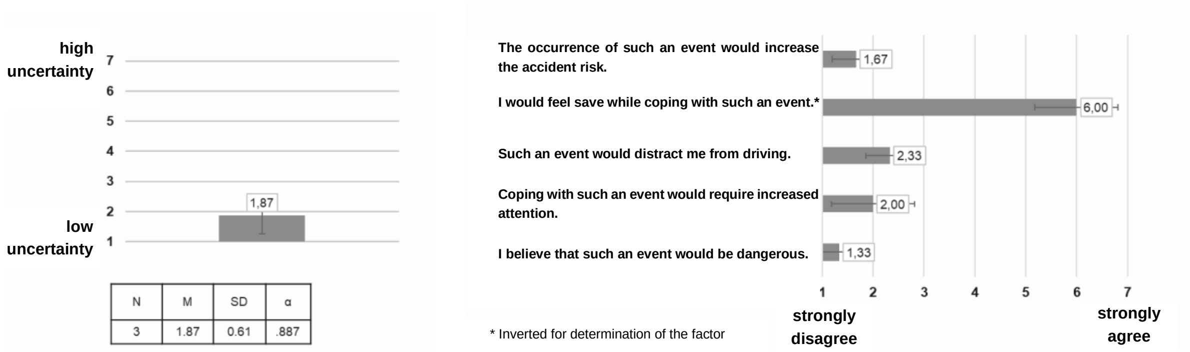
Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the blocked RWA failure event.



The perceived uncertainty with noticing the event was also classified in a low range of the 7-point scale (1 = low uncertainty - 7 = high uncertainty) (M = 1.87, SD = 0.61). For a presentation of the results of the factor and the underlying individual items, see Illustration 63.

Illustration 63

Results of the Perceived Safety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the blocked RWA failure event.

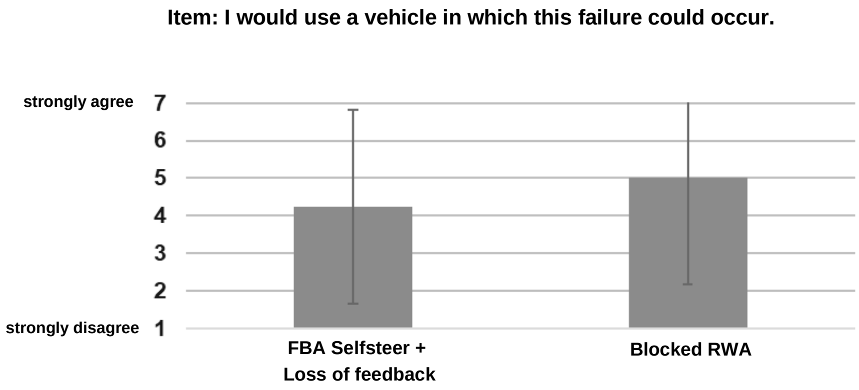


3.3.2.2.8 Follow-up survey

After the test subjects had been informed about the subject of the experiment, they were asked follow-up questions. The questions were only answered by the test subjects if they had noticed the respective failure during the test. First, the test subjects were asked to indicate how much they agreed with the statement "I would use a vehicle in which this failure could occur" on a 7-point scale (1 = strongly disagree - 7 = strongly agree). The *FBA Selfsteer + Loss of Feedback failure* was noticed by all test subjects (N = 13). The average intention to use is in the middle of the scale (M = 4.23, SD = 2.59). The failure *Blocked SHE* was only noticed by n = 2 people. They would be more likely to use a vehicle in which such a failure could occur (M = 5.00, SD = 2.83). For a presentation of the results, see Illustration 64.

Illustration 64

Intention to use a vehicle that could exhibit the experienced failure FBA Selfsteer & Loss of Feedback or Blocked RWA, surveyed using a 7-point Likert scale.



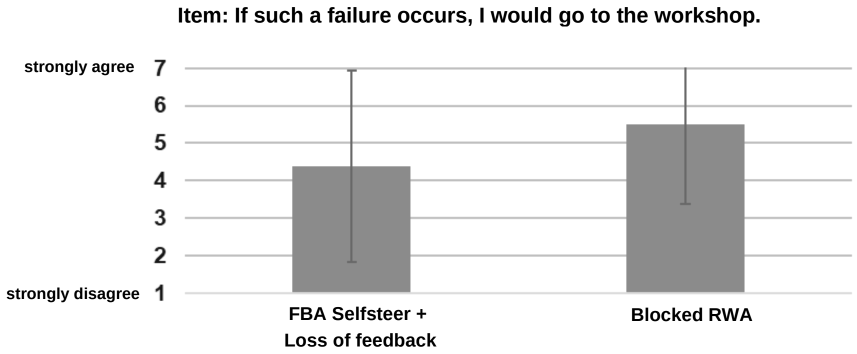
Failure	N	M	SD	Min	Max
FBA Selfsteer + Loss of Feedback	13	4,23	2,59	1	7
Blocked **RWA**	2	5,00	2,83	3	7

Note. The item was only answered by test subjects who noticed the respective failure during the test run.

The test subjects were also asked to indicate the extent to which they would agree to take the vehicle to a garage if such a failure occurred. For the FBA failure, there was neither a clear agreement nor a clear rejection of the statement (*M* = 4.38, *SD* = 2.57). The two test subjects who noticed the RWA failure were more likely to agree that they would like to visit a workshop (*M* = 5.50, *SD* = 2.12). See Illustration 65 for an illustration of the results.

Illustration 65

Intention to visit a workshop if the experienced failures FBA Selfsteer + Loss of Feedback or Blocked RWA occur in own vehicle, measured using a 7-point Likert scale.



Failure	N	M	SD	Min	Max
FBA Selfsteer + Loss of Feedback	13	4,38	2,57	1	7
Blocked **RWA**	2	5,50	2,12	4	7

3.3.3 Summary Study 3

The following is an overview of the main results of sub-study 3 (vehicle: compact class 1; failure: *FBA Selfsteer + Loss of Feedback / Blocked RWA*).

- The design of the course appears appropriate, the test subjects are neither under- nor overchallenged.
- A learning curve of the test subjects over the course of the test becomes clear.
- The research hypothesis was retained for the *FBA Selfsteer + Loss of Feedback failure.* The hypothesis cannot be tested for the *Blocked RWA failure,* as the specified sample size of valid data was not available for evaluation.
- The *FBA Selfsteer + Loss of Feedback* failure
    - was noticed by all test subjects;
    - is associated with a significant increase in the perceived mental and physical demands, the perceived own performance remains constant in the respective driving maneuver;
    - is mostly associated with shock and surprise on an affective level;
    - provokes a behavioral reaction in some of the test subjects (speed reduction/steering correction), while other test subjects stated that they continued driving unchanged;
    - is classified in terms of criticality at a high level of "noticeability" up to a high level of the category "disturbance of driving".

- The failure *Blocked RWA*
  - was not noticed by the majority of test subjects (80%);
  - does not cause a significant increase in the perceived mental and physical demands or a significant deterioration in the subjectively perceived performance in the driving maneuver in question;
  - did not lead to a specific affective or behavioral reaction in the test subjects who detected the failure;
  - is classified at a low to medium level of "noticeability" in terms of criticality if it was noticed by test subjects.

The discussion of the content of the results of the sub-studies for the entire test series is provided in Section 4.

3.4 Study 4

The fourth part of the study was carried out together with the compact class 3 vehicle in the period from April 20 to 25, 2022, using the associated test vehicle. The selected failure patterns were the *blocked FBA,* which was activated in the slalom maneuver, and the RWA failure *square-wave oscillation* in the associated straight exit maneuver. The results of this part of the study are presented in detail below.

3.4.1 Sample

The sample size is N = 28, 18 of whom are male. The mean age is M = 40.18 years (SD = 14.53 years), with the youngest subject being 22 years old and the oldest 64 years old. The average annual mileage is M = 15,089.29 km (SD = 11,120.60 km) with a range of 3,000 km to 60,000 km. One test subject reported a visual impairment in the form of a reduced field of vision, which, however, did not represent an obstacle. None of the other participants had any uncorrected visual or hearing impairments.

3.4.2 Results

The presentation of the results is divided into two sections: failure-independent and failure-specific results. Firstly, the results are reported, which relate to the general difficulty of the selected maneuvers in terms of testing the selected test design, independently of the failure setups. The failure-specific results are then presented. In this section, the results for testing the controllability of both failures (FBA and RWA) are discussed first. This is followed by the subjective test data regarding the experience of the failure activation and the objective test data regarding the driver and vehicle reaction as a result of the failure activation, first for the FBA failure and then for the RWA failure.

3.4.2.1 Failure-independent results

First, the perceived difficulty of the maneuvers was assessed based on the NASA-TLX instrument (Hart, 2006) using three items relating to the mental and physical demands, as well as the subjective assessment of one's own performance on a 20-point response scale (1 = very low to 20 = very high). The survey was conducted at two points in time, at the beginning of the test (after the familiarization ride) and before the end of the test (after ride 5). For inferential statistical analysis of the learning curve, the requirements at the beginning and end of the ride were compared using a paired *t-test*. All test subjects with valid data sets for both failure cases were included in the analysis.

At both survey times, the three maneuvers of driving straight-ahead, circular drive and slalom were classified in a low to medium range on the 20-point scale in terms of the mental demands involved in completing them. The maneuver driving straight-ahead appears to be associated with the comparatively lowest mental demands. For a presentation of the characteristic values, see Table 19. Compared to the beginning of the test, the maneuvers straight-ahead and slalom are perceived as mentally less demanding before the last run (after lap 5) (straight-ahead: $t(19) = 2.28$, $p = .018$, $d = 0.51$, slalom: $t(19) = 2.7$, $p < .01$, $d = 0.6$).

Table 19

Mental requirements of the maneuvers straight-ahead, circular drive and slalom based on a 20-point TLX scale (1 = very low to 20 = very high requirements) based on the NASA instrument (Hart, 2006).

Mental requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	20	6,05	4,01	1	13
	before the last round	20	4,35	2,80	1	10
Circular drive	after familiarization	20	10,15	4,16	2	15
	before the last round	20	8,95	3,97	2	15
Slalom	after familiarization	20	12,10	4,28	4	17
	before the last round	20	9,90	4,46	2	17

The physical requirements for completing the maneuvers are also rated in the low to medium range of the 20-point response scale over the course of the test. The maneuver driving straight-ahead again appears to be the least demanding maneuver. For a presentation of the characteristic values, see Table 20. The physical demands in the maneuvers driving straight-ahead and circular drive remain constant over the course of the test (driving straight-ahead: $t(19) = 0.0$, $p = .1.0$, circular drive: $t(19) = -0.05$, $p = .96$), whereas they decrease significantly for the slalom maneuver ($t(19) = 2.13$, $p = .023$, $d = 0.48$).

Table 20

Physical demands of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high demands) based on the NASA TLX instrument (Hart, 2006).

Physical requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	20	3,60	2,35	1	8
	before the last round	20	3,60	2,28	1	9
Circular drive	after familiarization	20	7,75	3,74	2	17
	before the last round	20	7,80	3,87	2	15
Slalom	after familiarization	20	10,95	4,38	3	17
	before the last round	20	9,40	4,65	3	18

The participants rated their own performance with regard to completing the three maneuvers at a high level on the 20-point response scale over the entire course of the test. For a presentation of the characteristic values, see Table 21. At the end of the test, performance in all maneuvers was rated significantly better than at the beginning of the test (straight driving: $t(19) = -3.20$, $p = .003$, $d = -0.72$, circular drive: $t(19) = -2.56$, $p < .01$, $d = -0.57$, slalom: $t(19) = -6.27$, $p < .001$, $d = -1.4$).

Table 21

Assessment of own performance in the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high performance) based on the NASA TLX instrument (Hart, 2006).

Assessment of own performance						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	20	15,45	3,66	7	20
	before the last round	20	17,40	1,93	13	20
Circular drive	after familiarization	20	13,45	3,49	8	20
	before the last round	20	15,70	2,32	10	19
Slalom	after familiarization	20	13,05	2,93	7	20
	before the last round	20	16,40	1,79	12	20

3.4.2.2 Failure-specific results

In this section, the results for testing the controllability of the two failures are discussed first. This is followed by a separate presentation of the subjective results for both failures with regard to the experience of the failure activation and then the objective test data with regard to the driver and vehicle reaction as a result of the failure events.

3.4.2.2.1 Controllability - hypothesis testing

In the context of the fourth sub-study, both for the *blocked FBA* failure activated in the slalom and for the *RWA square-wave oscillation* failure, which was activated when driving straight-ahead, 100% of the test subjects with valid data records did not leave the lane with the failure activated. 100 % of the data records comprise at least N = 20. The hypothesis can therefore be maintained for the combination of vehicle and failures selected in this part of the study. Within the scope of the study conducted, both failures were controllable for all test subjects at the C0 level defined by experts.

3.4.2.2.2 Failure type: *Blocked FBA*

In relation to the failure pattern of *blocked FBA,* the objective results for the description of the vehicle and driver reaction as a result of the failure are reported below, as well as the results with regard to the subjective perception of the connection.

3.4.2.2.3 Results of objective vehicle measurement data

The statistical distributions of the characteristic objective values for the *blocked FBA* failure pattern in the compact class 3 determined on the basis of the recorded vehicle measurement variables are shown in Illustration 66. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle of the failure pattern under investigation are also shown in Table 22. For a qualitative comparison of the objective parameters determined, the following Illustration 66 also shows the statistical evaluations for the second vehicle, which was examined with the same defect pattern as part of the overall study. However, due to vehicle and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.

In the present case, the Compact Class 3 compared to the SUV 2 shows significantly less Disturbance with regard to the *blocked FBA* failure pattern in relation to the driver's steering response and the resulting vehicle dynamics. It should be noted here that the failure activation time of 100 ms in the compact class 3 is only half as long as the failure activation time in the SUV 2. In addition, the failures were activated at different points on the slalom course.

Illustration 66

Statistical evaluation of the calculated objective parameters of Disturbance steering angle, Disturbance steering rate, Disturbance lateral acceleration and Disturbance yaw rate for the blocked FBA failure pattern
(gray: compact class 3; white: SUV 2)

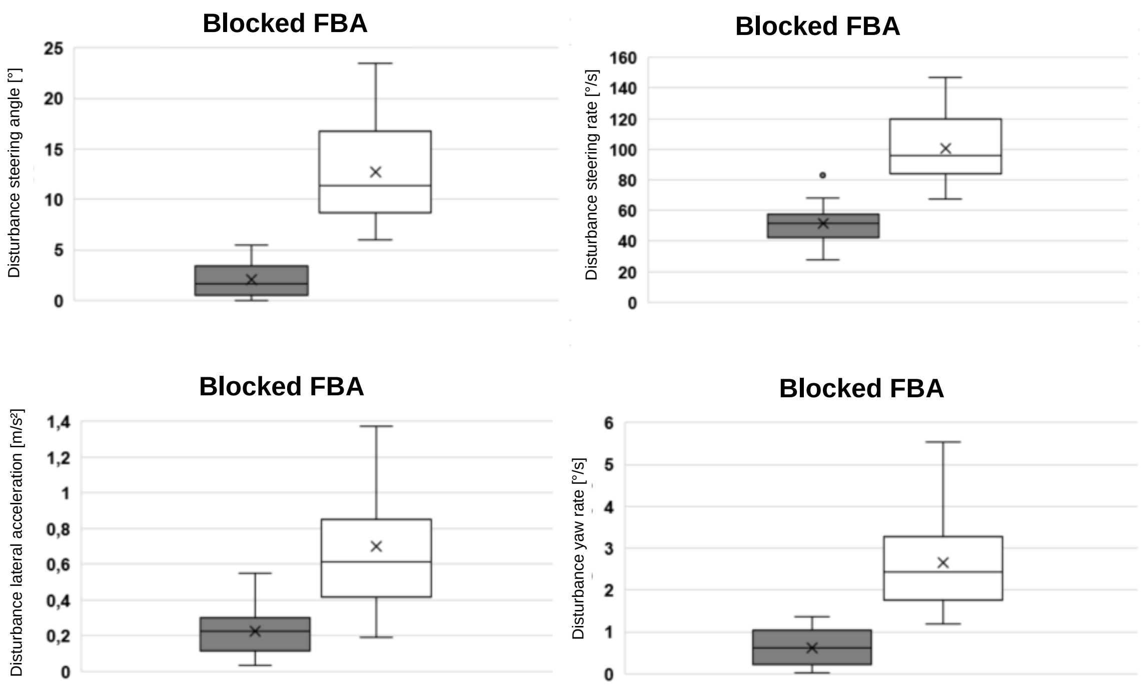


Table 22

Descriptive statistical parameters Failure pattern blocked FBA (compact class 3)

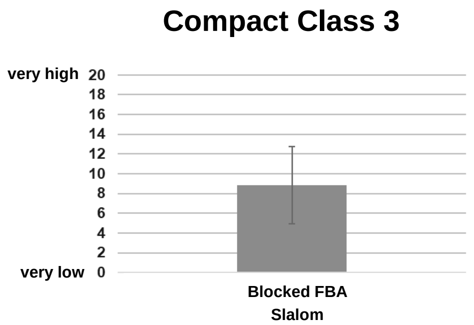
Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
Compact Class 3	Disturbance steering angle [°]	20	2,04	1,60	1,64	0,00	5,44
	Disturbance steering rate[°/s]	20	51,32	12,38	51,28	27,40	82,80
	Dist. lateral acceleration [m/s²]	20	0,23	0,13	0,23	0,03	0,55
	Disturbance yaw rate [°/s]	20	0,62	0,43	0,61	0,01	1,37

3.4.2.2.4 Results of subjective measures

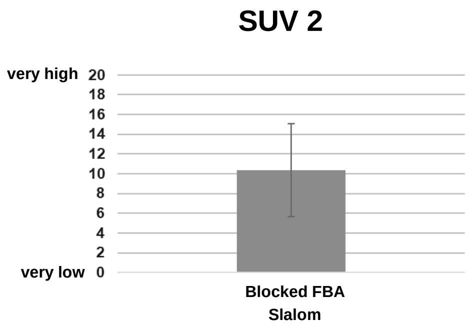
First of all, the perceived difficulty of the slalom maneuver with failure activation is discussed. This was recorded using three items based on the NASA-TLX instrument (Hart, 2006) on the perceived mental and physical demands as well as the subjective assessment of one's own performance on a 20-point response scale (1 = very low to 20 = very high). As the same type of failure was also used in sub-study 6 (for details see 3.6) with the SUV 2 vehicle using their test vehicle, the results of this sub-study with regard to the perceived difficulty of the maneuver in the event of a failure are already presented here as a basis for comparison. An explanation of potential differences in the results of both partial studies is not possible due to the parallel variation of test vehicle and failure parameterization. All valid data records were included in the analysis. When the *blocked FBA* failure was activated, the mental requirements for completing the slalom maneuver were classified in the middle range of the 20-point response scale (compact class 3: M = 8.85, SD = 3.91; SUV 2: M = 10.35, SD = 4.70). For a presentation of the results, see Illustration 67.

Illustration 67

Mental requirements of the slalom maneuver with activation of the blocked FBA failure were assessed on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
20	8,85	3,91	2	15



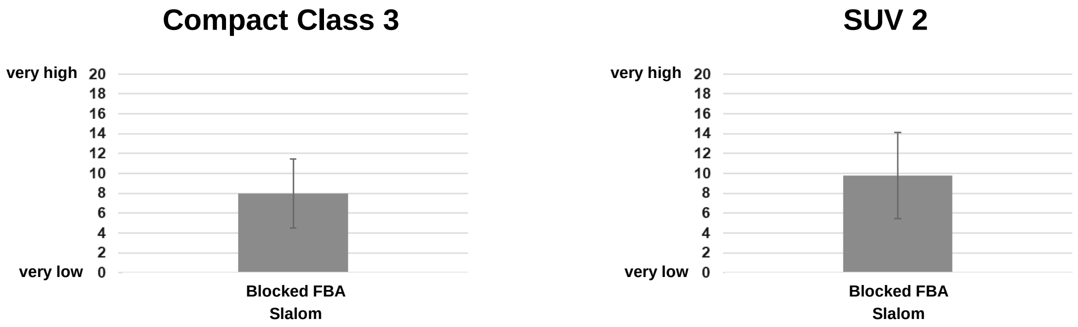
N	M	SD	Min	Max
26	10,35	4,70	3	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The physical requirements for completing the slalom maneuver with failure activation were classified in the middle range of the 20-point response scale in both sub-studies (compact class 3: *M* = 8.00, *SD* = 3.48, SUV 2: *M* = 9.77, *SD* = 4.32). For a presentation of the results, see Illustration 68.

Illustration 68

Physical demands of the slalom maneuver with activation of the blocked FBA failure were assessed on a 20-point response scale (1 = very low to 20 = very high demands) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
20	8,00	3,48	2	13

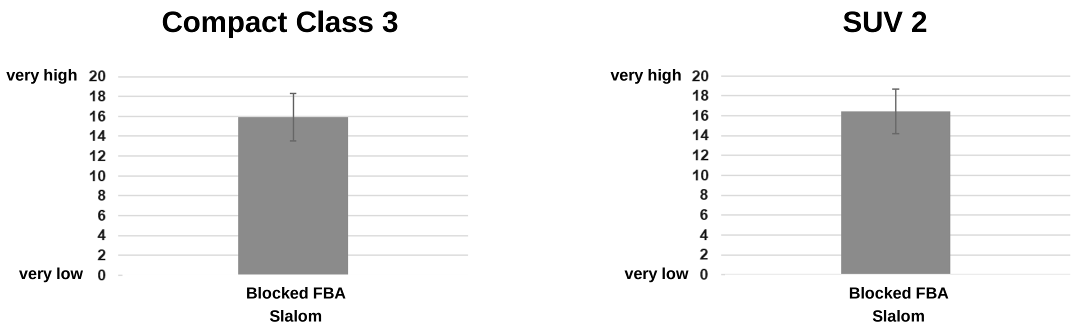
N	M	SD	Min	Max
26	9,77	4,32	3	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The subjective assessment of one's own performance when completing the slalom maneuver with failure activation was at a high level on the 20-point response scale in both sub-studies (compact class 3: *M* = 15.90, *SD* = 2.38; SUV 2: *M* = 16.46, *SD* = 2.23). For a presentation of the results, see Illustration 69.

Illustration 69

Subjective performance in completing the slalom maneuver with activation of the blocked FBA failure measured on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
20	15,90	2,38	12	20

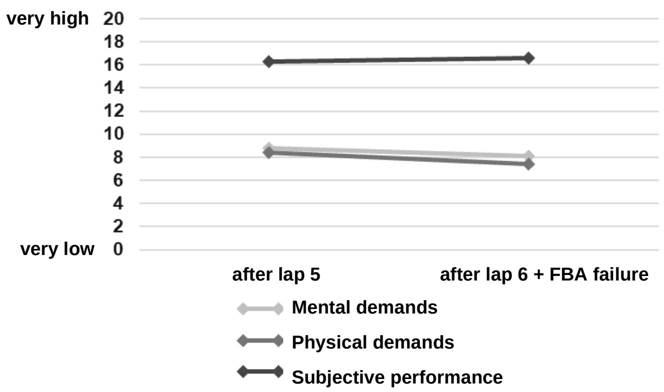
N	M	SD	Min	Max
26	16,46	2,23	11	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

In order to be able to describe the influence of the failure on the perceived difficulty of the maneuvers, the information on the perceived mental and physical demands as well as the subjective assessment of one's own performance after round 5 without failure activation were compared with those after failure activation in round 6. This approach was based on the assumption that the participants' learning curve with regard to completing the maneuvers was flattened at this point in the test and that differences between the assessments were attributable to the activation of the failure. Due to the randomized activation of both failures (RWA and FBA failures) in rounds 3 and 6, it should be noted that only about half of the valid data sets (in which the respective failure was activated in round 6) can be used for the analysis. The inferential statistical analysis of the comparison of both survey times was carried out using a paired *t-test*. In the sub-study with the compact class 3, the mental and physical demands remained constant in the last round with the activation of the *blocked FBA* failure compared to the previous round 5 (mental demands: t(9) = 1.21, *p* = .13; physical demands: *t*(9) = 1.63, *p* = .07). Similarly, the comparison of the subjective performance assessment shows no significant difference (*t(*9) = -0.67, *p* = .26). For a graphical representation of the progression and the associated characteristic values, see Illustration 70.

Illustration 70

*Subjectively perceived difficulty measured using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands and the subjectively perceived performance in the slalom maneuver in lap 5 without the failure activation compared to lap 6 after the failure blocked FBA was activated. (Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



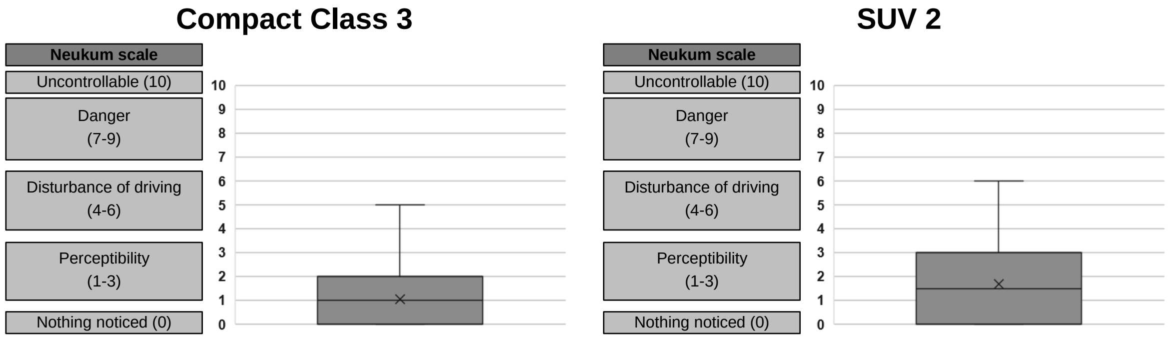
Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	**10**	8,80	4,73	2	16
	after lap 6 + failure	10	8,10	3,93	4	15
Physical demands	after lap 5	10	8,40	5,25	3	18
	after lap 6 + failure	10	7,40	4,01	2	13
Subjective performance	after lap 5	10	16,30	2,31	12	20
	after lap 6 + failure	10	16,60	2,50	12	20

When the failure was activated, the behavior of the test subjects was observed by the test administrator with the aim of recording the initial reaction of the participants. In this sub-study, no reaction noticeable to the experimenter was observed in any of the test subjects. When asked, 55% of the participants (n = 11) stated that they had noticed something special, and the majority described this as an intervention in the steering. The participants who reported noticing the failure were then asked about their affective, cognitive and behavioral reactions to the event. To record the affective reaction, they were asked to describe their feelings at the time of the event. Most subjects did not report any specific affective reaction (n = 8), with a few (n = 3) mentioning surprise. To assess the cognitive reaction, the test subjects were asked to describe the thoughts they had at the time of perceiving the event. No specific reaction was evident in this regard either. A few people (n = 3) stated that they had thought about the causes of the event. To record the behavioral reaction to the failure activation, the test subjects were asked what they did during the event. Most participants (n = 7) described subjectively continuing their drive unchanged. Individual test subjects reported corrective driving behavior: Steering wheel held tighter (n = 1), counter-steering (n = 1).

The subjectively perceived criticality of the event was assessed using the Neukum scale (Neukum & Krüger, 2003). Here too, the results of the sub-study with the compact class 3 were compared with those of sub-study 6 with SUV 2 (for details see 3.6), as the same failure type was used in both. However, potential differences in the data cannot be explained due to the simultaneous variation of failure parameterization and test vehicle. All valid data sets were considered for the analysis. The mean subjectively perceived criticality when experiencing the failure activation in the compact class 3 was classified on the 11-point scale (0 = nothing noticed - 10 = vehicle not controllable) at a low level in the noticeability category (M = 1.05, SD = 1.32). When experiencing the failure activation in the SUV 2, criticality was reported at a low to medium level of the noticeability category (M = 1.69, SD = 1.57). For a representation of both distributions see Illustration 71.

Illustration 71

Subjectively perceived criticality of the blocked FBA failure type in the two sub-studies with the compact class 3 and SUV 2 vehicles, measured using the Neukum scale (Neukum & Krüger, 2003).



N	M	SD	Median	Percentile		
				25	50	75
20	1,05	1,32	1,00	0,00	1,00	2,00

N	M	SD	Median	Percentile		
				25	50	75
26	1,69	1,57	1,50	0,00	1,50	3,00

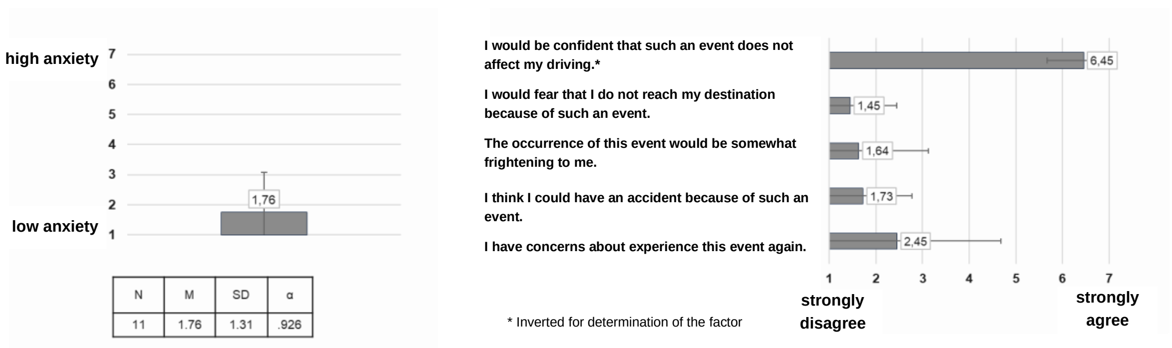
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The two rating scales of the factors *Anxiety* and *Perceived Safety* of the *Car Technology Acceptance Model* (Osswald et al., 2012) were selected and modified to survey the situational experience of the test subjects who perceived the failure event. This data was collected from all test subjects who stated that they had noticed the failure.

Anxiety with the occurrence of the failure event was assessed at a low level on the 7-point scale (1 = low anxiety - 7 = high anxiety) (*M* = 1.76, *SD* = 1.31). For a presentation of the results of the *Anxiety* factor and the individual items used for the calculation, see Illustration 72.

Illustration 72

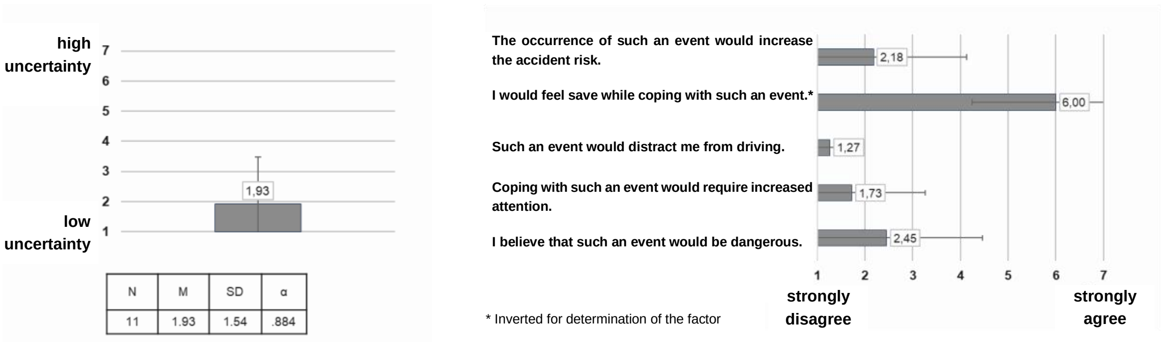
Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the Blocked FBA failure event.



The perceived uncertainty with the occurrence of the event was also assessed at a low level on the 7-point scale (1 = low uncertainty - 7 = high uncertainty) (M = 1.93, SD = 1.54). For a presentation of the results of the Perceived Safety factor and the individual items used for the calculation, see Illustration 73.

Illustration 73

Results of the factor Perceived Safety (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the failure event Blocked FBA.



3.4.2.2.5 Failure pattern: *RWA* square-wave *oscillation*

The objective results for describing the vehicle and driver reaction as a result of the activation of the *RWA square-wave oscillation* failure pattern in the straight-ahead maneuver are described below, as well as the results with regard to the subjective perception of the failure activation.

3.4.2.2.6 Results of objective vehicle measurement data

The statistical distributions of the characteristic objective values for the *RWA square-wave oscillation* failure pattern in the compact class 3 determined on the basis of the recorded vehicle measurement variables are shown in Illustration 74. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern investigated are also shown in Table 23. For a qualitative comparison of the objective parameters determined, the following Illustration 74 also shows the statistical evaluations for the second vehicle, which was examined with the same failure pattern as part of the overall study. However, due to vehicle and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.

In the present case, the compact class 3 shows greater disturbance influences with regard to the driver's steering response compared to the SUV 2 with regard to the *RWA square-wave oscillation* failure pattern. Both the disturbance influence on the steering angle and the disturbance steering rate are greater than the comparative values in the SUV 2. There are also higher disturbance influences with regard to lateral acceleration and yaw rate in terms of vehicle reaction.

Illustration 74

Statistical evaluation of the calculated objective parameters for disturbance influence steering angle, disturbance steering rate, disturbance lateral acceleration and disturbance yaw rate for the RWA square-wave oscillation *failure pattern*
(gray: compact class 3; white: SUV 2)

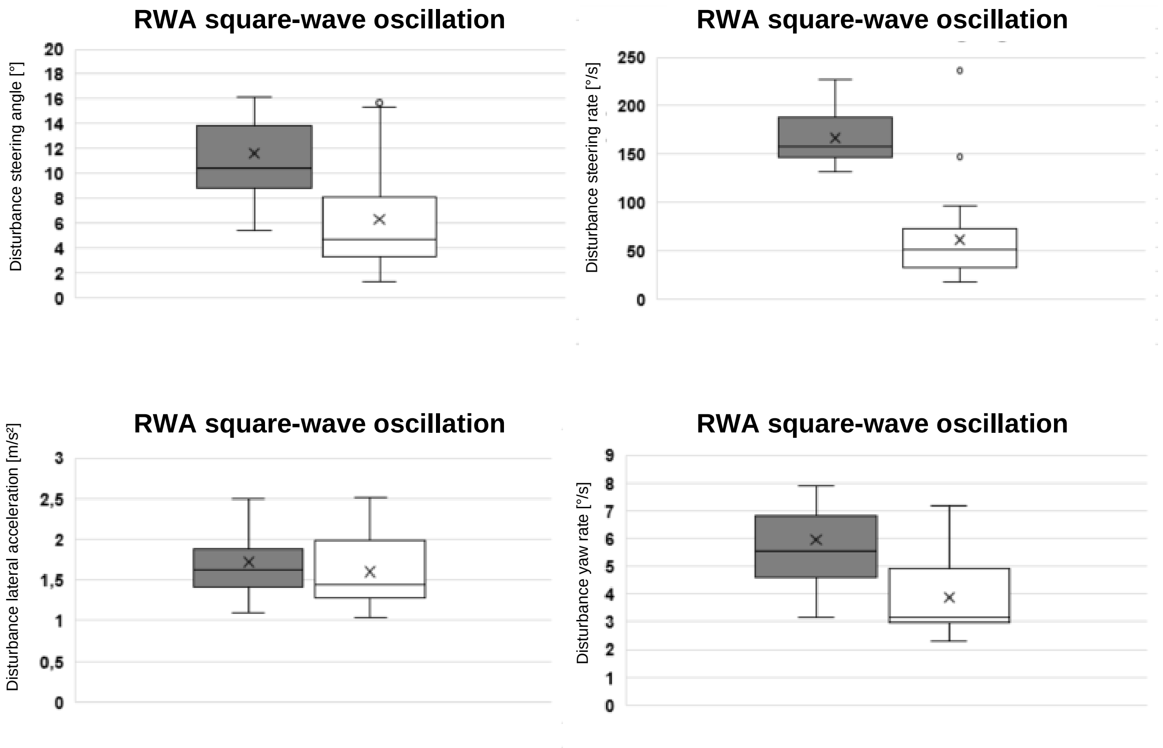


Table 23

Descriptive statistical parameters Failure pattern RWA square-wave oscillation (compact class 3)

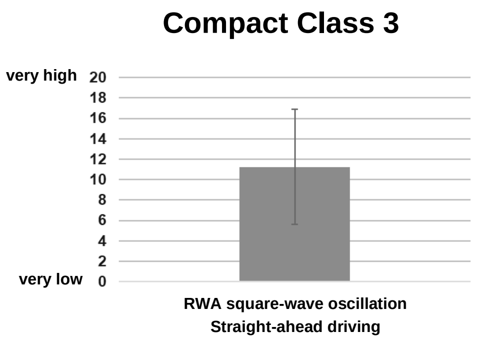
Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
Compact Class 3	Disturbance steering angle [°]	28	11,40	4,38	10,30	5,41	23,75
	Disturbance steering rate[°/s]	28	166,44	25,95	157,53	131,78	227,17
	Dist. lateral acceleration [m/s²]	28	1,72	0,50	1,63	1,10	3,73
	Disturbance yaw rate [°/s]	28	5,96	2,08	5,55	3,17	12,37

3.4.2.2.7 Results of subjective measures

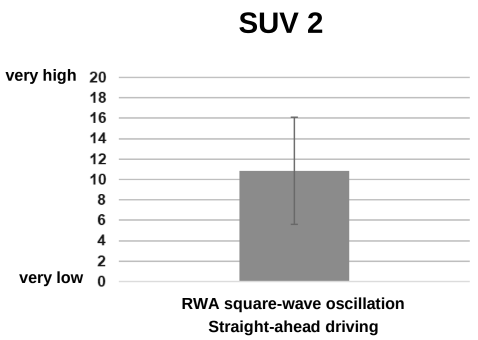
In the following, the perceived difficulty of the maneuver straight-ahead with activation of the failure is discussed first, which was assessed on a 20-point scale (1 = very low - 20 = very high) using three items based on the NASA-TLX instrument with regard to mental and physical demands, as well as the subjective assessment of one's own performance. The results are presented in comparison with the results of sub-study 6 (for details see 3.6) with the SUV 2 vehicle, as the same type of failure was considered in both sub-studies. However, an explanation of potential differences is not possible due to the simultaneous manipulation of failure parameterization and test vehicle. All valid data sets were included in the analysis. In both sub-studies, the mental requirements for driving straight-ahead with the RWA failure activated were assessed in the middle range of the 20-point response scale (1 = very low - 20 = very high) (compact class 3: *M* = 11.25, *SD* = 5.65; SUV 2: *M* = 10.83, *SD* = 5.26). For a presentation of the results, see Illustration 75.

Illustration 75

Mental requirements of the maneuver straight-ahead with activation of the failure RWA square-wave oscillation collected on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
28	11,25	5,65	1	20



N	M	SD	Min	Max
24	10,83	5,26	3	20

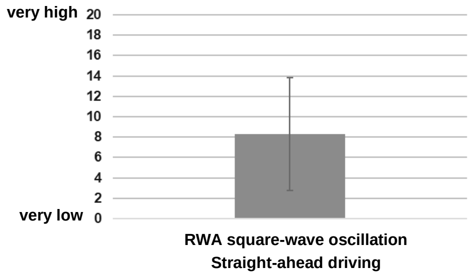
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

Likewise, the physical requirements for completing a straight-line exit with a failure event were assessed in both sub-studies in a medium range of the 20-point response scale (1 = very low - 20 = very high) (compact class 3: M = 8.29, SD = 5.52; SUV 2: M = 9.42, SD = 5.03). For a presentation of the results, see Illustration 76.

Illustration 76

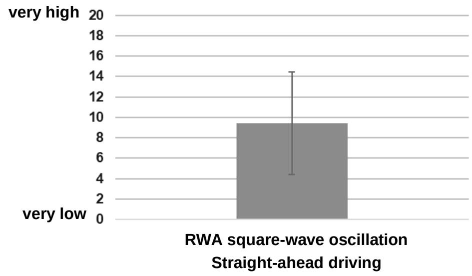
Physical requirements of the straight-ahead maneuver with activation of the RWA square-wave oscillation failure were determined on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).

Compact Class 3



N	M	SD	Min	Max
28	8,29	5,52	1	20

SUV 2



N	M	SD	Min	Max
24	9,42	5,03	2	20

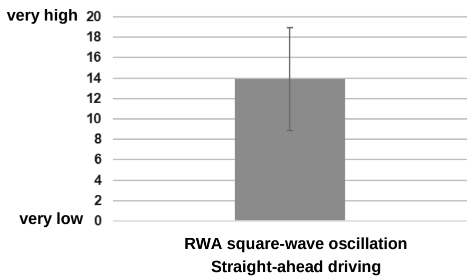
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The assessment of subjective performance when completing the maneuver of driving straight-ahead with failure activation is at a medium to high level on the 20-point response scale (1 = very low - 20 = very high) in both sub-studies (compact class 3: $M = 13.89$, $SD = 5.04$; SUV 2: $M = 15.33$, $SD = 3.66$). For a presentation of the results, see Illustration 77.

Illustration 77

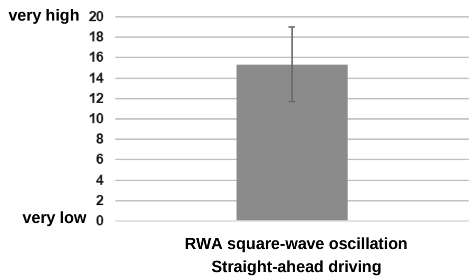
Subjective performance when completing the straight-ahead maneuver with activation of the RWA square-wave oscillation failure measured on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).

Compact Class 3



N	M	SD	Min	Max
28	13,89	5,04	1	20

SUV 2



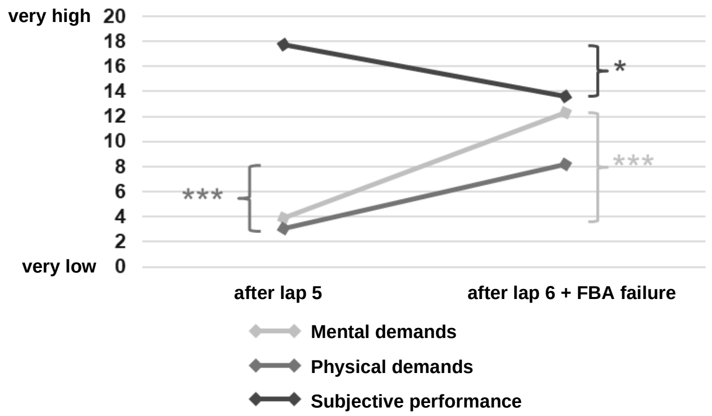
N	M	SD	Min	Max
24	15,33	3,66	5	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

With the aim of identifying the influence of the failure activation on the perceived difficulty of completing the maneuver straight-ahead, the mental and physical demands as well as the subjective assessment of one's own performance after lap 5 without failure activation were compared with those after activation in lap 6. The inferential statistical analysis of the comparison between the two survey times was carried out using a paired *t-test*. Both the mental and physical demands increased significantly in round 6 with failure activation compared to the previous round without failure activation (mental demands: $t(13) = -6.5$, $p < .001$, $d = -1.74$; physical demands: $t(13) = -4.18$, $p < .001$, $d = -1.12$). In addition, the subjectively perceived performance decreased significantly when completing the maneuver ($t(13) = 2.44$, $p = .02$, $d = -0.65$). For a graphical representation of the course and the associated characteristic values, see Illustration 78.

Illustration 78

*Subjective perceived difficulty measured using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands as well as the subjectively perceived performance in the straight-ahead maneuver in lap 5 without failure activation compared to lap 6 after activation of the RWA square-wave oscillation failure. (Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	14	3,93	2,43	1	8
	after lap 6 + failure	14	12,29	5,25	2	20
Physical demands	after lap 5	14	3,07	1,73	1	6
	after lap 6 + failure	14	8,21	5,31	1	20
Subjective performance	after lap 5	14	17,71	1,82	15	20
	after lap 6 + failure	14	13,57	6,44	1	20

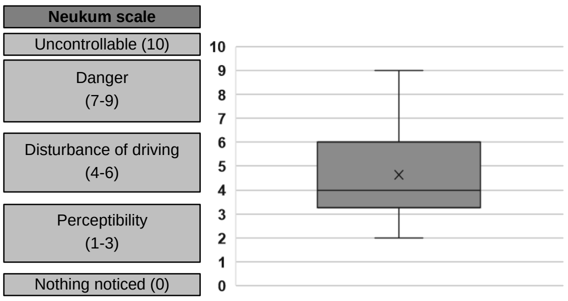
Observation of the test subjects by the experimenter at the time of the failure activation revealed that 60.71% ($n = 17$) of the participants showed a noticeable reaction, which in most cases ($n = 13$) took the form of a verbal response. When asked, 100% of participants ($n = 28$) stated that they had noticed the event. The majority described it as the vehicle swerving to the right and left ($n = 21$), or with an intervention in the steering / correction ($n = 12$). To describe the reaction to the event on an affective level, surprise/fright was mentioned most frequently (*n* = 18), followed by uncertainty (*n* = 6). On a cognitive level, the subjects most frequently reported having thought about having driven over a cone / something else (*n* = 11), or having thought about the cause of the event (*n* = 9). With regard to the behavioral reaction, the most frequently mentioned response was to have counter-steered / corrected (*n* = 16), followed by the statement to have held the steering wheel tighter (*n* = 7). If the event were to happen again, 17.86% of participants (*n* = 5) would react differently and reduce their speed (*n* = 3) or stop the vehicle (*n* = 2).

The results of the assessment of subjectively perceived criticality using the Neukum scale (Neukum & Krüger, 2003) are compared below for the sub-studies with the compact class 3 and SUV 2 vehicles (for details see 3.6), as the *RWA square-wave* failure type was taken into account in both. All valid data sets were taken into account for the analysis. In both sub-studies, criticality was classified on the 11-point scale (0 = nothing noticed - 10 = vehicle not controllable) at a low to medium level in the category of driving disturbance (compact class 3: *M* = 4.64, *SD* = 1.68; SUV 2: *M* = 4.38, *SD* = 1.50). For a representation of both distributions, see Illustration 79.

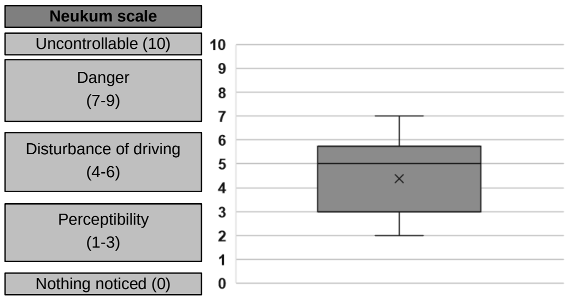
Illustration 79

Subjectively perceived criticality of the RWA square-wave oscillation failure type in the two sub-studies with the compact class 3 and SUV 2 vehicles, measured using the Neukum scale (Neukum & Krüger, 2003).

Compact Class 3



SUV 2



N	M	SD	Median	Percentile		
				25	50	75
28	4,64	1,68	4,00	3,25	4,00	6,00

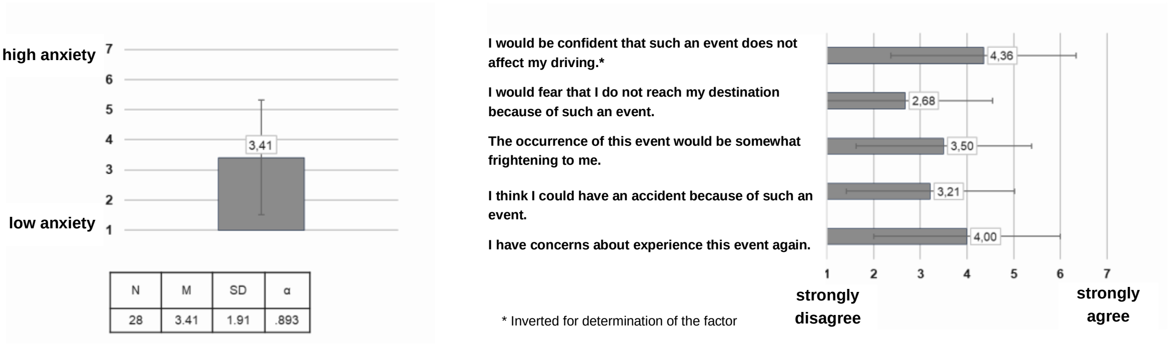
N	M	SD	Median	Percentile		
				25	50	75
24	4,38	1,50	5,00	3,00	5,00	5,75

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

Anxiety when noticing the failure event was classified on a medium range of the 7-point scale (1 = low anxiety - 7 = high anxiety) (*M* = 3.41, *SD* = 1.91). For a representation of the factor and the individual items used for the calculation, see Illustration 80.

Illustration 80

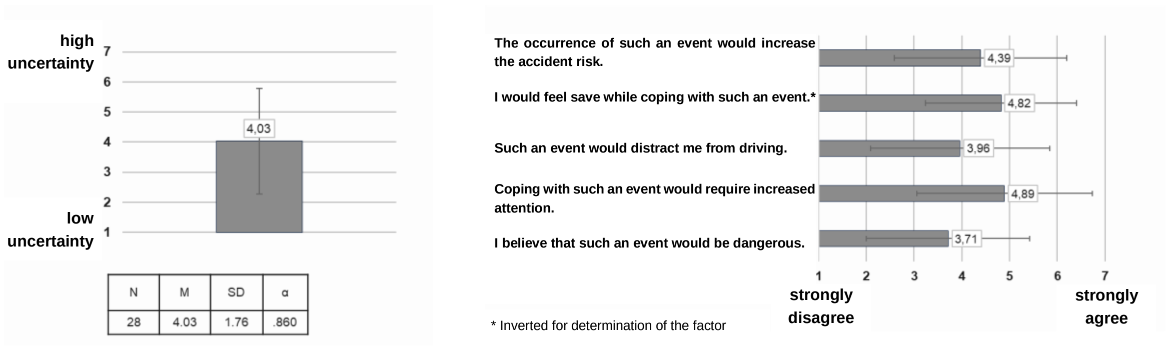
Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the RWA square-wave failure event.



The perceived uncertainty with noticing the failure event was reported on a medium level of the 7-point scale (1 = low uncertainty - 7 = high uncertainty) (*M* = 4.03, *SD* = 1.76). For a presentation of the factor and the individual items on which the calculation is based, see Illustration 81.

Illustration 81

Results of the Perceived Safety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the RWA square-wave failure event.

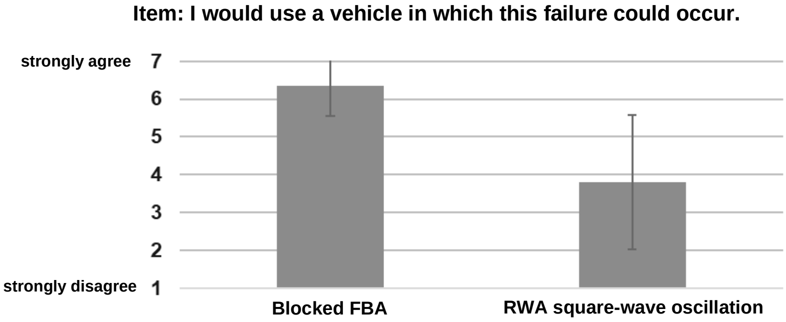


3.4.2.3 Follow-up survey

After the study was conducted, the participants were informed about the subject of the study. Based on this, a follow-up survey was conducted, whereby the participants only answered the questions if they had noticed the respective failure during the course of the experiment. First, the test subjects were asked to indicate on a 7-point response scale (1 = strongly disagree - 7 = strongly agree) how much they would agree to use a vehicle that might have the failure they had experienced. With regard to the *blocked FBA failure,* the intention to use the vehicle was high among those who noticed the failure during the test (*n* = 11) (*M* = 6.36, *SD* = 0.81). The *RWA square-wave* failure was noticed by all test subjects (N = 20). The willingness to use a vehicle that could have this failure was classified slightly lower in the middle range of the scale (*M* = 3.80, *SD* = 1.77). For a presentation of the results, see Illustration 82.

Illustration 82

Intention to use a vehicle that could exhibit the experienced failure of blocked FBA or RWA square-wave oscillation, surveyed using a 7-point Likert scale.



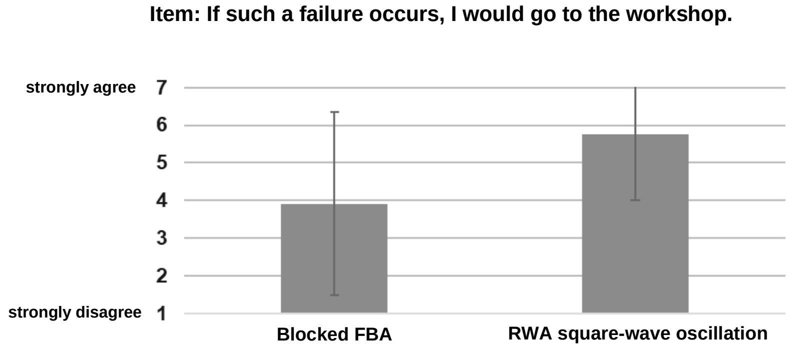
Failure	N	M	SD	Min	Max
Blocked **FBA**	11	6,36	0,81	5	7
RWA square-wave	20	3,80	1,77	1	6

Note. The item was only answered by test subjects who noticed the respective failure during the test run.

Furthermore, the test subjects were asked to indicate how much they agreed with the statement on a 7-point response scale (1 = strongly disagree - 7 = strongly agree) to visit the workshop when the respective failure occurs. The people who noticed the FBA failure (*n* = 11) neither agree nor disagree with this statement (*M* = 3.91, *SD* = 2.43). With regard to the occurrence of the RWA failure, the test subjects tended to agree with the statement (*M* = 5.75, *SD* = 1.74). For a presentation of the results, see Illustration 83.

Illustration 83

Intention to visit a workshop if the experienced failures Blocked FBA or RWA square-wave oscillation in own vehicle occur, assessed using a 7-point Likert scale.



Failure	N	M	SD	Min	Max
Blocked **FBA**	11	3,91	2,43	1	7
RWA square-wave	20	5,75	1,74	2	7

Note. The item was only answered by test subjects who noticed the respective failure during the test run.

3.4.3 Summary Study 4

The following is an overview of the main results of sub-study 4 (vehicle: compact class 3; failure: *blocked FBA* / *RWA square-wave oscillation*).

- The design of the course appears appropriate, the test subjects are neither under- nor overchallenged.
- A learning curve of the test subjects over the course of the test becomes clear.
- The research hypothesis was retained for both failure patterns.
- The *blocked FBA* failure
  - was noticed by 55% of the test subjects;
  - does not cause a significant increase in the perceived mental and physical demands or a significant deterioration in the subjectively perceived performance in the driving maneuver in question;
  - does not provoke any specific affective or behavioral response in the test subjects;
  - is classified at a low level of "noticeability" in terms of criticality.
- The *RWA square-wave* failure
  - was noticed by all test subjects, 60% of the test subjects showed a directly observable reaction to the failure activation;
  - is associated with a significant increase in the perceived mental and physical demands, and the perceived own performance also drops significantly with failure activation;
  - is mostly associated with shock and surprise on an affective level;

- provokes a behavioral reaction (counter-steering/correction) in the majority of test subjects; individual test subjects report that they would react by reducing their speed if they experienced the failure again;
- is classified in terms of criticality at a medium level of the category "disturbance of driving", ranging from a medium level of "noticeability" to a high level of the category "dangerousness".

The content of the results of the sub-studies is discussed for the entire test series in Section 4.

3.5 Study 5

The fifth part of the study was carried out together with the sedan vehicle in the period from April 26 to 29, 2022, using the associated test vehicle. The selected failure patterns were the *FBA step* failure, which was activated in the slalom maneuver, and the *uncontrolled RWA* failure in the circling maneuver. The results of this partial study are described in detail below.

3.5.1 Sample

The sample size is N = 27, of which 13 are female. The mean age is M = 39.70 (SD = 13.82 years), with the youngest subject being 20 and the oldest 62 years old. The average mileage is M = 17,253.85 km (SD = 13810.55 km) with a range of 2,500 km to 55,000 km. None of the test subjects stated that they had an uncorrected visual or hearing impairment.

3.5.2 Results

The presentation of the results is divided into two sections: failure-independent and failure-specific results. Firstly, the results are reported which relate to the general difficulty of the selected maneuvers in the sense of testing the selected test design, independently of the failure setups. The failure-specific results are then presented. In this section, the results for testing the controllability of both failures (FBA and RWA) are discussed first. This is followed by the subjective test data regarding the experience of the failure activation and the objective test data regarding the driver and vehicle reaction as a result of the failure activation, first for the FBA failure and then for the RWA failure.

3.5.2.1 Failure-independent results

The perceived difficulty of the maneuvers was assessed using three adapted questions from the NASA-TLX instrument (Hart, 2006) regarding the mental and physical demands, as well as the subjective assessment of one's own performance on a 20-point response scale (1 = very low to 20 = very high) at the beginning of the test after the familiarization ride and before the last lap following ride 5. For the inferential statistical analysis of the learning curve, the demands at the beginning and end of the ride were compared using a paired *t-test*. All test subjects with valid data sets for both failure cases were included in the analysis.

In terms of mental demands, all three maneuvers are classified in a low to medium range on the 20-point scale over the course of the test, with the straight-line maneuver appearing to be the least mentally demanding and the slalom the most mentally demanding. For a presentation of the characteristic values, see Table 24. Before the last drive, all maneuvers were perceived as significantly less mentally demanding than at the beginning of the test (straight driving: $t(20) = 2.60$, $p < .01$, $d = 0.57$; circular drive: $t(20) = 3.27$, $p = .002$, $d = 0.71$; slalom: $t(20) = 2.70$, $p = .007$, $d = 0.59$).

Table 24

Mental requirements of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high requirements) based on the NASA TLX instrument (Hart, 2006).

Mental requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	21	6,71	4,52	1	17
	before the last round	21	4,57	2,62	1	10
Circular drive	after familiarization	21	9,81	3,76	1	17
	before the last round	21	7,14	3,58	1	16
Slalom	after familiarization	21	12,29	2,83	5	17
	before the last round	21	9,62	3,88	3	17

In terms of physical demands, the three maneuvers were also classified in a low to medium range on the 20-point response scale at both survey times. Driving straight-ahead appears to be associated with the lowest physical demands and the slalom with the highest. For a presentation of the results, see Table 25. The physical demands of driving straight-ahead and circular drive remain constant when comparing the two survey times (driving straight-ahead: $t(20) = 0.42$, $p = .34$; circular drive: $t(20) = 1.38$, $p = .09$), whereas they decrease significantly for completing the slalom over the course of the test ($t(20) = 1.76$, $p = .047$, $d = 0.38$).

Table 25

Physical demands of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high demands) based on the NASA TLX instrument (Hart, 2006).

Physical requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	21	3,19	2,71	1	11
	before the last round	21	3,00	1,70	1	7
Circular drive	after familiarization	21	7,24	3,40	1	13
	before the last round	21	6,33	3,50	1	13
Slalom	after familiarization	21	10,10	4,22	1	16
	before the last round	21	8,67	4,56	1	15

Overall, the test subjects rated their own performance in completing the maneuvers at a high level on the 20-point response scale at both survey times. For a presentation of the results, see Table 26. At the end of the test, the subjects rated their own performance in all three maneuvers as significantly better than at the beginning of the test (driving straight-ahead: $t(20) = -2.79$, $p < .01$, $d = -0.61$; circular drive: $t(20) = -2.84$, $p < .01$, $d = -0.62$; slalom: $t(20) = -3.57$, $p = .001$, $d = -0.78$).

Table 26

Assessment of own performance in the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high performance) based on the NASA TLX instrument (Hart, 2006).

Assessment of own performance						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	21	16,62	3,28	10	20
	before the last round	21	18,00	1,92	13	20
Circular drive	after familiarization	21	15,14	3,21	9	20
	before the last round	21	16,90	2,47	11	20
Slalom	after familiarization	21	13,81	4,20	6	20
	before the last round	21	16,19	3,01	9	20

3.5.2.2 Failure-specific results

In the following passage, the results of the test of the controllability of the two failures *FBA step* and *uncontrolled RWA are* discussed first. This is followed by a separate presentation of the subjective results for both failures with regard to the experience of the failure activation and then the objective test data with regard to the driver and vehicle reaction as a result of the failure events.

3.5.2.2.1 Controllability - hypothesis testing

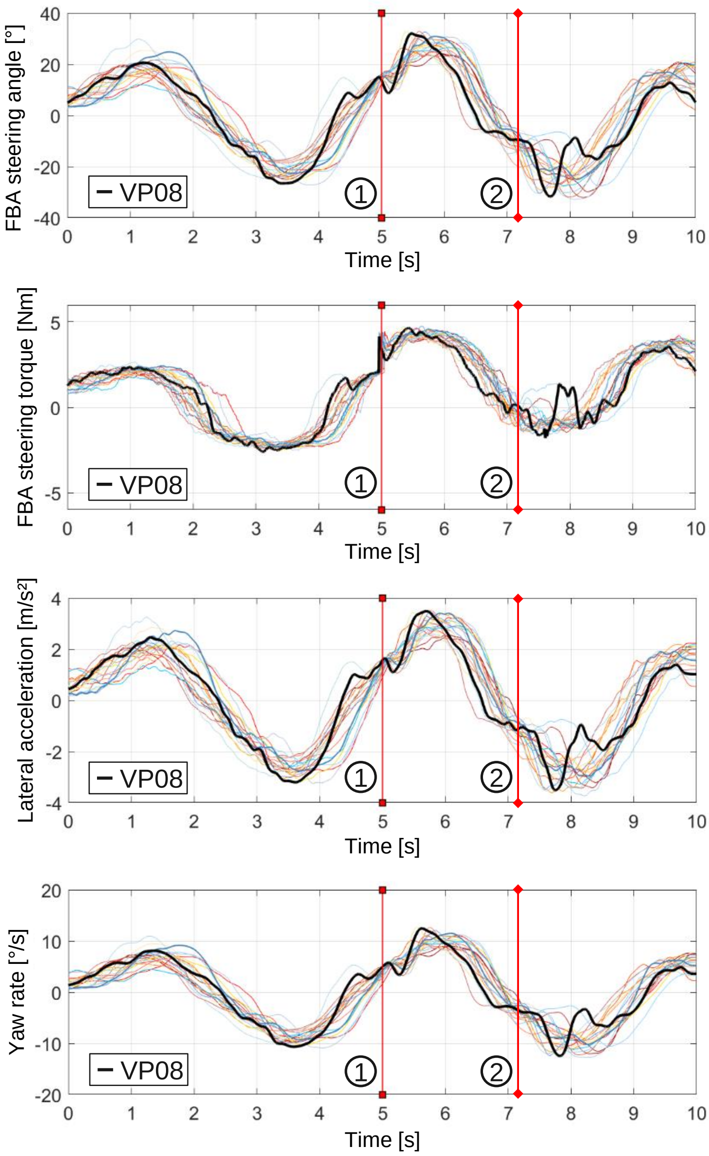
For the fifth sub-study, 100% of the test subjects with valid data records did not leave the lane with failure activation for the failure *Uncontrolled RWA* activated in the maneuver Circular drive. 100 % of the data records comprise at least *N* = 20 data records. The hypothesis can therefore be maintained for the selected combination of vehicle and RWA failure. Within the scope of the study, this failure was controllable for all test subjects at the C0 level defined by experts. For the *FBA step* failure activated in the slalom maneuver, one test subject left the lane with the failure activated. The hypothesis must therefore be rejected for the combination of vehicle and FBA failure. The C0 level defined by experts for this failure was not checked by all test subjects.

3.5.2.2.1.1 Lane departure analysis

The driving situation described can be analyzed in more detail using the available objective vehicle measurement variables. For this purpose, selected vehicle measurement variables from test subject 8 are compared with the other valid measurement data sets of the *FBA* step *failure activation* in Illustration 84. The time of the failure activation (1) and that of the pylon contact (2) are marked on the time axis. With regard to the driver inputs (steering angle, steering torque), test subject 8 does not show any conspicuous driving behavior compared to the other test subjects, both before the failure activation (1) and the contact with the pylon (2). This is also reflected in the resulting vehicle dynamics (lateral acceleration, yaw rate), which also showed no abnormalities before the pylon contact. As a result of the pylon contact (2), a steering reaction of the test subject is recognizable. This steering reaction also affects the vehicle dynamics, which deviate from the other measurement data sets in terms of yaw rate and lateral acceleration as a result of the pylon contact. For a graphical representation of the situation, see appendix 6.5.

Illustration 84

Vehicle measurement data FBA step failure activation (sedan) - VP08 vs. valid FBA step measurement data; (1) time of failure activation, (2) time of pylon contact



3.5.2.2.2 Failure type: *FBA step*

In relation to the *FBA step* failure pattern, the objective results describing the vehicle and driver reaction as a result of the failure and the subjective results regarding the subjective perception of the connection are reported below.

3.5.2.2.3 Results of objective vehicle measurement data

The statistical distribution of the characteristic objective values determined on the basis of the recorded vehicle measurement variables for the *FBA step* failure pattern are shown in Illustration 85 for the sedan. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern investigated are also shown in Table 27 summarized. For a qualitative comparison of the objective parameters determined, the following are shown in Illustration 85 also shows the statistical evaluations for the second vehicle, which was examined with the same failure pattern as part of the overall study. However, due to vehicle and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.

Illustration 85

Statistical evaluation of the calculated objective parameters steering angle, disturbance steering rate, disturbance lateral acceleration and disturbance yaw rate for the FBA step failure pattern
(white: sedan; gray: compact class 2)

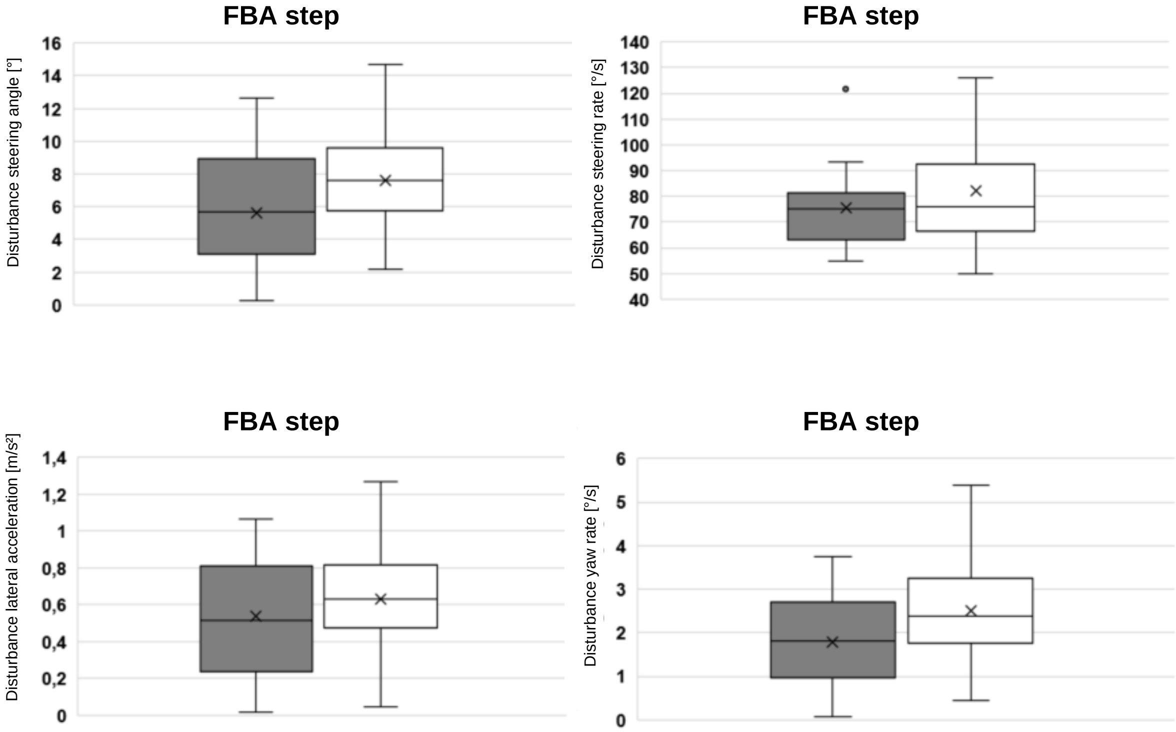


Table 27

Descriptive statistical characteristics Failure pattern FBA step (sedan)

Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
Sedan	**Disturbance steering angle [°]**	22	7,58	3,20	7,57	2,16	14,72
	Disturbance steering rate[°/s]	22	82,11	22,48	75,88	49,68	146,50
	Dist. lateral acceleration [m/s²]	22	0,63	0,30	0,63	0,05	1,27
	Disturbance yaw rate [°/s]	22	2,51	1,21	2,39	0,45	5,39

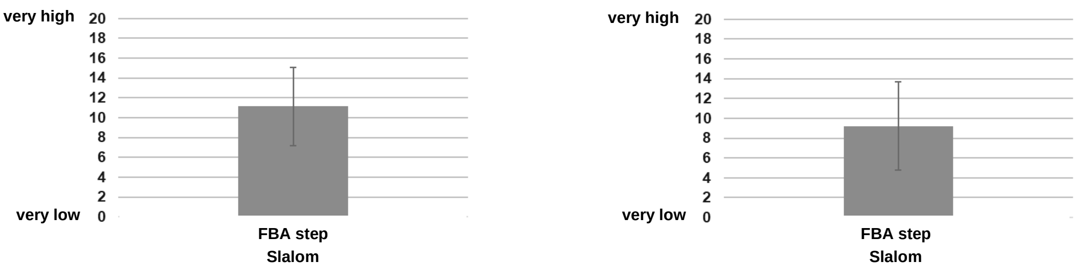
3.5.2.2.4 Results of subjective measures

In this section, the perceived difficulty of the maneuvers with failure activation is reported first. The survey was also conducted using three questions based on the NASA-TLX instrument (Hart, 2006) regarding the mental and physical demands as well as the assessment of one's own performance when completing the maneuvers using a 20-point response scale (1 = very low to 20 = very high). As the *FBA step* failure type was also the subject of sub-study 2 (for details see 3.2) with the compact class 2, the results of both sub-studies are presented in comparison. However, potential differences between the two sub-studies can only be presented and not explained due to the simultaneous variation of failure parameterization and test vehicle. All valid data sets were included in the analysis. In the sub-study with the sedan vehicle, the mental requirements of the slalom maneuver with the FBA failure were classified in the middle range of the 20-point scale. (*M* = 11.14, *SD* = 3.96). In the sub-study with the compact class 2, the mental demands in this situation were also estimated to be in the medium range (*M* = 9.23, *SD* = 4.46). For a presentation of the results, see Illustration 86.

Illustration 86

Mental requirements of the slalom maneuver with activation of the FBA step failure assessed on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).

Sedan **Compact Class 2**



N	M	SD	Min	Max
22	11,14	3,96	4	17

N	M	SD	Min	Max
26	9,23	4,46	2	18

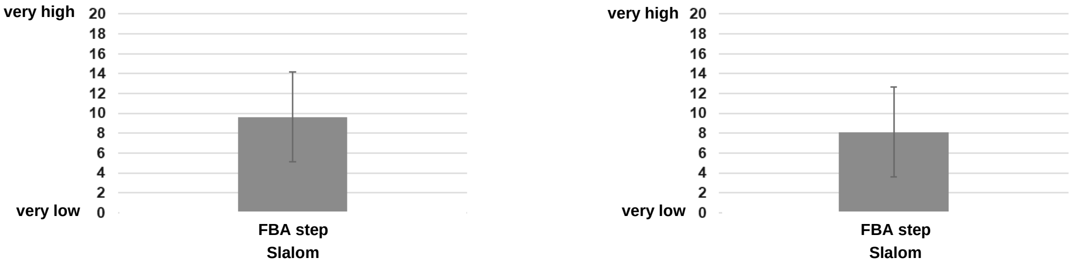
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The physical requirements for completing the slalom maneuver with failure activation are classified in both sub-studies at a medium level on the 20-point scale (sedan: *M* = 9.64, *SD* = 4.50; compact class 2: *M* = 8.12, *SD* = 4.50). For a presentation of the results, see Illustration 87.

Illustration 87

Physical demands of the slalom maneuver with activation of the FBA step failure assessed on a 20-point response scale (1 = very low to 20 = very high demands) based on the NASA-TLX instrument (Hart, 2006).

Sedan **Compact Class 2**



N	M	SD	Min	Max
22	9,64	4,50	1	17

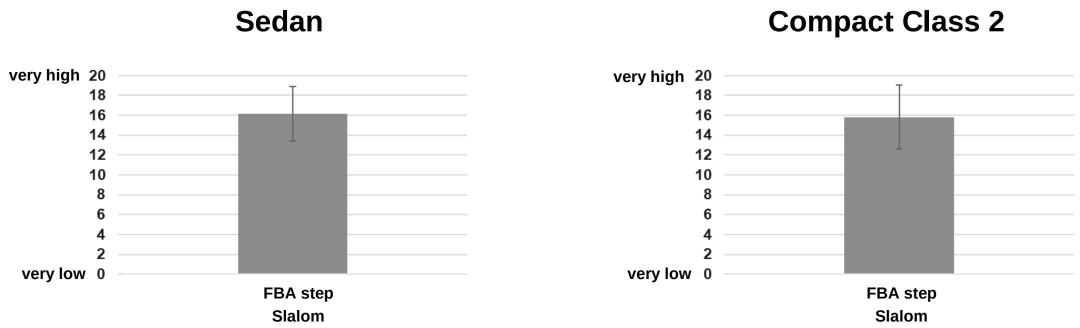
N	M	SD	Min	Max
26	8,12	4,50	1	17

In both sub-studies, the participants rated their own performance in completing the slalom maneuver with failure activation in a medium to high range on the 20-point scale (sedan: *M* =

16.14, *SD* = 2.73; compact class 2: *M* = 15.81, *SD* = 3.24). For a presentation of the results, see Illustration 88.

Illustration 88

Subjective performance in completing the slalom maneuver with activation of the FBA step failure on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).



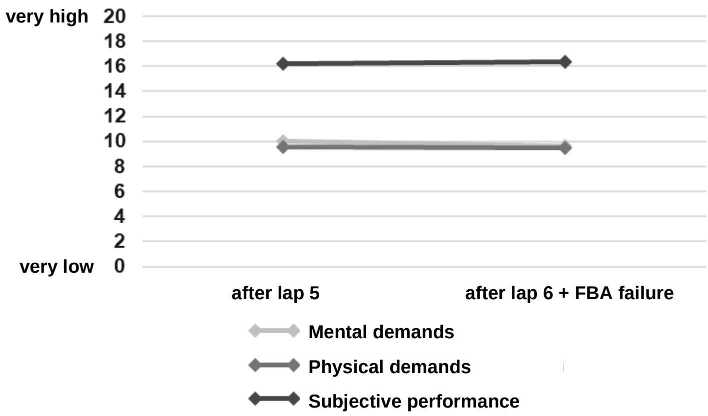
N	M	SD	Min	Max
22	16,14	2,73	9	20

N	M	SD	Min	Max
26	15,81	3,24	10	20

To identify the proportion of the difficulty of the slalom maneuver attributable to the failure activation, the perceived mental and physical demands as well as the subjective performance assessment after round 5 without failure activation were compared with those after failure activation in round 6. This approach was based on the assumption that the learning curve of the test subjects with regard to completing the maneuver was already flattened at this point in the test and that differences between the time points can be attributed to the experience of the failure. Due to the randomized activation of FBA and RWA failures in rounds 3 and 6, only about half of the valid data sets (in which the activation took place in round 6) can be considered in this analysis. The inferential statistical analysis of the comparison of the two survey times was carried out using a paired *t-test*. It became clear that both the mental and physical requirements remained constant with the activation of the *FBA step failure* compared to the previous round (mental requirements: $t(12) = 0.67$, $p = .26$; physical requirements: $t(12) = 0.11$, $p = .46$). Similarly, there was no significant difference in the comparison of the subjective performance assessment when completing the maneuver ($t(12) = -0.34$, $p = .37$). For a graphical representation of the course and the associated characteristic values, see Illustration 89.

Illustration 89

*Subjectively perceived difficulty measured using three items based on the NASA-TLX instrument (Hart, 2016) in relation to the mental and physical demands as well as the subjectively perceived performance in the slalom maneuver in round 5 without failure activation compared to round 6 after activation of the FBA step failure. (Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	13	10,00	3,16	5	15
	after lap 6 + failure	13	9,62	4,07	4	16
Physical demands	after lap 5	13	9,54	4,35	1	15
	after lap 6 + failure	13	9,46	3,99	1	14
Subjective performance	after lap 5	13	16,23	3,03	11	20
	after lap 6 + failure	13	16,38	2,60	12	20

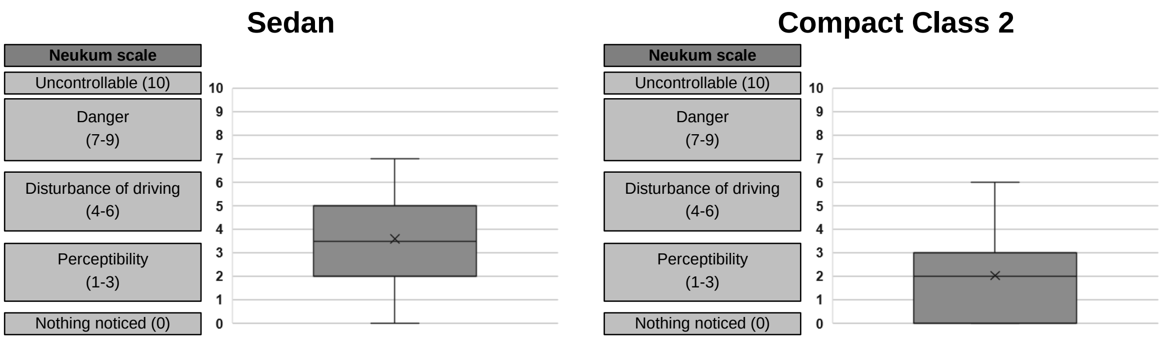
At the time of the failure activation, the behavior of the test subjects was observed by the test administrator with the aim of recording the initial reaction. 18.18% of the participants (*n* = 4) showed a reaction that was noticeable to the experimenter, which in *n* = 3 cases took the form of a verbal expression. When asked, 95% of participants (*n* = 21) stated that they had noticed something special. This was most frequently described as a blockage at the steering wheel (*n* = 17), followed by the observation of the vehicle steering in the opposite direction (*n* = 8). To record the participants' reaction to the experience of the failure, the people who noticed the failure (*n* = 21) were asked about their affective, cognitive and behavioral reaction. With regard to the affective reaction, they were asked to describe their feelings on noticing the failure. Most subjects did not report a specific reaction, with surprise/fright being mentioned less frequently (*n* = 8). With regard to the cognitive reaction, participants were asked what they were thinking at the time of the event. The most common answer (*n* = 5) was that something was different /

not right. With regard to the behavioral reaction, the majority of test subjects (n = 19) stated that they had counter-guided / corrected.

The subjective perception of the criticality of the failure recording was assessed using the Neukum scale (Neukum & Krüger, 2003). Here, too, the results are compared with those obtained in sub-study 2 with the compact class 2 (for details see 3.2) and taking into account the same type of failure. Again, potential differences in the results of both sub-studies due to the parallel manipulation of failure parameterization and test vehicle can only be illustrated and not explained. All valid data sets were taken into account for the analysis. In the sub-study with sedan, the subjectively perceived criticality of the *FBA step* failure was classified on average between a high level in the noticeability category and a low level in the driving disturbance category on the 11-point response scale (0 = nothing noticed - 10 = vehicle not controllable) (M = 3.59, SD = 1.65). In the sub-study with the compact class 2, criticality was reported on average in a medium range of the noticeability category (M = 2.04, SD = 1.71). For a representation of both distributions see Illustration 90.

Illustration 90

Subjectively perceived criticality of the failure type FBA step in the two sub-studies with the sedan and compact class 2 vehicles, measured using the Neukum scale (Neukum & Krüger, 2003).



N	M	SD	Median	Percentile		
				25	50	75
22	3,59	1,65	3,50	2,00	3,50	5,00

N	M	SD	Median	Percentile		
				25	50	75
26	2,04	1,71	2,00	0,00	2,00	3,00

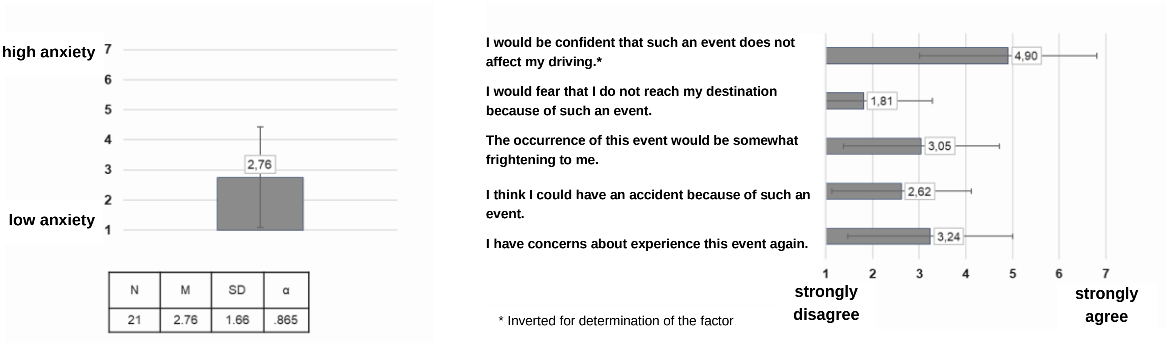
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The two factors *Anxiety* and *Perceived Safety* of the *Car Technology Acceptance Model* (Osswald et al., 2012) were chosen to assess the experience of the situation. This data was collected from all test subjects who stated that they had noticed the failure.

Anxiety when the failure event occurs is classified on a lower range of the 7-point scale (1 = low anxiety to 7 = high anxiety) (M = 2.76, SD = 1.66). For an illustration of the *anxiety* factor and the individual items used in the calculation, see Illustration 91.

Illustration 91

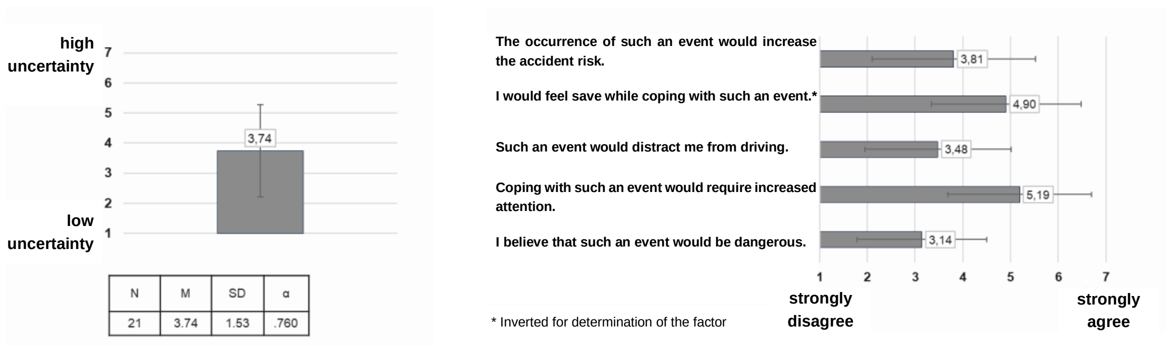
Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the FBA step failure event.



The perceived uncertainty with the occurrence of the failure was classified on a medium level of the 7-point scale (1 = low uncertainty - 7 = high uncertainty) (*M* = 3.74, *SD* = 1.53). For a presentation of the results of the *Perceived Safety* factor and the underlying individual items, see Illustration 92.

Illustration 92

Results of the Perceived Safety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the FBA step failure event.



3.5.2.2.5 Failure pattern *Uncontrolled* RWA

The following section reports the objective results for the description of the vehicle and driver reaction as a result of the activation of the failure pattern *Uncontrolled RWA* in the maneuver circular drive, as well as the results regarding the subjective perception of the activation.

3.5.2.2.6 Results of objective vehicle measurement data

The statistical distributions of the objective characteristics determined on the basis of the recorded vehicle measured variables for the failure pattern of *uncontrolled RWA* in the sedan are shown in Illustration 93. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern investigated are also shown in Table 28. For a qualitative comparison of the objective parameters determined, the following Illustration 93 also shows the statistical evaluations for the second vehicle, which was examined with the same failure pattern as part of the overall study. However, due to vehicle and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.

In the present case, the sedan tends to show slightly greater Disturbance with the driver's steering response and the resulting vehicle dynamics with regard to the *Uncontrolled RWA* failure pattern.

Illustration 93

Statistical evaluation of the calculated objective parameters of Disturbance steering angle, Disturbance steering rate, Disturbance lateral acceleration and Disturbance yaw rate for the uncontrolled RWA failure pattern
(white: sedan; gray: compact class 2)

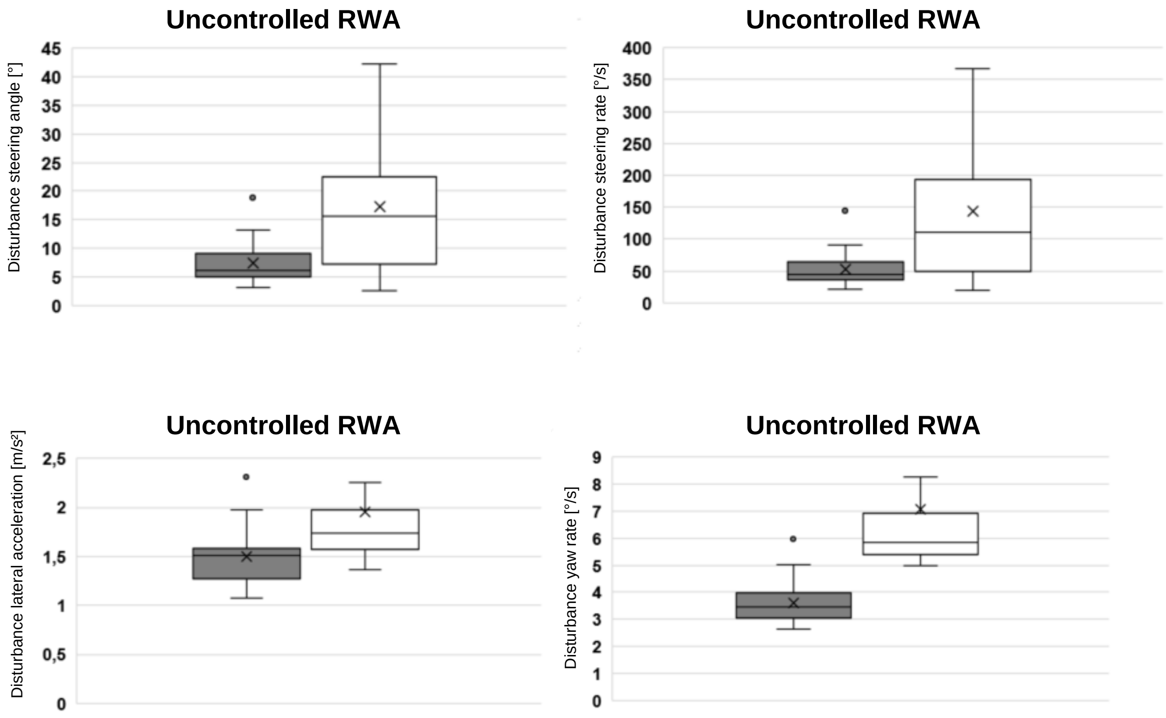


Table 28

Descriptive statistical characteristics Failure pattern Uncontrolled RWA (sedan)

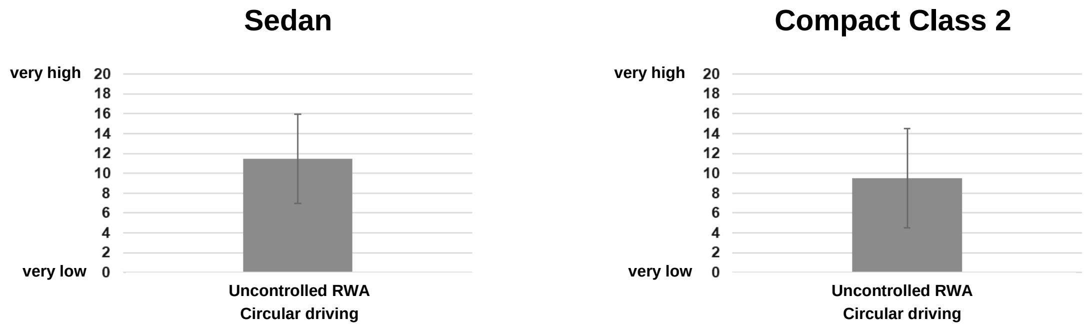
Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
Sedan	**Disturbance steering angle [°]**	25	16,13	12,36	13,19	2,64	51,36
	Disturbance steering rate[°/s]	25	144,08	111,60	111,15	20,28	451,98
	Dist. lateral acceleration [m/s²]	25	1,95	0,66	1,74	1,37	4,24
	Disturbance yaw rate [°/s]	25	7,06	2,93	5,83	5,00	16,63

3.5.2.2.7 Results of subjective measures

In the following, the perceived difficulty of completing the maneuver in a circular drive with the failure being activated is discussed. This was recorded using three items based on the NASA-TLX instrument (Hart, 2006) with regard to mental and physical demands, as well as the assessment of one's own performance on a 20-point response scale (1 = very low to 20 = very high). Again, the results are presented here together with the results of sub-study 2 (for details see 3.2) with the compact class 2, as the same type of failure was taken into account in this study. An explanation of potential differences between the sub-studies is not possible due to the simultaneous variation of vehicle and failure parameterization. All valid data sets were included in the analysis. In both sub-studies, the mental demands of experiencing the RWA failure were estimated in the middle range of the 20-point response scale (sedan: *M* = 11.48, *SD* = 4.47; compact class 2: *M* = 9.50, *SD* = 4.97). For a presentation of the results, see Illustration 94.

Illustration 94

Mental requirements of maneuvering in a circular drive with activation of the failure Uncontrolled RWA collected on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
25	11,48	4,47	3	20

N	M	SD	Min	Max
22	9,50	4,97	3	19

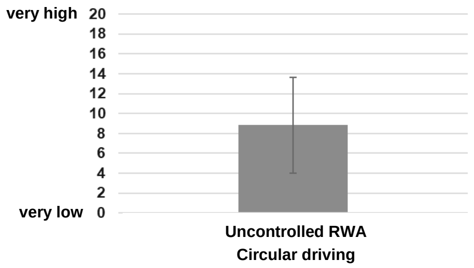
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The physical requirements for completing the maneuver of driving in a circular drive with activation of the RWA failure were assessed in the low to medium range of the 20-point scale in both sub-studies (sedan: *M* = 8.84, *SD* = 4.82; compact class 2: *M* = 8.00, *SD* = 4.91). For a presentation of the results, see Illustration 95.

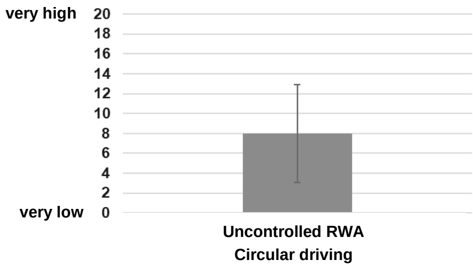
Illustration 95

Physical requirements of the maneuver Circular motion with activation of the failure Uncontrolled RWA collected on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).

Sedan



Compact Class 2



N	M	SD	Min	Max
25	8,84	4,82	1	19

N	M	SD	Min	Max
22	8,00	4,91	1	17

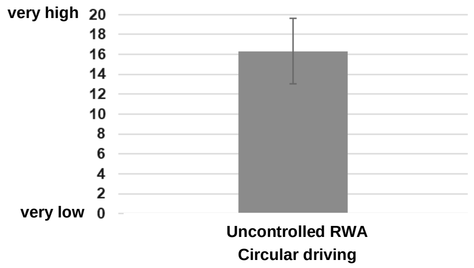
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

In both sub-studies, the assessment of the participants' own performance in completing the maneuver of circular drive with failure activation was in the medium to high range of the 20-point response scale (sedan: *M* = 16.32, *SD* = 3.28; compact class 2: *M* = 16.00, *SD* = 2.76). For a presentation of the results, see Illustration 96.

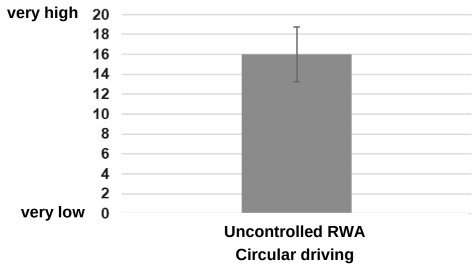
Illustration 96

Subjective performance in completing the maneuver of circular drive with activation of the failure Uncontrolled RWA measured on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).

Sedan



Compact Class 2



N	M	SD	Min	Max
25	16,32	3,28	8	20

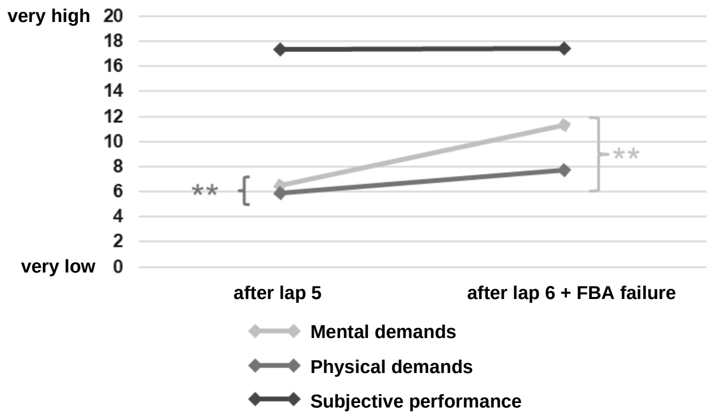
N	M	SD	Min	Max
22	16,00	2,76	11	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

To identify the proportion of the perceived difficulty of the circular drive maneuver that can be attributed to the experience of the *Uncontrolled RWA* failure, a comparison was made of the perceived mental and physical demands as well as the subjective performance assessment after round 5 without failure activation with those after failure activation in round 6. The inferential statistical analysis of the comparison of both survey times was carried out using a paired *t-test*. It became clear that both the mental and physical demands in the last round with failure activation increased significantly compared to the previous round (mental demands: $t(12) = -3.91$, $p = .001$, $d = -1.08$, physical demands: $t(12) = -4.01$, $p = .001$, $d = -1.12$). At the same time, the assessment of one's own performance remains constant ($t(12) = -0.17$, $p = .44$). For a graphical representation of the progression and the associated characteristic values, see Illustration 97.

Illustration 97

*Subjectively perceived difficulty measured using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands as well as the subjectively perceived performance in the maneuver circular drive in round 5 without failure activation compared to round 6 after activation of the failure Uncontrolled RWA. (Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



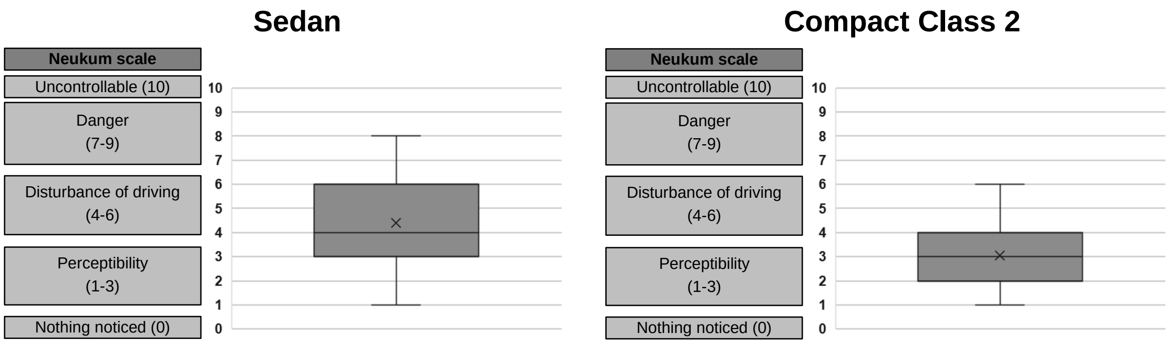
Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	13	6,46	3,71	2	16
	after lap 6 + failure	13	11,31	5,11	3	18
Physical demands	after lap 5	13	5,92	3,69	1	13
	after lap 6 + failure	13	7,69	4,73	1	17
Subjective performance	after lap 5	13	17,38	1,71	13	19
	after lap 6 + failure	13	17,46	2,03	14	20

The behavioral observation of the test subjects at the time of the failure activation revealed that 80% of the participants ($n = 20$) showed a reaction that was noticeable to the test supervisor. This was shown in most cases ($n = 14$) by verbal statements and second most frequently ($n = 7$) by a glance in the side/rear view mirror. When asked, 100% of participants ($n = 25$) stated that they had noticed something special and most frequently ($n = 12$) described this as the vehicle swerving/skidding. The most frequently mentioned description of the affective reaction was surprise/fright (*n* = 18). To describe the reaction on a cognitive level, the thought of having driven over a pylon/something else was mentioned most frequently (*n* = 14). On a behavioral level, some test subjects (*n* = 15) stated that they had counter-steered/corrected after the failure had occurred. If the event were to happen again, 80% of the participants (*n* = 20) would react in the same way, whereas some people would stop and look at the reason for the event (*n* = 3) or slow down more quickly (*n* = 2).

The evaluation of the subjectively perceived criticality using the Neukum scale (Neukum & Krüger, 2003) showed that the failure *of uncontrolled RWA* in the sub-study with the sedan on the 11-point scale (0 = nothing noticed - 10 = vehicle not controllable) was on average at a low level in the category disturbance of driving (*M* = 4.40, *SD* = 1.66). All valid data sets were included in the analysis. For comparison, the results of the sub-study with the compact class 2 (for details see 3.2), in which the same type of failure was investigated. However, it is not possible to explain differences between the results of the two sub-studies due to the simultaneous variation of failure parameterization and test vehicle. In the compact class 2 sub-study, the same failure type was reported on average at a high level in the noticeability category (*M* = 3.05, *SD* = 1.29). For a representation of both distributions, see Illustration 98.

Illustration 98

Subjectively perceived criticality of the failure type Uncontrolled SHE in the two sub-studies with the sedan and compact class 2 vehicles, measured using the Neukum scale (Neukum & Krüger, 2003).



N	M	SD	Median	Percentile		
				25	50	75
25	4,40	1,66	4,00	3,00	4,00	6,00

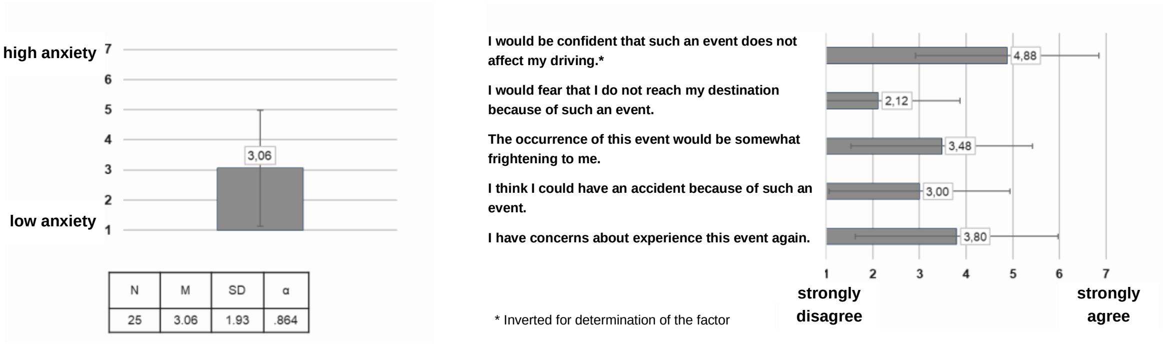
N	M	SD	Median	Percentile		
				25	50	75
22	3,05	1,29	3,00	2,00	3,00	4,00

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

Anxiety with noticing the event appears at a medium level on the 7-point scale (1 = low anxiety to 7 = high anxiety) (*M* = 3.06, *SD* = 1.93). For an illustration of the factor and the underlying items, see Illustration 99.

Illustration 99

Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the Uncontrolled RWA failure event.

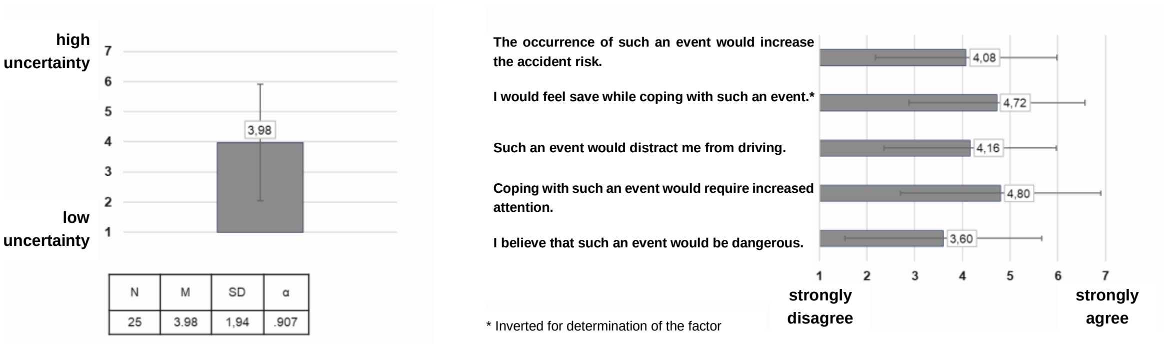


N	M	SD	α
25	3.06	1.93	.864

The perceived uncertainty when noticing the failure was in the middle of the 7-point scale (1 = low uncertainty to 7 = high uncertainty) (*M* = 3.98, *SD* = 1.94). For a presentation of the results of the factor and the individual items, see Illustration 100.

Illustration 100

Results of the Perceived Safety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the failure event Uncontrolled RWA.



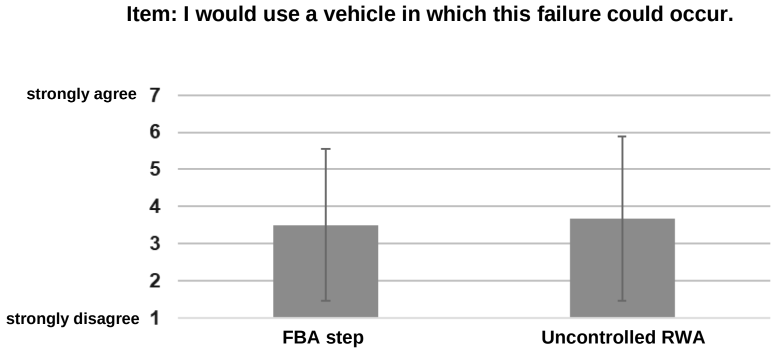
N	M	SD	α
25	3.98	1,94	.907

3.5.2.2.8 Follow-up survey

After informing the participants about the test object following the test drives, a follow-up survey was conducted. For each defect noticed, the participants were first asked to indicate to what extent they would agree to use a vehicle that might have this defect. The FBA failure was noticed by all but one participant (*n* = 20). On a 7-point scale (1 = strongly disagree to 7 = strongly agree), the participants neither clearly agreed nor disagreed (*M* = 3.50, *SD* = 2.04) with the use of a vehicle that might have this defect. The RWA failure was noticed by all participants (*n* = 21) and the use of a vehicle that could have this failure was neither clearly approved nor rejected (*M* = 3.67, *SD* = 2.22). For a presentation of the results, see Illustration 101.

Illustration 101

Intention to use a vehicle that could exhibit the experienced failure FBA step or uncontrolled RWA, surveyed using a 7-point Likert scale.



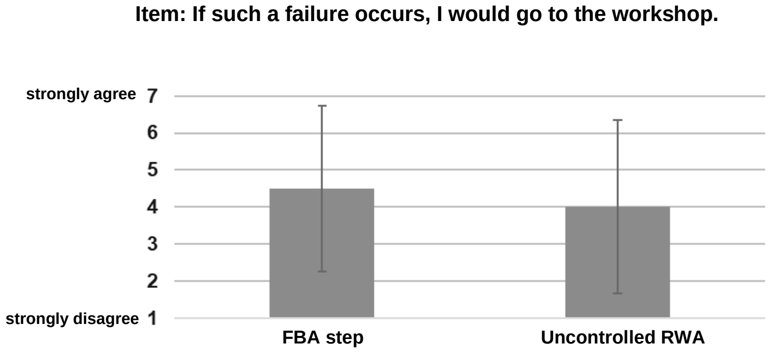
Failure	N	M	SD	Min	Max
FBA step	20	3,50	2,04	1	6
Uncontrolled **RWA**	21	3,67	2,22	1	7

Note. The item was only answered by test subjects who noticed the respective failure during the test run.

Participants were also asked to indicate the extent to which they would agree to take the vehicle to a workshop if such a failure occurred for each failure they noticed. For both failures, there was neither a clear agreement nor a clear rejection with regard to visiting a workshop (*FBA step*: $M = 4.50$, $SD = 2.24$; *uncontrolled RWA*: $M = 4.00$, $SD = 2.35$). For a presentation of the results, see Illustration 102.

Illustration 102

Intention to visit a workshop if the experienced failures FBA step or uncontrolled SHE in own vehicle occur, measured using a 7-point Likert scale.



Failure	N	M	SD	Min	Max
FBA step	20	4,50	2,24	1	7
Uncontrolled **RWA**	21	4,00	2,35	1	7

Note. The item was only answered by test subjects who noticed the respective failure during the test run.

3.5.3 Summary Study 5

The following is an overview of the main results of sub-study 5 (vehicle: sedan, failure: *FBA step I uncontrolled RWA)*.

- The design of the course appears appropriate, the test subjects are neither under- nor overchallenged.
- A learning curve of the test subjects over the course of the test becomes clear.
- The research hypothesis could be maintained for the failure *Uncontrolled RWA.* The hypothesis must be rejected for the *FBA step failure, as in the* case of *n* = 1, the track was left after the failure was activated.
- The *FBA step* failure
  - was noticed by 95% of the test subjects;
  - does not cause a significant increase in the perceived mental and physical demands or a significant deterioration in the subjectively perceived performance in the driving maneuver in question;
  - is mostly associated on an affective level with fright and surprise or no specific affective reaction;
  - provokes a behavioral response in the majority of test subjects (counter-steering/corrective behavior)
  - is classified in terms of criticality at a high level of "noticeability" to a low level of the category "disturbance of driving".
- The failure *Uncontrolled RWA*
  - was noticed by all test subjects;

- is associated with a significant increase in the perceived mental and physical demands, the perceived own performance decreases significantly in the respective driving maneuver;
- is mostly associated with shock and surprise on an affective level;
- provokes a behavioral response in the majority of test subjects (counter-steering/corrective behavior)
- is categorized at a low level of criticality in the category "disruption to driving".

The content of the results of the sub-studies is discussed for the entire test series in Section 4.

3.6 Study 6

The sixth sub-study was carried out from May 17 to 20, 2022 together with the SUV 2 vehicle and their test vehicle. The selected failure patterns were the *blocked FBA* in the slalom maneuver and the *RWA square-wave oscillation,* which was activated in the associated straight-ahead maneuver. The results are presented below.

3.6.1 Sample

N = 28 people took part in the study, 22 of whom were male. The mean age is *M* = 34.79 years (*SD* = 14.65), with the youngest subject being 20 years old and the oldest 63 years old. The mean annual mileage is *M* = 18,111.11 km (*SD* = 12,864.18 km) with a range of 7,000 km - 60,000 km. None of the test subjects stated that they had an uncorrected visual or hearing impairment.

3.6.2 Results

The presentation of the results is divided into two sections: failure-independent and failure-specific results. Firstly, the results are reported, which relate to the general difficulty of the selected maneuvers in terms of testing the selected test design, independently of the failure setups. The failure-specific results are then presented. In this section, the results for testing the controllability of both failures (FBA and RWA) are discussed first. This is followed by the subjective test data regarding the experience of the failure activation and the objective test data regarding the driver and vehicle reaction as a result of the failure activation, first for the FBA failure and then for the RWA failure.

3.6.2.1 Failure-independent results

The general difficulty of the maneuvers was recorded using three questions on the mental and physical demands as well as the subjective assessment of one's own performance based on the NASA-TLX instrument (Hart, 2006) and on a 20-point response scale (1 = very low to 20 = very high). The survey points were at the beginning of the test (after the familiarization drive) and before the end of the test (after drive 5). For the inferential statistical analysis of the learning curve, the requirements at the beginning and end of the ride were compared using a paired t-test. All persons who had valid data sets for both failure cases were included in the analysis.

All three maneuvers are classified on a low to medium range of the 20-point scale in terms of mental demands, whereby driving straight-ahead appears to be the least mentally demanding. For a presentation of the characteristic values, see Table 29. Before the last ride, the maneuvers circular drive and slalom are perceived as significantly less mentally demanding than at the beginning of the test (circular drive: $t(22) = 3.2$, $p = .002$, $d = 0.67$; slalom: $t(22) = 1.97$, $p = .03$, $d = 0.41$; straight-ahead: $t(22) = 1.08$, $p = .15$).

Table 29

Mental requirements of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high requirements) based on the NASA TLX instrument (Hart, 2006).

Mental requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	23	6,17	3,61	1	14
	before the last round	23	5,52	4,05	1	15
Circular drive	after familiarization	23	10,83	3,74	4	20
	before the last round	23	8,35	3,97	2	15
Slalom	after familiarization	23	12,52	4,05	3	20
	before the last round	23	10,74	5,00	2	20

In terms of physical demands, the three maneuvers are also classified in a low to medium range on the response scale. Again, driving straight-ahead appears to be the least demanding. For a presentation of the characteristic values, see Table 30. Over the course of the test, the physical demands of driving straight-ahead remain constant ($t(22) = 0.09$, $p = .46$), whereas they decrease significantly for the maneuvers circular drive and slalom (circular drive: $t(22) = 1.82$, $p = .04$, $d = 0.38$; slalom: $t(22) = 2.27$, $p = .017$, $d = 0.47$).

Table 30

Physical demands of the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high demands) based on the NASA TLX instrument (Hart, 2006).

Physical requirements						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	23	3,65	3,27	1	13
	before the last round	23	3,61	2,71	1	10
Circular drive	after familiarization	23	8,57	4,21	1	20
	before the last round	23	7,48	3,78	2	15
Slalom	after familiarization	23	11,22	4,12	3	20
	before the last round	23	9,78	4,47	3	20

Over the course of the test, the test subjects rated their own performance at a high level on the response scale. For a presentation of the characteristic values, see Table 31. At the end of the test, the performance in all three maneuvers was rated as significantly better compared to the beginning of the test (driving straight-ahead: $t(22) = -1.78$, $p = .045$, $d = -0.37$; circular drive: $t(22) = -3.67$, $p < .001$, $d = -0.76$; slalom: $t(22) = -3.22$, $p = .002$, $d = -0.67$).

Table 31

Assessment of own performance in the maneuvers straight-ahead, circular drive and slalom based on a 20-point scale (1 = very low to 20 = very high performance) based on the NASA TLX instrument (Hart, 2006).

Assessment of own performance						
Maneuver	**Time of measurement**	**N**	**M**	**SD**	**Min**	**Max**
Straight-ahead	after familiarization	23	16,35	2,35	12	20
	before the last round	23	17,39	2,39	13	20
Circular drive	after familiarization	23	14,48	2,68	10	20
	before the last round	23	16,09	2,17	12	20
Slalom	after familiarization	23	14,35	2,64	7	18
	Before the last round	23	15,91	2,47	11	20

3.6.2.2 Failure-specific results

In the following, the results for testing the controllability of the two failure patterns *blocked FBA* and *RWA square-wave oscillation are* discussed first. This is followed by a separate presentation of the subjective results for both failures with regard to the experience of the failure activation and then the objective test data with regard to the driver and vehicle reaction as a result of the failure events.

3.6.2.2.1 Controllability - hypothesis testing

In the context of the sixth sub-study, 100% of the test subjects with valid data records did not leave the lane with failure activation for the *blocked FBA* failure activated in the slalom maneuver. 100% of the data records comprise at least N = 20 data records. The hypothesis can therefore be maintained for this selected combination of failure parameterization and vehicle. Within the scope of the study conducted, this failure was controllable for all test subjects at the C0 level defined by experts.

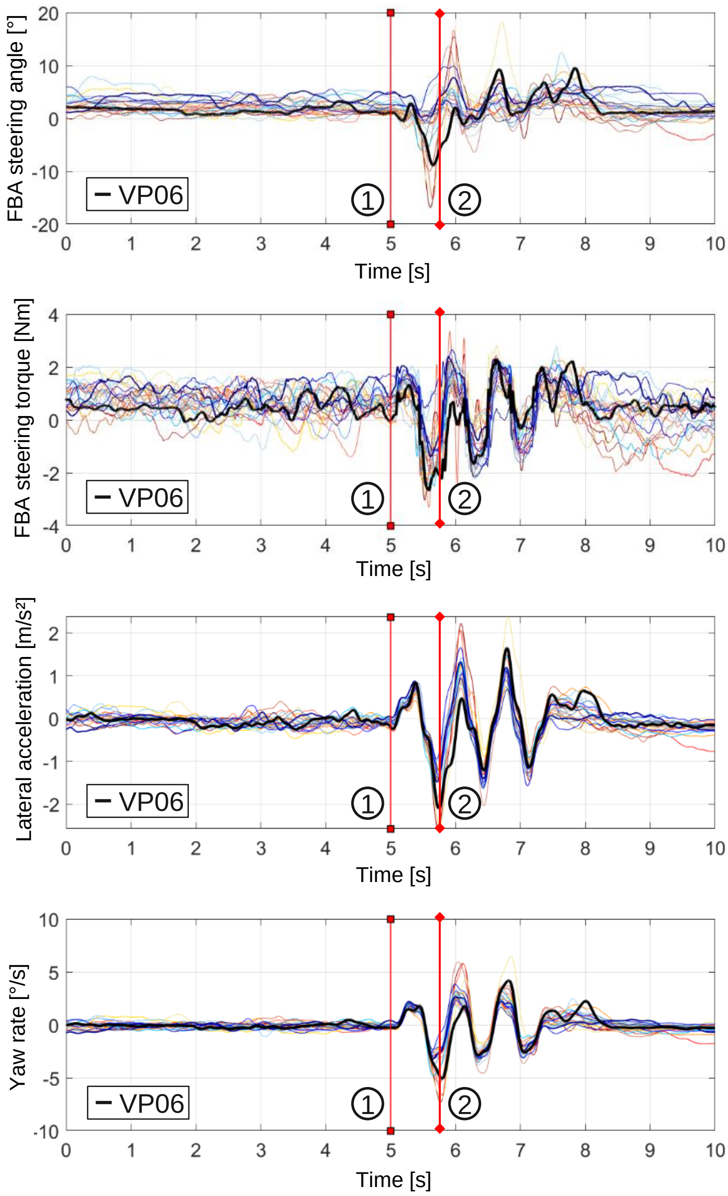
When the *RWA square-wave oscillation* failure was activated, one case (n = 1) resulted in a lane departure in the form of a touched pylon. The hypothesis must therefore be rejected for the combination of test vehicle and RWA failure. The C0 level defined by experts for this failure could not be controlled by all test subjects.

3.6.2.2.1.1 Lane departure analysis

Illustration 103 shows selected vehicle measurement variables of test subject 06 in comparison to the other valid measurement data sets of the *RWA square-wave oscillation* failure activation in the SUV 2. The time of the failure activation (1) and the time of the pylon contact (2) are marked on the time axis. With regard to the driver inputs (steering angle, steering torque), test subject 06 does not show any conspicuous driving behavior compared to the other test subjects, both before the failure activation (1) and the contact with the pylon (2). This is also reflected in the resulting vehicle dynamics (lateral acceleration, yaw rate), which also shows no abnormalities in comparison to the other test subjects before contact with the pylon. For a graphical representation of the situation, see appendix 6.6.

Illustration 103

Vehicle measurement data RWA square-wave failure activation (SUV 2) - VP06 vs. valid RWA square-wave measurement data; (1) time of failure activation, (2) time of pylon contact



3.6.2.2.2 Failure type: *Blocked FBA*

The objective vehicle measurement data for describing the vehicle and driver reaction as a result of the *blocked FBA* failure, as well as the results with regard to the subjective perception of the activation, are reported below.

3.6.2.2.3 Results of objective vehicle measurement data

The statistical distributions of the objective characteristics determined on the basis of the recorded vehicle measurement variables for the *blocked FBA* failure pattern in the SUV 2 are shown in Illustration 104. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern under investigation are also shown in Table 32. For a qualitative comparison of the objective parameters determined, the following Illustration 104 also shows the statistical evaluations for the second vehicle, which was examined as part of the overall study with the same failure pattern. However, due to vehicle and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.

In the present case, the SUV 2 shows significantly greater disturbances in terms of the driver's steering response and the resulting vehicle dynamics compared to the compact class 3. It should be noted here that the failure activation time of 200 ms in the SUV 2 was also twice as long as the failure activation time in the compact class 3. In addition, the failures were activated at different points in the slalom course.

Illustration 104

Statistical evaluation of the calculated objective parameters of Disturbance steering angle, Disturbance steering rate, Disturbance lateral acceleration and Disturbance yaw rate for the blocked FBA failure pattern
(white: SUV 2; gray: compact class 3)

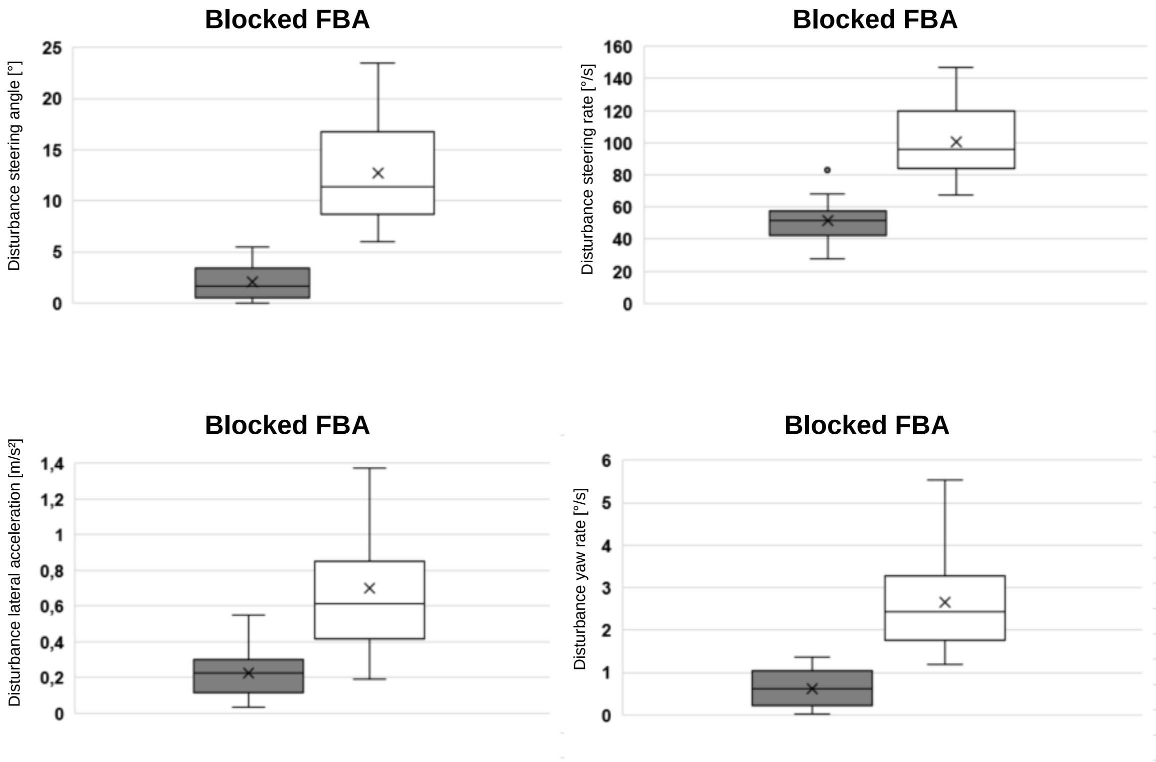


Table 32

Descriptive statistical characteristics Failure pattern blocked FBA (SUV 2)

Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
SUV 2	**Disturbance steering angle [°]**	26	12,61	4,54	11,32	6,01	23,43
	Disturbance steering rate[°/s]	26	100,58	22,13	96,09	67,40	146,99
	Dist. lateral acceleration [m/s²]	26	0,70	0,37	0,61	0,19	1,69
	Disturbance yaw rate [°/s]	26	2,67	1,08	2,42	1,20	5,53

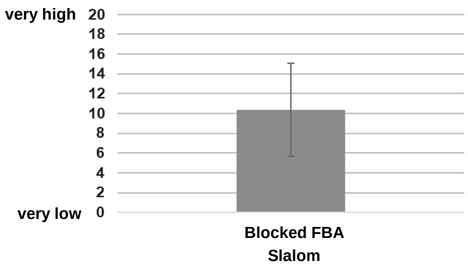
3.6.2.2.4 Results of subjective measures

In the following, the perceived difficulty of the slalom maneuver with activation of the failure is discussed. Three questions were used to record the perceived mental and physical demands as well as the subjective assessment of one's own performance on a 20-point response scale (1 = very low to 20 = very high) based on the NASA-TLX instrument (Hart, 2006). The results are presented together with those of sub-study 4 (for details see 3.1) with the compact class 3 vehicle, as the same type of failure was taken into account here. However, differences between the two sub-studies can only be described and cannot be explained due to the simultaneous manipulation of failure parameterization and vehicle. All valid data records were included in the analysis. Both in the sub-study using the SUV 2 as the test vehicle and in the sub-study with the compact class 3 and the use of their test vehicle, the mental requirements for completing the slalom maneuver with failure activation were reported in a medium range of the 20-point response scale (SUV 2: *M* = 10.35, *SD* = 4.70; compact class 3: *M* = 8.85, *SD* = 3.91). For a presentation of the results, see Illustration 105.

Illustration 105

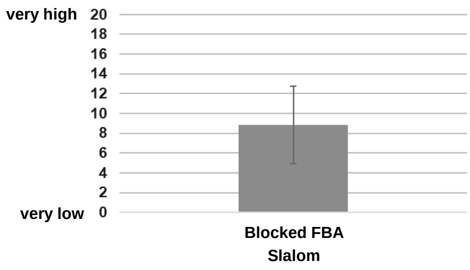
Mental requirements of the slalom maneuver with activation of the blocked FBA failure were assessed on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).

SUV 2



N	M	SD	Min	Max
26	10,35	4,70	3	20

Compact Class 3



N	M	SD	Min	Max
20	8,85	3,91	2	15

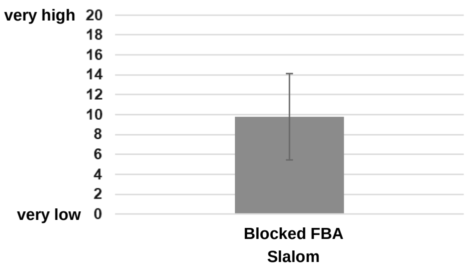
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The physical requirements for completing the slalom maneuver in the context of the experience of the failure circuit were assessed in the middle range of the response scale in both sub-studies (SUV 2: *M* = 9.77, *SD* = 4.32; compact class 3: *M* = 8.00, *SD* = 3.48). For a presentation of the results, see Illustration 106.

Illustration 106

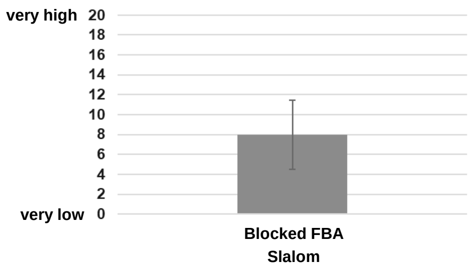
Physical demands of the slalom maneuver with activation of the blocked FBA failure were assessed on a 20-point response scale (1 = very low to 20 = very high demands) based on the NASA-TLX instrument (Hart, 2006).

SUV 2



N	M	SD	Min	Max
26	9,77	4,32	3	20

Compact Class 3



N	M	SD	Min	Max
20	8,00	3,48	2	13

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The subjective assessment of one's own performance with regard to completing the slalom maneuver with activation of the failure is in the medium to high range of the 20-point response scale in both sub-studies (SUV 2: M = 9.77, SD = 4.32, compact class 3: M = 8.00, SD = 3.48). For a presentation of the results, see Illustration 107.

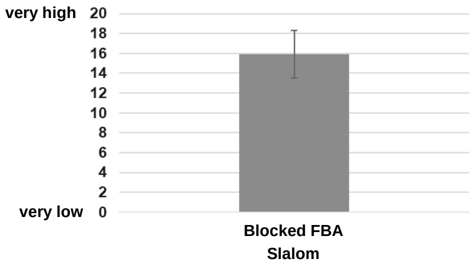
Illustration 107

Subjective performance in completing the slalom maneuver with activation of the blocked FBA failure measured on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).

SUV 2

N	M	SD	Min	Max
26	16,46	2,23	11	20

Compact Class 3



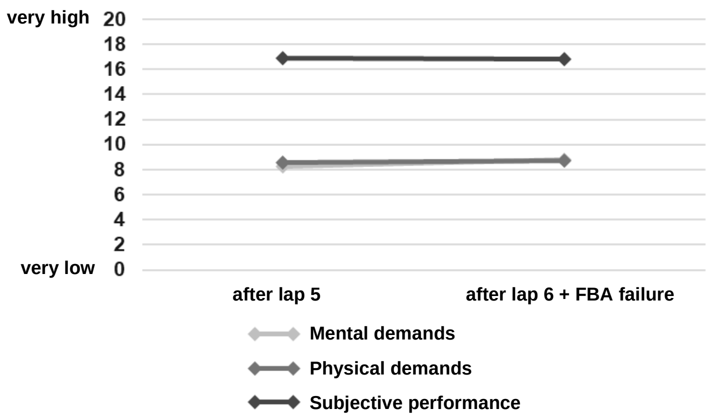
N	M	SD	Min	Max
20	15,90	2,38	12	20

Remark. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization

With the aim of identifying the influence of the failure activation on the perceived difficulty of the maneuver, the perceived mental and physical demands as well as the subjective assessment of one's own performance after round 5 without failures were compared with those after the failure activation in round 6. The underlying assumption was that at this advanced stage of the test, the learning curve of the test subjects with regard to completing the maneuver had already flattened out and that differences between the assessments could be attributed to the experience of the failure. Due to the randomized activation of both failures in this sub-study (FBA and RWA failures) in rounds 3 and 6, the number of data sets included in the analysis is half of all valid data sets (in which the respective failure was activated in round 6). The inferential statistical analysis of the comparison of the two survey times was carried out using a paired *t-test*. For the sub-study with the SUV 2, it was shown that the mental and physical demands, as well as the assessment of one's own performance, remained constant across both survey times (mental demands: $t(12) = -0.5$, $p = .31$, physical demands: $t(12) = -0.14$, $p = .44$, subjective performance: $t(12) = 0.22$, $p = .41$). For a graphical representation of the course and the associated characteristic values, see Illustration 108.

Illustration 108

*Subjectively perceived difficulty measured using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands and the subjectively perceived performance in the slalom maneuver in lap 5 without the failure activation compared to lap 6 after the failure blocked FBA was activated. (Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



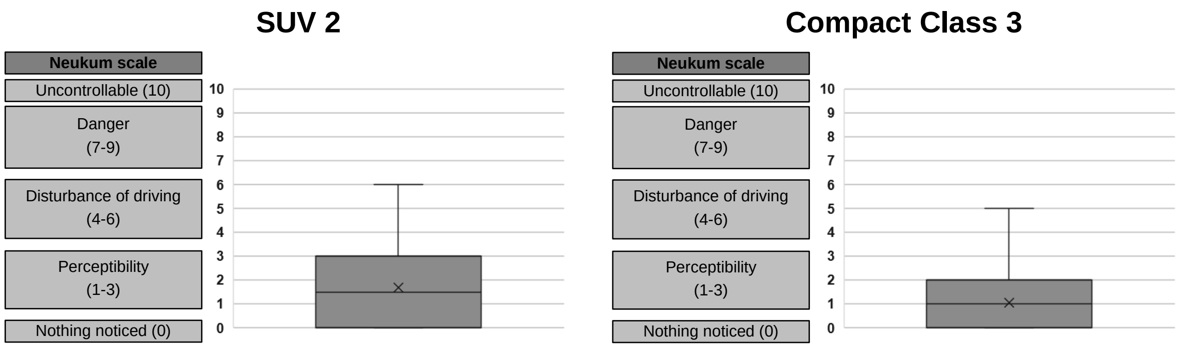
Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	13	8,23	4,29	2	16
	after lap 6 + failure	13	8,77	4,25	3	15
Physical demands	after lap 5	13	8,54	3,87	3	15
	after lap 6 + failure	13	8,69	4,21	3	15
Subjective performance	after lap 5	13	16,92	2,87	11	20
	after lap 6 + failure	13	16,85	2,70	11	20

At the time the failure was triggered, the experimenter observed the behavior of the test subjects to record an initial reaction to the event. In this case, however, none of the test subjects showed a noticeable reaction for the experimenter. After completing the maneuver, the test subjects were asked whether they had noticed anything special. This was answered in the affirmative by 73.08% of the test subjects (*n* = 19). The event was most frequently described as a blockage / change in the steering wheel (*n* = 14). Following on from this, the test subjects who had noticed the failure were asked to describe their reaction on an affective, cognitive and behavioral level. To record the affective reaction, they were asked to describe the feelings that accompanied the perception of the failure. However, the most common response was that there was no specific emotional reaction (*n* = 11). A (slight) surprise/fright (*n* = 3) was mentioned much less frequently, but was the second most frequently mentioned emotion. To record the cognitive reaction, the test subjects were asked to describe their thoughts at the time of the event. The thought that something was different / wrong was mentioned most frequently (*n* = 5). *n* = 3 times the question "was there something wrong?" was asked. To describe the behavioral reaction, the test subjects were asked to explain what they did as a result of the event. The most common responses were to continue driving normally (n = 10) or to steer/correct (*n* = 9).

The Neukum scale (Neukum & Krüger, 2003) was chosen to record the subjectively perceived criticality associated with the failure activation. At this point, the results of this sub-study are compared again with those from sub-study 4 (for details see 3.4) with the compact class 3 vehicle, as this dealt with the same type of failure. Possible differences between the results of the two sub-studies can only be described, but cannot be explained due to the simultaneous variation of vehicle and failure parameterization. All valid data sets were taken into account for the analysis. The mean subjectively perceived criticality of the activation of the *blocked FBA* failure in sub-study 6 with the SUV 2 was in the low to medium range of the noticeability category (*M* = 1.69, *SD* = 1.57) on the 11-point response scale (0 = nothing noticed - 10 = vehicle not controllable). In the sub-study with the compact class 3, the criticality is descriptively classified somewhat lower at a low level of the noticeability category (*M* = 1.05, *SD* = 1.32). For a representation of both distributions, see Illustration 109.

Illustration 109

Subjectively perceived criticality of the blocked FBA failure type in the two sub-studies with the SUV 2 and compact class 3 vehicles, measured using the Neukum scale (Neukum & Krüger, 2003).



N	M	SD	Median	Percentile 25	Percentile 50	Percentile 75
26	1,69	1,57	1,50	0,00	1,50	3,00

N	M	SD	Median	Percentile 25	Percentile 50	Percentile 75
20	1,05	1,32	1,00	0,00	1,00	2,00

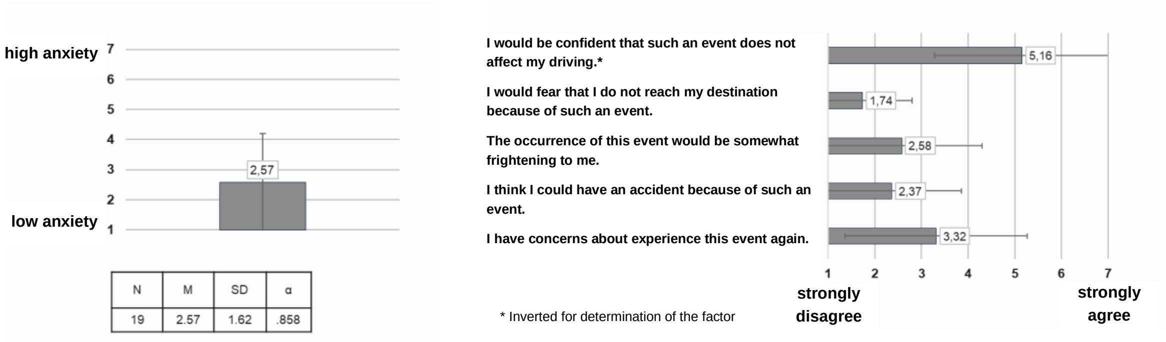
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The two factors *Anxiety* and *Perceived Safety* from the *Car Technology Acceptance Model* (Osswald et al., 2012) were selected and adapted to assess the experience of the situation. This data was collected from all test subjects who stated that they had noticed the failure.

Anxiety with the occurrence of the failure event was classified at a low level on the 7-point scale (1 = low anxiety - 7 = high anxiety) (*M* = 2.57, *SD* = 1.62). For a presentation of the results for the *anxiety* factor and the individual items used for the calculation, see Illustration 110.

Illustration 110

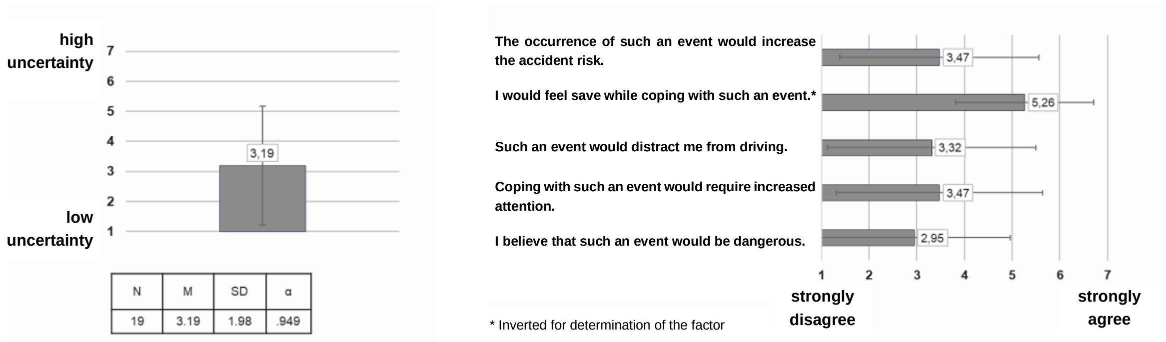
Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the Blocked FBA failure event.



The perceived uncertainty with the occurrence of the failure event is reported on a medium level of the 7-point response scale (1 = low uncertainty - 7 = high uncertainty) (*M* = 3.19, *SD* = 1.98). For a presentation of the results of the *Perceived Safety* factor and the individual items, see Illustration 111.

Illustration 111

Results of the factor Perceived Safety (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the failure event Blocked FBA.



3.6.2.2.5 Failure pattern: *RWA square-wave oscillation*

The objective results describing the vehicle and driver reaction as a result of the activation of the *RWA square-wave oscillation* failure pattern in the straight-ahead maneuver are described below, as well as the results regarding the subjective perception of the failure activation.

3.6.2.2.6 Results of objective vehicle measurement data

The statistical distributions of the characteristic objective values for the *RWA square-wave oscillation* failure pattern in the SUV 2 determined on the basis of the recorded vehicle measurement variables are shown in Illustration 112. The figure shows the distributions for the determined disturbance influence on the steering angle, the disturbance steering rate, the disturbance lateral acceleration and the disturbance yaw rate. The descriptive statistics of the objective characteristics determined for the vehicle for the failure pattern under investigation are also shown in Table 33 summarized. For a qualitative comparison of the objective parameters determined, the following Illustration 112 also shows the statistical evaluations for the second vehicle, which was examined as part of the overall study with the same failure pattern. However, due to vehicle- and system-specific differences and not completely identical failure patterns, a direct comparison of the objective characteristic values is only possible to a limited extent and is of limited significance.

In the present case, the SUV 2 shows lower disturbance influences with regard to the driver's steering response compared to the compact class 3 with regard to the *RWA square-wave square-wave oscillation* failure pattern. Both the disturbance influence on the steering angle and the disturbance steering rate are lower than the comparative values in the compact class 3. There are also lower disturbance influences in terms of lateral acceleration and yaw rate with regard to the vehicle reaction.

Illustration 112

Statistical evaluation of the calculated objective parameters steering angle, disturbance steering rate, disturbance lateral acceleration and disturbance yaw rate for the RWA square-wave oscillation failure pattern
(white: SUV 2, gray: compact class 3)

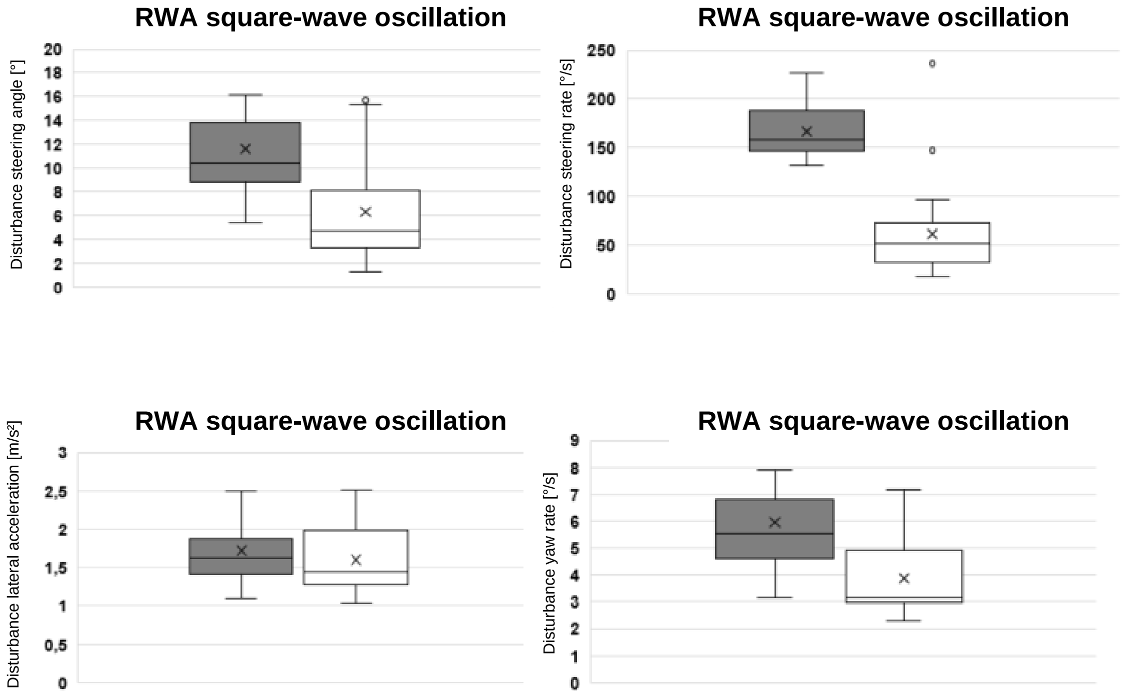


Table 33

Descriptive statistical characteristics Failure pattern RWA square-wave oscillation (SUV 2)

Vehicle	Disturbance variable	N	M	SD	Median	Min.	Max.
SUV 2	**Disturbance steering angle [°]**	24	6,23	4,10	4,68	1,30	15,68
	Disturbance steering rate[°/s]	24	61,42	46,75	50,59	17,22	236,28
	Dist. lateral acceleration [m/s²]	24	1,60	0,44	1,44	1,04	2,50
	Disturbance yaw rate [°/s]	24	3,88	1,44	3,14	2,30	7,17

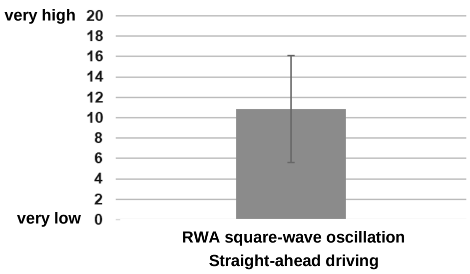
3.6.2.2.7 Results of subjective measures

The following passage deals with the perceived difficulty of the maneuver straight-ahead with failure activation, which was surveyed using three questions based on the NASA-TLX inventory regarding the mental and physical demands, as well as the subjective assessment of one's own performance using a 20-point response scale (1 = very low - 20 = very high). The results are compared with those of sub-study 4 (for details see 3.4) with the compact class 3, as the *RWA square-wave* failure type was also considered here. An explanation of possible differences between the two sub-studies is not possible due to the simultaneous variation of failure parameterization and vehicle. All valid data sets were included in the analysis. In both studies, the mental requirements for completing the maneuver of driving straight-ahead with failure activation were assessed in the middle range of the 20-point response scale (1 = very low - 20 = very high) (SUV 2: *M* = 10.83, *SD* = 5.26; compact class 3: *M* = 11.25, *SD* = 5.65). For a presentation of the results, see Illustration 113.

Illustration 113

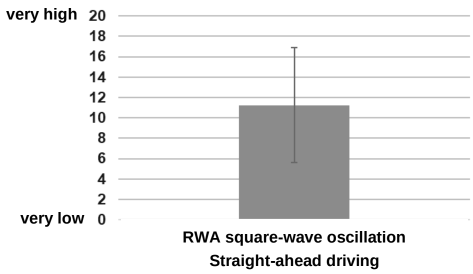
Mental requirements of the maneuver straight-ahead with activation of the failure RWA square-wave oscillation collected on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).

SUV 2



N	M	SD	Min	Max
24	10,83	5,26	3	20

Compact Class 3



N	M	SD	Min	Max
28	11,25	5,65	1	20

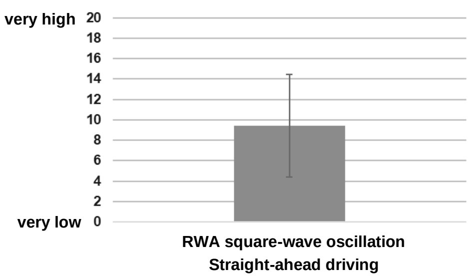
Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

The physical requirements for completing the maneuver with activation of the failure are also classified in the middle range of the response scale in both sub-studies (SUV 2: *M* = 9.42, *SD* = 5.03; compact class 3: *M* = 8.29, *SD* = 5.52). For an illustration see Illustration 114.

Illustration 114

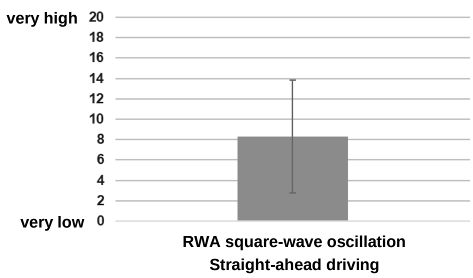
Physical requirements of the straight-ahead maneuver with activation of the RWA square-wave oscillation failure were determined on a 20-point response scale (1 = very low to 20 = very high requirements) based on the NASA-TLX instrument (Hart, 2006).

SUV 2



N	M	SD	Min	Max
24	9,42	5,03	2	20

Compact Class 3



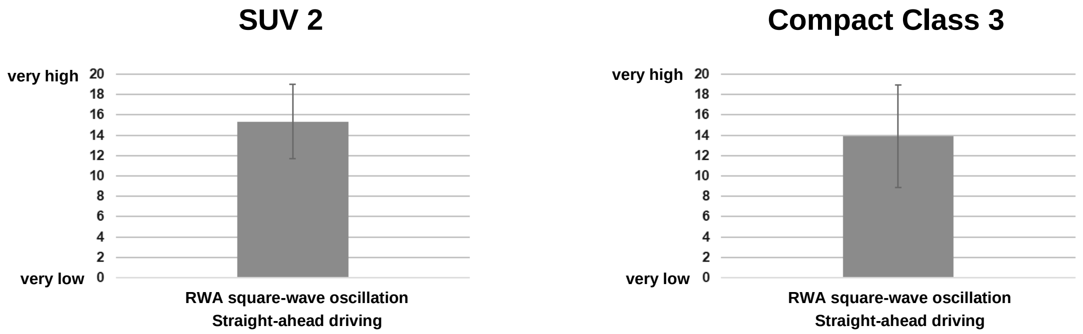
N	M	SD	Min	Max
28	8,29	5,52	1	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

In both sub-studies, the assessment of one's own performance when driving in a straight-ahead with failure activation is classified in a medium to high range on the 20-point response scale (SUV 2: *M* = 15.33, *SD* = 3.66; compact class 3: *M* = 13.89, *SD* = 5.04). For a presentation of the results, see Illustration 115.

Illustration 115

Subjective performance when completing the straight-ahead maneuver with activation of the RWA square-wave oscillation failure measured on a 20-point response scale (1 = very low to 20 = very high performance) based on the NASA-TLX instrument (Hart, 2006).



N	M	SD	Min	Max
24	15,33	3,66	5	20

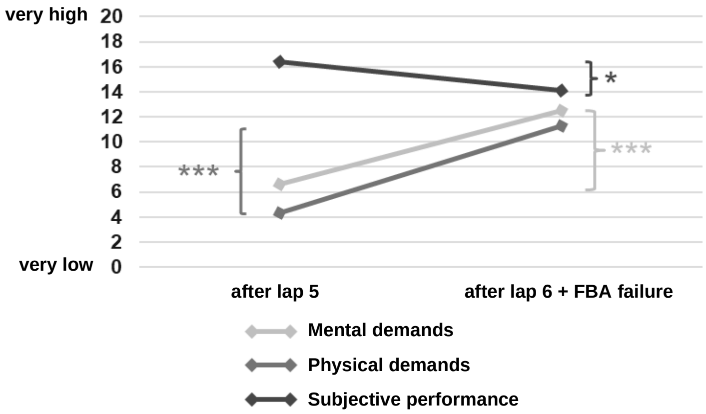
N	M	SD	Min	Max
28	13,89	5,04	1	20

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

To identify the influence of the failure event on the perceived difficulty of the maneuver straight-ahead, the mental and physical demands as well as the subjective assessment of one's own performance after lap 5 without failure activation were compared with those after activation in lap 6. The inferential statistical analysis of the comparison between the two survey times was carried out using a paired *t-test*. It was found that mental and physical demands increased significantly in the last round with activation of the RWA failure compared to the previous round without failure activation (mental demands: $t(13) = -5.86$, $p < .001$, $d = -1.57$; physical demands: $t(13) = -8.05$, $p < .001$, $d = -2.15$). Subjectively perceived performance also decreased significantly ($t(13) = -2.32$, $p = .19$, $d = 0.62$). For a graphical representation of the course and the associated characteristic values, see Illustration 116.

Illustration 116

*Subjective perceived difficulty measured using three items based on the NASA-TLX instrument (Hart, 2006) in relation to the mental and physical demands as well as the subjectively perceived performance in the straight-ahead maneuver in lap 5 without failure activation compared to lap 6 after activation of the RWA square-wave oscillation failure. (Significance levels: * = 5% level, ** = 1% level, *** = 0.1% level)*



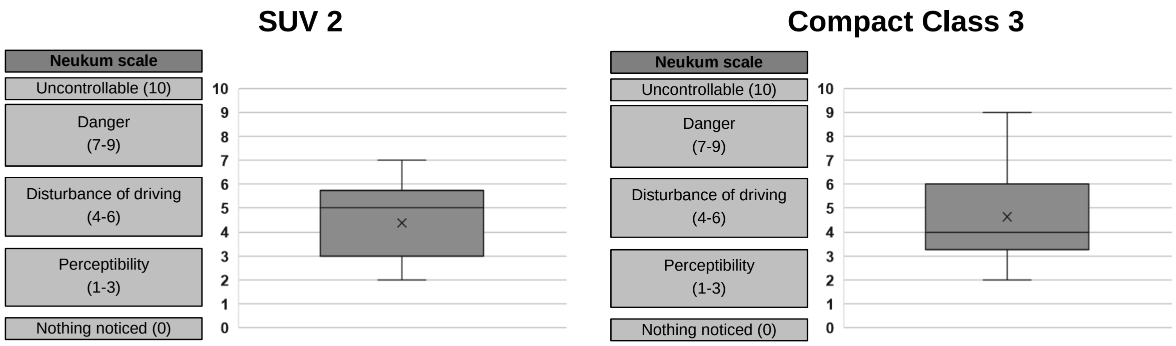
Construct	Time of measurement	N	M	SD	Min	Max
Mental demands	after lap 5	14	6,64	4,18	1	15
	after lap 6 + failure	14	12,50	4,65	4	20
Physical demands	after lap 5	14	4,36	2,98	1	10
	after lap 6 + failure	14	11,29	4,25	6	20
Subjective performance	after lap 5	14	16,43	2,82	10	20
	after lap 6 + failure	14	14,14	3,74	5	18

Observation of the test subjects at the time of the failure activation revealed that 37.50% (*n* = 9) showed an observable reaction, most frequently in the form of a verbal utterance (*n* = 5). When asked, 100% of the test subjects stated that they had noticed the event. The majority described this as the vehicle swerving / interfering with the steering (*n* = 21), while jerking was mentioned much less frequently, although it was the second most common (*n* = 4). To record the affective reaction to the event, the test subjects were asked to describe their feelings when they noticed the event. The most frequent response was surprise/fright (*n* = 9), followed by excitement/tension as the second most frequent response (*n* = 4). On a cognitive level, the test subjects reported dealing with the cause of the event ("What was that?") (*n* = 8), or giving themselves mental driving instructions ("Don't turn the steering wheel") (*n* = 6). With regard to the behavioral reaction, counter-steering/correcting was reported most frequently (*n* = 15), as well as holding the steering wheel more firmly (*n* = 6). If the event were to happen again, 75% of the test subjects (*n* = 18) would behave in the same way again.

The Neukum scale (Neukum & Krüger, 2003) was used to assess the subjectively perceived criticality. The results are presented below in comparison with those of sub-study 4 (for details see 3.4) with the compact class 3 vehicle, as the same type of failure was taken into account here. All valid data records were taken into account for the analysis. However, it should be noted that differences between the sub-studies can only be shown, but not explained. The reason for this is the simultaneous variation of vehicle and failure parameterization. In both sub-studies, criticality was classified on the 11-point scale (0 = nothing noticed - 10 = vehicle not controllable) at a low to medium level in the category of driving disturbance (SUV 2: *M* = 4.38, *SD* = 1.50; compact class 3: *M* = 4.64, *SD* = 1.68;). For a representation of both distributions, see Illustration 117.

Illustration 117

Subjectively perceived criticality of the RWA square-wave oscillation failure type in the two sub-studies with the SUV 2 and compact class 3 vehicles measured using the Neukum scale (Neukum & Krüger, 2003).



N	M	SD	Median	Percentile		
				25	50	75
24	4,38	1,50	5,00	3,00	5,00	5,75

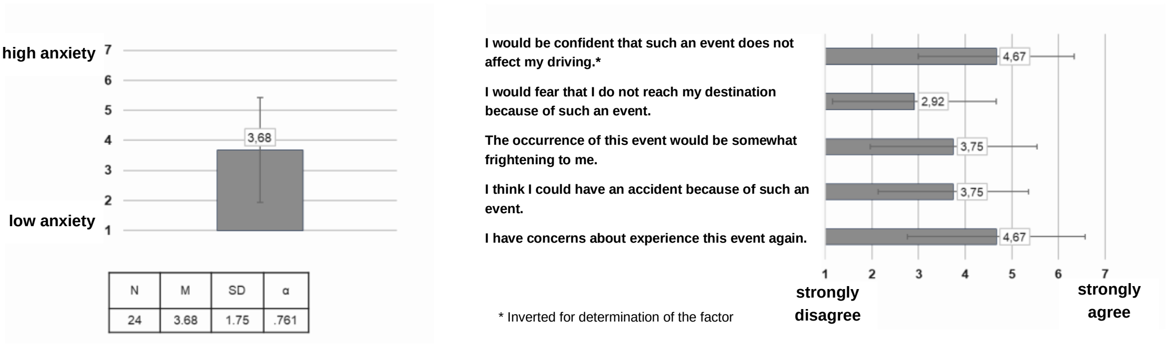
N	M	SD	Median	Percentile		
				25	50	75
28	4,64	1,68	4,00	3,25	4,00	6,00

Note. A comparison of the results of both studies is only possible to a limited extent due to the simultaneous variation of vehicle and failure parameterization.

Anxiety on noticing the event was reported at a medium level on the 7-point response scale (1 = low anxiety - 7 = high anxiety) (*M* = 3.68, *SD* = 1.75). For a presentation of the factor and the individual items used in the calculation, see Illustration 118.

Illustration 118

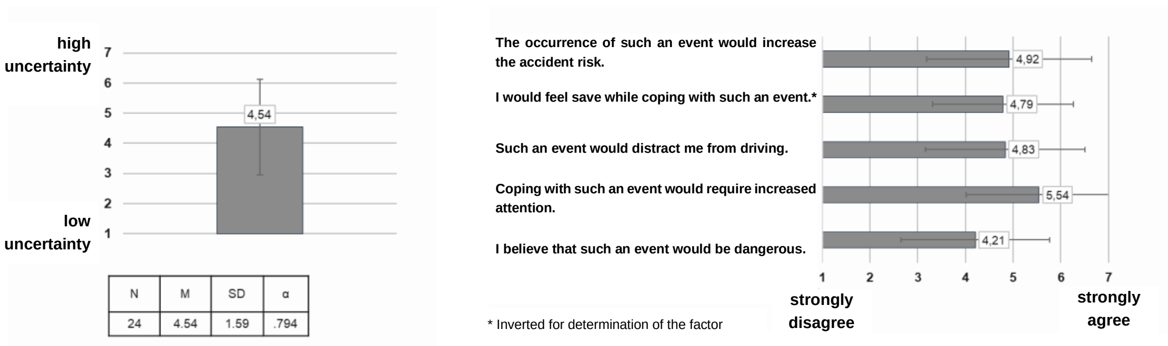
Results of the Anxiety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the RWA square-wave failure event.



The perceived uncertainty with noticing the failure was classified in a medium range of the 7-point scale (1 = low uncertainty - 7 = high uncertainty) (*M* = 4.54, *SD* = 1.59). For a representation of the factor and the individual items on which the calculation is based, see Illustration 119.

Illustration 119

Results of the Perceived Safety factor (left) and the underlying individual items (right) based on the Car Technology Acceptance Model (Osswald et al., 2012) and related to the activation of the RWA square-wave failure event.

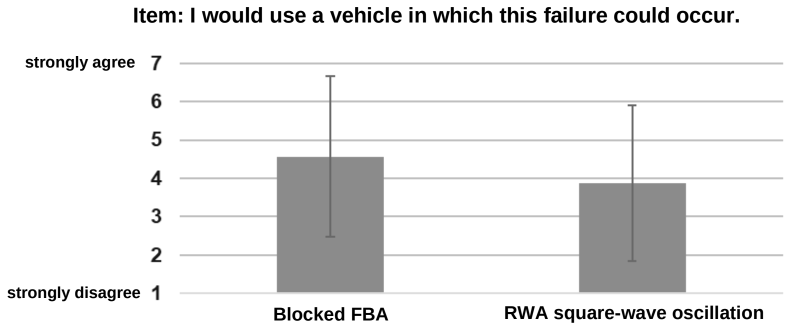


3.6.2.3 Follow-up survey

After the test subjects had been informed about the subject of the study, a follow-up survey was conducted. The questions were only answered by the test subjects if they had noticed the respective failure during the test. First, the test subjects were asked to assess whether they would use a vehicle in which the failure they had experienced could occur (response scale: 1 = strongly disagree - 7 = strongly agree). The test subjects who had noticed the failure blocked *FBA* (n = 18) neither clearly agreed nor clearly disagreed with its use (*M* = 4.56, *SD* = 2.09). For the *RWA square-wave oscillation failure,* the intention to use was at a slightly lower, but also neutral, level (*M* = 3.87, *SD* = 2.03). This failure was noticed by all test subjects. For a presentation of the results, see Illustration 120.

Illustration 120

Intention to use a vehicle that could exhibit the experienced failure of blocked FBA or RWA square-wave oscillation, surveyed using a 7-point Likert scale.



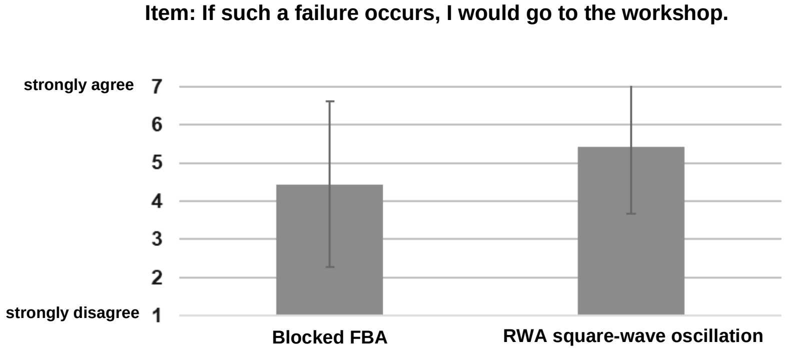
Failure	N	M	SD	Min	Max
Blocked **FBA**	18	4,56	2,09	1	7
RWA square-wave	23	3,87	2,03	1	7

Note. The item was only answered by test subjects who noticed the respective failure during the test run.

The test subjects were then asked to indicate how much they would agree with the statement to visit a workshop if they noticed the respective failure (7-point response scale: 1 = strongly disagree - 7 = strongly agree). The people who noticed the FBA failure (*n* = 18) neither clearly agreed nor disagreed with the statement (*M* = 4.44, *SD* = 2.18). With regard to the *RWA square-wave oscillation,* which was noticed by all test subjects, they tended to agree with the statement on average (*M* = 5.43, *SD* = 1.75). See Illustration 121 for an illustration of the results.

Illustration 121

Intention to visit a workshop if the experienced failures Blocked FBA or RWA square-wave oscillation in own vehicle occur, assessed using a 7-point Likert scale.



Failure	N	M	SD	Min	Max
Blocked **FBA**	18	4,44	2,18	2	7
RWA square-wave	23	5,43	1,75	2	7

Note. The item was only answered by test subjects who noticed the respective failure during the test run.

3.6.3 Summary Study 6

The following is an overview of the main results of sub-study 6 (vehicle: SUV 2; failure: *blocked FBA* / *RWA square-wave oscillation*).

- The design of the course appears appropriate, the test subjects are neither under- nor overchallenged.
- A learning curve of the test subjects over the course of the test becomes clear.
- The research hypothesis was retained for the *blocked FBA failure.* The hypothesis must be rejected for the *RWA square-wave oscillation failure,* as in the case of $n = 1$, track deviation occurred after failure activation.
- The *blocked FBA* failure
    - was noticed by 77% of the test subjects;
    - does not cause a significant increase in the perceived mental and physical demands or a significant deterioration in the subjectively perceived performance in the driving maneuver in question;
    - does not provoke any specific affective or behavioral response in the test subjects;
    - is classified at a medium level of "noticeability" in terms of criticality.
- The *RWA square-wave oscillation* failure
    - was noticed by all test subjects;
    - is associated with a significant increase in the perceived mental and physical demands, and the perceived own performance also drops significantly with failure activation;
    - is mostly associated with shock and surprise on an affective level;

- provokes a behavioral reaction (counter-steering/correction) in the majority of the test subjects, a quarter of the test subjects report that they would react by reducing their speed if they experienced the failure again;
- is classified in terms of criticality at a medium level in the "impaired driving" category, ranging from a medium level of "noticeability" to a low level in the "dangerousness" category.

The content of the results of the sub-studies is discussed for the entire test series in section 4.

4 Discussion

The aim of the six individual studies was to examine the assessment of the controllability of various failure patterns in a steer-by-wire steering system by experts as part of a test subject study on the test track and in the context of the driving speeds that can be displayed there. The focus here was on the hypothesis that a C0 level defined by experts with regard to a failure pattern in normal drivers in predefined maneuvers does not lead to a departure from a driving corridor specified by the experts within a defined period of time from failure activation. As shown in the reports on the results of the individual studies, this hypothesis was tested in a total of twelve cases (for both failures in all six individual studies). Illustration 122 shows an overview of all twelve failure-vehicle combinations.

Illustration 122
Overview of hypothesis testing

FBA Selfsteer & Loss of Feeback	**FBA step**	**Blocked FBA**	**Blocked RWA**	**Uncontrolled RWA**	**RWA square-wave oscillation**
SUV 1 ✓	Compact Class 2 ✓	Compact Class 3 ✓	SUV 1 ✓	Compact Class 2 ✓	Compact Class 3 ✓
Compact Class 1 ✓	Sedan X *n = 1 lane departure*	SUV 2 ✓	Compact Class 1 X *n < 20*	Sedan ✓	SUV 2 X *n = 1 lane departure*

In nine out of twelve cases, the hypothesis could be maintained based on the study results (marked in green). In the case of testing the failure *blocked RWA* in the compact class 1 vehicle (marked in gray), the test could not be carried out because the minimum number of 20 valid data records was not available in the evaluation. In the remaining and analyzed n = 15 valid data records, no lane departure occurred with a test subject. Both the *FBA step* in the sedan and the *RWA square-wave oscillation* in the SUV 2 vehicle (marked in red) resulted in one case of lane departure as a result of the failure activation by one test subject in each case, so that the hypothesis for these two failure-vehicle combinations had to be rejected.

In addition to the objective criterion for controllability, the measured vehicle parameters and questionnaire data provide further information about the initial driver and vehicle reaction to the various failure events, both on an objectively measurable and subjectively perceived level. This showed that

- drivers are able to discriminate between different failure events;
- These failure events are assessed with different levels of criticality;
- the objectively measurable driver and vehicle reactions per failure-vehicle combination exhibit similar behavior;
- some failure events are not perceived by the majority of drivers despite being objectively verifiable;

- individual test subjects also perceive failures as very critical, but the mean value of the sample is in the low to medium range of subjective criticality, depending on the failure pattern;
- test subjects subjectively report a feeling of shock or surprise in response to some failures.

The results of the study must always be understood, interpreted and discussed against the background of the method used. The most influential decisions and assumptions made for the study series are discussed below.

From a methodological point of view, the number of failure recordings per test subject can be cited here first (see section 2.1). The study results show that it is not advisable to consider more than two different failure patterns per subject, even for future experiments. For the investigation of controllability, the provocation and recording of the initial reaction of the test subjects is particularly decisive, but this is influenced by an expectation effect with an increasing number of failure displays. Test subjects know that their behavior is being observed and thus develop a certain expectation with regard to any special events that are not expected in everyday road traffic.

In addition, experience has shown that test subjects behave very attentively when conducting a study on a test site. This phenomenon is well known and described in the psychological literature under the term "reactivity problem". When creating the experimental design, the reactivity problem was discussed with regard to the possibility of presenting several failure activation situations. For example, the increased vigilance of the test subjects could already stand in the way of a second failure activation in the sense of an increased probability of detection. Accordingly, three methodological countermeasures were installed in the present study, which, however, cannot completely resolve the reactivity problem. Firstly, the location and time of the activation were systematically varied and, by combining the three maneuvers into an overall course, could not be predicted by the test subjects. Furthermore, a cover story was presented to draw attention to a focus not associated with the study. On the other hand, so-called distractor runs were carried out in which no failure activation was carried out, but the questioning was kept constant. In this way, the failure activation could be masked in the best possible way. Increased vigilance and thus responsiveness of the test subjects is nevertheless given due to driving on a test site.

It should also be borne in mind that controllability is mapped exclusively via objectively measurable criteria (see section 2.4.1). The attempt to additionally depict the controllability of a failure situation by normal drivers using subjective criteria was not pursued further in this study for the following reasons. The decisive factor for this decision was that the subjective perception of a situation is independent of whether or not drivers were objectively in control of the situation. Subjective judgments are naturally subject to greater inter- and intra-individual variability, as they are influenced by factors such as previous experience, form on the day, mood, etc.. The aforementioned factors are also proven to influence the driving performance of drivers (cf. Precht et al., 2017). However, a similar influence on the objective measure of controllability was found by controlling the factors described in section 2.6 and the familiarization rounds that were carried out. Thus, the objective measurement of controllability represents a highly reliable measure of controllability. This assumption was confirmed in the post-hoc analysis of the statistical correlation between the objective vehicle measurement variables after failure activation and the subjective judgments given on the Neukum scale (Neukum & Krüger, 2003). In the course of this, only individual significant correlations could be shown. The descriptive comparison also shows cases in which a comparatively low failure reaction of the vehicle was rated very critically, while in others a very pronounced driving dynamics as a consequence of the failure activation was rated as uncritical. In the descriptive cases cited here, however, the failure was objectively controlled. On the one hand, this shows that the choice of an objective criterion for measuring controllability appears suitable, but on the other hand it also highlights the additional value of subjective measurement methods. As a result, 3.5% (n = 10 evaluations) of the total number of valid failure placements (n = 279) were evaluated with the attribute "dangerous" (scale values 7, 8 and 9; Neukum scale). The results show that the objective controllability can be fundamentally opposed by a high perceived criticality of the drivers. In the course of system development and design, this aspect should also be taken into account with regard to the probability of a failure occurring in the event of corresponding market penetration, as a negative correlation between the constructs of perceived criticality and acceptance can be assumed (Numan, 1998; Chavaillaz et. al., 2016).

With regard to the interpretation of the study results, the definition of the operationalization of the observable controllability of a driving situation is one of the most important aspects. In the present study, a critical situation was defined by a lane departure, which was operationalized by touching or driving over the pylons marking the lane. Unlike the alternative measure of absolute failure-induced lane departure, the selected criterion does not include the initial offset in the pylon lane. As a result, a comparatively small lane offset for a test subject who, despite the instruction to drive through the center of the lane, was oriented further out to one side, is more likely to lead to a pylon contact than a larger lane offset for a test subject driving in the center. In this context, a comparison of the objective vehicle measurements of all test subjects, as described in sections 3.5.2.2.1.1 and 3.6.2.2.1.1 there was no noticeably worse driver reaction to the failure activation in those test subjects who left the lane. The use of a pylon lane was chosen by the experts for good reason and set as a boundary condition for the test subject study. The core of the justification is, on the one hand, the comprehensible consistency with the methodological procedure of the expert tests of the OEMs, even though higher dynamics are explicitly applied there. Secondly, a test procedure had to be used that could be applied to all vehicles from the various OEMs with different measurement technology requirements. In addition, the procedure has a fit with the reality of everyday drivers and their usual driving style. It can be assumed that test subjects who show a driving style oriented to one side in the test on the test track will also show this in everyday road traffic. This means that even a comparatively small lane offset could cause a critical situation for individual drivers. This makes it clear that the actual, absolute lane offset is not an unrestrictive measure of the criticality of a driving situation, as this is always in relation to the given driving situation.

A methodological point to be discussed based on this is the topic of the selected maneuvers and the capping of the maximum speed at 80 km/h. The maneuvers selected by the expert panel and the speeds agreed there and reduced in comparison to expert tests were aimed at designing a driving task appropriate to the driving skills of normal drivers that would enable a noticeable effect of the failure activation (see section 2.3). From this point of view, the study results must be evaluated against the background of the selected speed range and cannot be extrapolated directly to higher speeds, as the dynamics and the corresponding steering behavior do not behave linearly with increasing speeds. In this context, it must be taken into account that higher recommended speeds in real traffic are generally accompanied by wider control cross-sections compared to the test person study. In view of the fact that higher speeds can in principle also occur with normal drivers, a safe, standardized test environment would be necessary for corresponding subject studies, which should not be at the expense of dynamics. For example, a study in a dynamic driving simulator with correspondingly elaborate maneuvers could represent a further useful component for evaluating the controllability of errors by normal drivers in a safe study environment.

However, the course proved to be suitable for a test subject study with normal drivers on the test track and the reduction in speed was necessary. The selected course, consisting of the three maneuvers of driving straight-ahead, circular drive and slalom, was neither too challenging nor too demanding for the test subjects. In addition, there was a learning curve in driving performance with increasing number of repetitions of the maneuvers. In order to minimize the influence of this learning process on the results, the procedure of carrying out at least two familiarization rounds before starting the actual experimental test is recommended as a necessary prerequisite. As expected, the slalom proved to be the most challenging of the three maneuvers. As described in section 2.7 it was not possible for the majority of the test subjects to maintain the recommended speed of 60 km/h over the entire length of the maneuver. In addition, more pylons were touched or driven over without the influence of failure activation compared to the slalom. One way to reduce the potential influence of failure-independent pylon contacts on the study results is to further reduce the recommended speed for the slalom maneuver or to extend the lane width in subsequent studies. However, it should be noted that both the change in dynamics and the mitigation of the required driving task can have an influence on the perceived criticality of a failure. After conducting the test, the fundamental decision to choose the slalom as the test maneuver for the failures requiring a changing steering angle appears to be justified. The maneuver proved to be fundamentally feasible for normal drivers and appears to be more suitable for the test than the alternative of a dynamically driven double lane change, which was discussed and tested in the preparation phase of the study.

Along with the selection of the various maneuvers, the procedure chosen to exclude test subjects due to poor driving performance must also be discussed. The aim of the chosen procedure was to define an objective criterion, applicable to all studies, which excludes those test subjects from the evaluation who, in comparison to the rest of the sample, were conspicuous in terms of the number of times they touched the pylons, irrespective of a failure activation. It is conceivable, for example, that such an anomaly is based on the fact that a person who drives a small car in everyday life has problems estimating the dimensions of a significantly larger test vehicle and therefore touches or drives over individual cones even without a failure activation. For this reason, the mean value of the maneuvers that were not completed without failures was calculated for each sample and all test subjects who were more than two standard deviations above their respective sample mean value were excluded. As a result, although the number of "permissible" cone touches varied per failure-vehicle combination, the criterion was uniformly defined across all parts of the study and can also be objectively applied to future studies.

The reason for the necessity of such a criterion lies in the minimization of possible disturbing influences on the criterion of lane departure as a measure of the controllability of a driving situation. When completing a maneuver with failure activation, contact with the limiting pylons can basically have various causes. On the one hand, contact can be caused by the effects of the failure pattern. On the other hand, there is a random component, for example due to temporary distraction or carelessness on the part of the driver. Apart from this, however, inadequate driving skills can also lead to contact with cones (lack of experience, incorrect assessment of vehicle dimensions, etc.). However, the evaluation of the experimental measurement should only focus on the first factor, the effects of the failure activation. In order to minimize the probability of Disturbance as a cause of possible pylon contact, in addition to the targeted definition of certain sample characteristics (see section 2.6), the aforementioned exclusion procedure was chosen. In this way, it was possible to exclude those persons who, by leaving the lane, regardless of the failure circuits, indicated an increased probability of Disturbance influences on the controllability measurement compared to the rest of the sample.

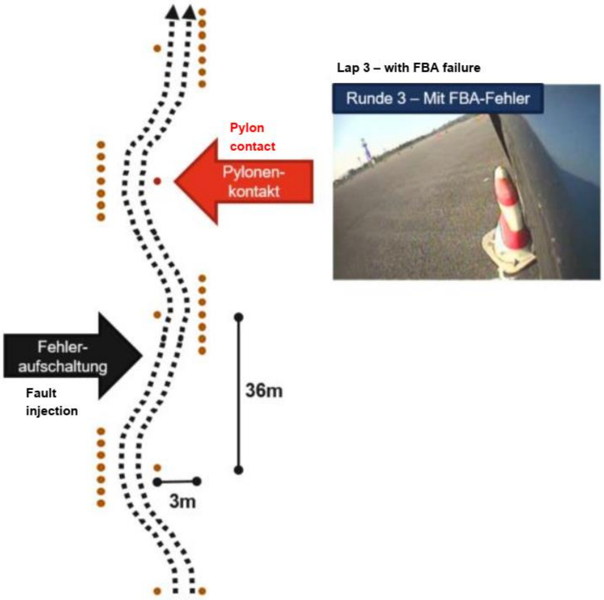
An alternative (or even complementary) procedure at this point could be an expert-based exclusion (or re-inclusion), in which, for example, the test supervisor driving along assesses the driving ability of the test subjects over the entire driving test on the basis of their expertise in order to achieve a practical assessment that is independent of the statistics. However, the disadvantage of such a procedure could be the loss of objectivity in the evaluation.

As summarized above, two failure-vehicle combinations each resulted in lane departure in connection with the failure activation. These are described and discussed below in the light of the methodological assumptions of the subject study.

If we consider the case of leaving the lane in the test vehicle of the sedan vehicle under the influence of the *FBA step* failure (see Illustration 123 / larger illustration in Appendix 6.5), the following picture emerges. The test subject did not leave the lane in the slalom in any of the five laps without a failure activation. It also showed no other conspicuous driving behavior (documented by the camera), neither with regard to the passage of other maneuvers nor with regard to the objective vehicle measurements that were recorded during the relevant failure activation. It is therefore reasonable to assume that this test subject did not leave the lane due to a lack of driving skills during the failure activation. In comparison, the other n = 21 valid test subjects did not record any failure-independent or failure-associated contact with cones in the slalom. If the entire tested sample of the individual study is included in this comparison, three test subjects had failure-independent cone touches. However, these were excluded due to the a priori defined criteria (*n* = 1 incorrect measurement data and too high number of failure-independent torn pylons, *n* = 1 too low speed and too high number of failure-independent torn pylons, *n* = 1 too high number of failure-independent torn pylons). At the level of the subjective data, the test subject classified the failure on the Neukum scale as a 5 (average level of the category "disturbance of driving") and is therefore in the 75th percentile of the comparison sample. It should be noted, however, that the test subject noticed the pylon being driven over in the situation and rated the failure under this impression.

Illustration 123

Graphical representation of the case of lane departure in the sedan under the influence of the failure FBA step.

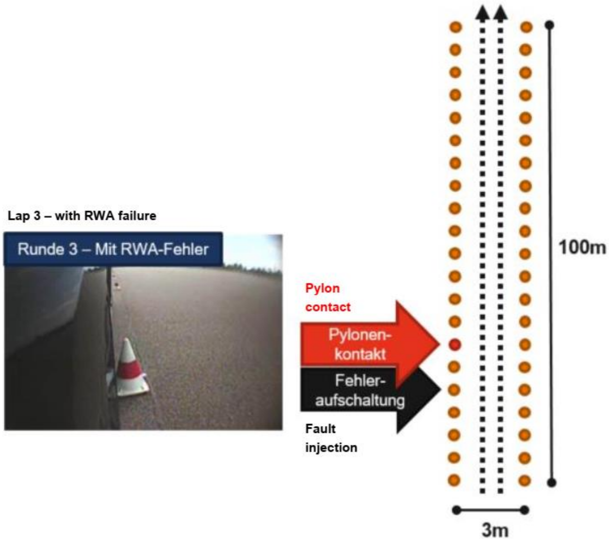


In the second case in the SUV 2 vehicle of the associated vehicle during the *RWA square-wave oscillation* (see Illustration 124 / larger illustration in Appendix 6.6), the test subject in the critical case also only touched a cone in the round with failure activation. In all other rounds, the maneuver was completed without failure, as were the other two maneuvers in all six rounds. At the subjective data level, the test subject classified the failure in the perceptibility range (2) and is therefore in the minimum range of the comparison sample. In the other *n* = 23 valid data sets, there was no failure-dependent or failure-associated lane deviation. As part of the exclusion via criterion 3 (lane keeping), two test subjects were excluded who did not complete the maneuver of driving straight-ahead in one lap without touching a cone, regardless of the failure.

The two cases of lane deviation as a result of failure activation mentioned in the paragraph above basically show that the selected methodology is in principle capable of identifying critical situations as a result of a failure event. In the context of the assumptions of the test series, the hypothesis that experts can locate the parameterization of the failure in the area of the C0 level must therefore be rejected in both cases.

Illustration 124

Graphical representation of the case of lane departure in the SUV 2 vehicle of the associated vehicle under the influence of the RWA square-wave oscillation failure.



Even if the majority of the hypothesis tests speak for a fundamental reliability of the expert judgments, in these cases new tests and readjustments of the failure patterns appear necessary under the framework conditions discussed above. It also remains to be seen to what extent experts can anticipate the controllability of failure events in steer-by-wire steering systems at higher driving speeds by normal drivers, as the lateral dynamic transmission behavior of a motor vehicle changes significantly depending on the vehicle speed.

The transfer of these findings to the actual system design should therefore be critically discussed at this point based on the assumptions of the study series described above. In the context of this discussion, however, a conservative approach seems appropriate in terms of the safety of future users. In addition, the evaluation of the objective overall vehicle reaction, e.g. by recording the metric "lane deviation" as a measure of the absolute failure severity, can serve as an important factor for interpretation. In addition, the continuation of the multi-eye principle applied in the study also appears to make sense when assessing the C0 level of a failure pattern in order to minimize the influence of possible residual subjectivity of an individual expert judgement.

In summary, the methodology developed in the series of tests described above gives experts the opportunity to gain a spotlight-like impression of the driving performance of everyday drivers as well as their subjective and objective handling of and reactions to failure events in steer-by-wire steering systems at limited driving speeds on a test site. Essentially, the results show that, with a few exceptions, experts in the speed range investigated are able to anticipate and assess the controllability of failure patterns by normal drivers in selected driving situations.

5 Bibliography

Brockmann (2009). Code of Practice for the Design and Evaluation of ADAS. PReVENT Report v5.0.

Chancey, E. T., Proaps, A., & Bliss, J. P. (2013, September). The role of trust as a mediator between signaling system reliability and response behaviors. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 57, No. 1, pp. 285-289). Sage CA: Los Angeles, CA: SAGE Publications.

Chavaillaz, A., Wastell, D., & Sauer, J. (2016). System reliability, performance and trust in adaptable automation. Applied Ergonomics, 52, 333-342.

Ghazizadeh, M., Peng, Y., Lee, J.D., and Boyle, L.N. (2012b), "Augmenting the technology acceptance model with trust: Commercial drivers' attitudes towards monitoring and feedback." Proceedings of the Human Factors and Ergonomics Society, 56(1), 2286-2290.

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. *In Proceedings of the human factors and ergonomics society annual meeting* (Vol. 50, No. 9, pp. 904-908). Sage CA: Los Angeles, CA: Sage publications.

Hegner, S.M., Beldad, A.D., and Brunswick, G.J. (2019), "In automatic we trust: Investigating the impact of trust, control, personality, characteristics, and extrinsic and intrinsic motivations on the acceptance of autonomous vehicles." International Journal of Human-Computer Interaction, 35(19), 1769-1780.

IBM Corp. released 2020. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp

Körber, M. (2018). Theoretical considerations and development of a questionnaire to measure trust in automation. In *Congress of the International Ergonomics Association* (pp. 13-30). Springer, Cham.

Neukum, A. & Krüger, H.-P. (2003). Driver reactions to steering system malfunctions - investigation methodology and evaluation criteria. *VDI Reports, 1791*, 297-318.

Numan, J.H. (1998), "Knowledge-based systems as companions: Trust, human computer interaction and complex systems." [Doctoral dissertation, University of Groningen]. http://irs.ub.rug.nl/ppn/169106586

Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., & Tscheligi, M. (2012, October). Predicting information technology usage in the car: towards a car technology acceptance model. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 51-58).

Pavlou, P.A. (2003), "Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model." International Journal of Electronic Commerce, 7(3), 101-134.

Precht, L., Keinath, A., & Krems, J. F. (2017). Effects of driving anger on driver behavior-Results from naturalistic driving data. Transportation research part F: traffic psychology and behavior, 45, 75-92.

6 Appendix

6.1 Preliminary survey

ID: Date:
VL: Time:
To be filled in by the investigator

Demography

V01: Age:

_____ years

V02: Gender:

☐ Female
☐ Male
☐ Miscellaneous
☐ Not specified

V03: Are there any uncorrected visual impairments today (e.g. due to forgotten glasses or color blindness)?

☐ No
☐ Yes, namely: ____

V04: Are there any uncorrected hearing impairments today (e.g. due to a forgotten hearing aid)?

☐ No
☐ Yes, namely: ____

Driving habits

V05: In which year did you obtain your car driver's license (class B)?

Year you got your driver's license: ___________

V06: Please state the make, model and year of manufacture of the vehicle you use most frequently in terms of mileage.

Make: ___________ Model: ___________ Year of manufacture: ___________

V07: How many kilometers did you drive as a driver last year?

___________ km

V08: Please indicate how much practical experience you have with the following driver assistance systems.

Which of the following driver assistance systems for **longitudinal control** do you use?[1] **:**	I use the system...				
	never	**rare**	**occasionally**	**often**	**always**
Cruise control: *The system maintains a speed set by the driver as far as possible. It does not adapt to the traffic, so that the driver has to brake himself, for example in traffic jams.*	☐	☐	☐	☐	☐
Adaptive cruise control (ACC): *The system supplements cruise control by adapting the speed to the traffic ahead, e.g. braking when a slow vehicle is in front.*	☐	☐	☐	☐	☐
Further system: ______________________	☐	☐	☐	☐	☐

Which of the following driver assistance systems for **lateral control** do you use?[2] **:**	I use the system...				
	never	**rare**	**occasionally**	**often**	**always**
Blind-spot warning/blind-spot warning: *The system gives a signal (warning light, sound, etc.) if another vehicle is in its own blind spot.*	☐	☐	☐	☐	☐
Lane-departure warning: *The system gives a signal (warning light, sound, steering wheel vibration, etc.) if the vehicle threatens to exceed the lane limits.*	☐	☐	☐	☐	☐
Lane-keeping assistant: *Extension of the lane-departure warning, in which the system performs a slight intervention if the vehicle threatens to exceed the lane boundary.*	☐	☐	☐	☐	☐
Lane centering: *The system keeps the vehicle in the center of the lane at all times.*	☐	☐	☐	☐	☐
Further system: ______________________	☐	☐	☐	☐	☐

Based on:

[1]: Radke, T. (2013). Energy-optimized longitudinal guidance of motor vehicles through the use of predictive driving strategies. *Karlsruher Schriftenreihe Fahrzeugsystemtechnik*, 19, 30-34. Retrieved from http://dx.doi.org/10.5445/KSP/1000035819

[2]: ADAS Systems. (2018, August). In *Specialty Equipment Market Association*. Retrieved from https://www.sema.org/sema-news/2018/08/adas-systems

6.2 Interim survey - survey dates 1 & 3

Stress

We would like to know how challenging you find completing maneuvers at the moment.

The following questions relate to one maneuver only. Please answer the questions with regard to the maneuver named by the experimenter.

Mental (intellectual) requirements: To what extent does completing the maneuver make mental demands, i.e. thinking, deciding, observing?

very low 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Very high 20
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

B.02: Physical requirements: How much physical activity is required to complete the maneuver?

very low 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Very high 20
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Performance: In your opinion, how successfully did you complete the maneuver?

very low 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Very high 20
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

How challenging do you find the course overall?

Not at all demanding 1	2	3	4	5	6	Very demanding 7
☐	☐	☐	☐	☐	☐	☐

How much do you agree with the following statement: I trust the vehicle?

Do not agree at all 1	2	3	4	5	6	Fully agree 7
☐	☐	☐	☐	☐	☐	☐

6.3 Interim survey - survey dates 2 & 4

Reaction

B.14: What did you feel during the event? Why?

B.15: What did you think during the event? Why?

B.16: What did you do as a result of the event? Why did you do it?

B.17: If the event you have just experienced were to happen again. Would you act differently?

Stress

We would like to know how challenging you found completing the maneuver at this point in time.

B.18: Mental (intellectual) demands: To what extent does completing the maneuver make mental demands, i.e. thinking, deciding, observing?

very low 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Very high 20
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

B.19: Physical requirements: How much physical activity is required to complete the maneuver?

very low 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Very high 20
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

B.20: Performance: In your opinion, how successfully did you complete the maneuver?

very low 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Very high 20
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

B.21: How challenging do you find the course overall?

Not at all demanding 1	2	3	4	5	6	Very demanding 7
☐	☐	☐	☐	☐	☐	☐

B.22: How much do you agree with the following statement: I trust the vehicle?

Do not agree at all 1	2	3	4	5	6	Fully agree 7
☐	☐	☐	☐	☐	☐	☐

Classification

B23: Please categorize the event in the scheme shown below. First select one of the categories on the left and then a number on the right.

Vehicle no longer controllable	**10**
Dangerousness	**9**
	8
	7
Interference with driving	**6**
	5
	4
Perceptibility	**3**
	2
	1
Nothing noticed	**0**

Consent

B.24: Please imagine that the event you have just experienced occurs in a normal driving context. How much do you agree with the following statements?

6.4 Follow-up survey

ID: Date:

VL: Time:

To be filled in by the investigator

	Do not						**ully gree 7**
I would have concerns about reliving this event.	☐	☐	☐	☐	☐	☐	☐
I suspect that I might have an accident as a result of such an event.	☐	☐	☐	☐	☐	☐	☐
The occurrence of this event would be frightening for me.	☐	☐	☐	☐	☐	☐	☐
I would fear that I would not reach my destination because of such an event.	☐	☐	☐	☐	☐	☐	☐
I would fear that I would not understand such an event.	☐	☐	☐	☐	☐	☐	☐
I would be confident that such an event would not have a negative impact on my driving style.	☐	☐	☐	☐	☐	☐	☐
I believe that such an event would be dangerous.	☐	☐	☐	☐	☐	☐	☐
Coping with this event would require increased attention.	☐	☐	☐	☐	☐	☐	☐
The event would distract me from the traffic.	☐	☐	☐	☐	☐	☐	☐
I would feel safe in dealing with the event.	☐	☐	☐	☐	☐	☐	☐
The occurrence of this event would increase the risk of accidents.	☐	☐	☐	☐	☐	☐	☐

Follow-up survey

N01: How much do you agree with the following statements, related to the first mistake you experienced?

	Do not agree at all 1	2	3	4	5	6	Fully agree 7
I would use a vehicle in which this failure could occur.	☐	☐	☐	☐	☐	☐	☐
If such a failure occurs, I would go to the workshop.	☐	☐	☐	☐	☐	☐	☐
I rule out using a vehicle in which this failure could occur.	☐	☐	☐	☐	☐	☐	☐

N02: How much do you agree with the following statements, related to the second failure you experienced?

	Do not agree at all 1	2	3	4	5	6	Fully agree 7
I would use a vehicle in which this failure could occur.	☐	☐	☐	☐	☐	☐	☐
If such a failure occurs, I would go to the workshop.	☐	☐	☐	☐	☐	☐	☐
I rule out using a vehicle in which this failure could occur.	☐	☐	☐	☐	☐	☐	☐

N03a: Which failure do you find more critical?

Failure 1	Failure 2	Both the same
☐	☐	☐

N03b: Please give reasons for your decision.

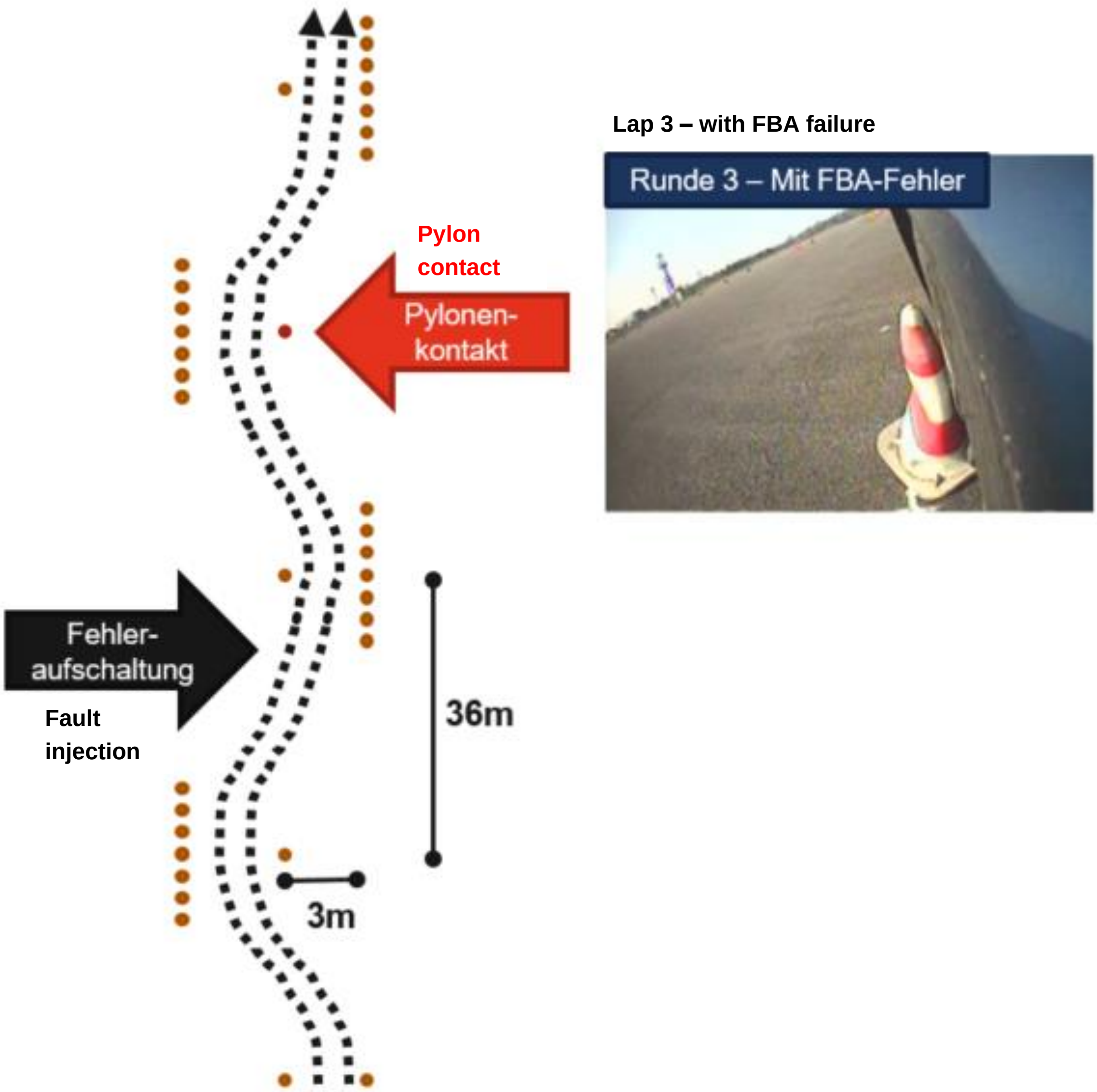
___

___

___

___

6.5 Representation of leaving the track - sedan (FBA step)



6.6 Representation of lane departure – SUV 2 (RWA square-wave)