

Weidner, Florian; Broll, Wolfgang

Stereoscopic 3D dashboards : an investigation of performance, workload, and gaze behavior during take-overs in semi-autonomous driving

Original published in: Personal and ubiquitous computing. - London : Springer. - 26 (2022), 3, p. 697-719.
Original published: 2020-08-25
ISSN: 1617-4917
DOI: [10.1007/s00779-020-01438-8](https://doi.org/10.1007/s00779-020-01438-8)
[Visited: 2023-03-16]



This work is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>



Stereoscopic 3D dashboards

An investigation of performance, workload, and gaze behavior during take-overs in semi-autonomous driving

Florian Weidner¹ · Wolfgang Broll¹

Received: 19 August 2019 / Accepted: 27 July 2020 / Published online: 25 August 2020
© The Author(s) 2020

Abstract

When operating a conditionally automated vehicle, humans occasionally have to take over control. If the driver is out of the loop, a certain amount of time is necessary to gain situation awareness. This work evaluates the potential of stereoscopic 3D (S3D) dashboards for presenting smart S3D take-over-requests (TORs) to support situation assessment. In a driving simulator study with a 4×2 between-within design, we presented 3 smart TORs showing the current traffic situation and a baseline TOR in 2D and S3D to 52 participants doing the n-back task. We further investigate if non-standard locations affect the results. Take-over performance indicates that participants looked at and processed the TORs' visual information and by that, could perform more safe take-overs. S3D warnings in general, as well as warnings appearing at the participants' focus of attention and warnings at the instrument cluster, performed best. We conclude that visual warnings, presented on an S3D dashboard, can be a valid option to support take-over while not increasing workload. We further discuss participants' gaze behavior in the context of visual warnings for automotive user interfaces.

Keywords Automotive user interfaces · Stereoscopic 3D · Driving simulation

1 Introduction

With level 2 automation on the road and higher levels being tested, it is important to look at the moments when humans suddenly have to take over control [62]. These situations require fast, correct, and safe responses to a potentially dangerous traffic situation. To make matters more complicated, it is very likely that humans engage in non-driving-related tasks when automation is enabled and by that, have reduced situation awareness (SA) [24]. Before the human can (or should) react to a take-over request

(TOR), a certain level of SA is necessary [20]. While motor readiness is almost instantaneously available, gaining SA takes a certain amount of time. The longer it takes to gain SA, the higher the risk of an accident [25].

Walch et al. [67] conclude their research with the statement that the majority of their participants simply took over as soon as a TOR was issued — without assessing the situation. Contrary to that, Telpaz et al. [66] mention that their haptic take-over support system did not perform well because drivers did a complete visual scan of their surroundings before acting. Bueno et al. [11] hint at a possible solution to avoid these two extremes: in their research, they conclude that providing information about the environment during a critical event could foster gaining SA. Miller et al. [44] come to a similar conclusion, recommending that sharing information with the driver could support SA.

Vehicle-to-vehicle (V2V) communication, where cars share information with each other — for example positional data [57]—can be a source of such information. Facilitating this knowledge, a vehicle could provide adaptive and holistic visualizations to drivers and support gaining SA. Related to that, recent research indicates that stereoscopic 3D (S3D) displays can improve the comprehension of

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00779-020-01438-8>) contains supplementary material, which is available to authorized users.

✉ Florian Weidner
florian.weidner@tu-ilmenau.de

Wolfgang Broll
wolfgang.broll@tu-ilmenau.de

¹ Virtual Worlds And Digital Games Group, Ilmenau University of Technology, Ilmenau, Germany

traffic scenario visualizations compared with traditional 2D presentations [17], making it an interesting candidate for such adaptive and holistic visualizations.

In this work, we explore smart TORs displayed on an S3D dashboard. “Smart” refers to two properties of our multimodal take-over requests: First, we use information from simulated V2V communication to show the relative location of other vehicles in the form of glyph-like icons. Second, we present the TOR not only at the instrument cluster — the default location for warnings — but also at other, potentially more suitable, locations of the dashboard. By that, we provide evidence for the effectiveness of smart TORs during the take-over process and demonstrate a positive effect of S3D on take-over performance.

Main contributions of this work are:

1. Evidence of the effectiveness of audiovisual smart S3D TORs during the take-over process.
2. Proof that well-designed visual warnings do not necessarily increase mental workload.
3. Three gaze patterns performed by participants when using visual in-vehicle warnings.

This article extends our previous work on smart take-over requests [71]. In addition to the influence of such notifications on take-over performance, this article reports on the design process that led to our visualization, on user experience, and on participants’ performance at the non-driving-related task. User experience and performance in the n-back task further strengthen our results by adding information about the state of participants during the experiment. In detail, it helps to judge their attitude towards the experiment and if they were out of the loop. In addition to that, the details about design process of the visualizations allow researchers to better understand the final take-over visualizations but also support them when creating new designs. We further provide an extensive analysis of participants’ gaze behavior and the new visualizations’ influence on mental workload. Both are important for safety critical features as they tell if the new technology has significant side effects besides potentially influencing take-over performance.

The following section provides related work to S3D, TORs, and mental workload. After that, we outline the research questions (Section 3) and the simulator environment (Section 4). Subsequently, we report on the pre-study that guided the design of the visualization in Section 5. Then, Section 6 describes the design of the driving simulator study. Section 7 presents results of our user study whereas Section 8 puts them in context and discusses them. Section 9 presents the derived gaze patterns, implications, and recommendations based on our results. Finally, Section 10 concludes our work and provides possible future research directions.

2 Background and related work

2.1 Perspective 3D and stereoscopic 3D

Stereoscopic displays provide two images to the human: one for the left eye and one for the right eye. Using special display hardware that ensures that each eye sees its corresponding image, the human visual system is able to fuse both images. This results in a perceived image with binocular depth cues. This type of 3D is called *stereoscopic 3D* (S3D).

Perspective displays operate on traditional screens and show one image for both eyes. Various cues but not binocular cues still make depth perception possible. Such displays often apply linear perspective to create the illusion of depth. Then, the perceived image is in *perspective 3D* (P3D or 2.5D) [47].

2.2 Mental workload while driving

In research, there is no clear definition of mental workload [12, 34, 38, 51]. However, many definitions refer to the information processing capacity necessary for executing a task [13, 59, 73]. Executing tasks that require similar mental resources (e.g., visual processing) in parallel can lead to decreased performance [72]. With regard to driving and take-over, it has been argued that the additional visual information of a visual take-over notification can impair driving performance by requiring visual processing capacity that is then not available for the primary driving task [5]. Hence, it is necessary to consider the influence of the added mental workload introduced by processing the visual TOR.

In the automotive domain, many techniques have been used to assess mental workload—all having advantages and disadvantages. Questionnaires like the Driver Activity Load Index (DALI, [50]) or the NASA TLX (Task Load Index, [26]) are widely used to assess mental workload because they are easy to use. However, these instruments can only be administered after the task or experiment which can lead to, for example, recall bias or post-rationalization bias [41].

To assess mental workload during the task or experiment, researchers have been using various physiological signals [13, 38]. Especially ocular measures promise to measure mental workload directly during a driving simulator experiment and in a non-invasive way [34]. That is because the eyes can rather easily be captured using head-mounted or external eye-trackers and the environment lighting can be controlled [23]. Especially the latter can be a problem and researchers need to control for it [55]. Prominent ocular measures for mental workload are based on the pupil diameter (e.g., index of cognitive activity — ICA [39],

mean pupil diameter change and change rate [49], index of pupillary activity — IPA [18]), on blink behavior (e.g., blink latency [30], blink rate [3]), or on the movement of the eye (e.g., microsaccades [31])

In this study, ocular measures based on pupil diameter and blink behavior are used to assess mental workload as they can be captured in a non-invasive way and have shown to deliver valid results in previous research.

2.3 In-vehicle stereoscopic 3D

Over the last few years, stereoscopic 3D displays are gaining popularity in the automotive domain [16, 54]. Previous work showed that warnings displayed on an S3D instrument cluster can lower reaction times in case of unexpected events [8]. S3D instrument clusters can also shorten glance times at the instrument cluster when identifying objects in non-critical situations [56]. Also, a navigation system in the form of an S3D display on the upper center stack can support drivers better than a presentation in perspective 3D [10]. It has also been shown that S3D can increase perceived urgency, and information structure of instrument clusters when used as a design element [9]. Results of Weidner et al. [70] further indicate that S3D visualizations as notifications and highlighting elements do not impair driving or secondary task performance. In addition to that, Dettmann et al. [17] found that using autostereoscopic 3D displays for the analysis of traffic situations leads to better perception and situational assessment compared with 2D displays. However, they did not perform their experiment in a driving simulator.

The scenarios in these experiments were not time-critical and potentially dangerous as an unscheduled take-over maneuver can be. Further, up to now, no previous study has analyzed the impact of S3D warning visualizations at other locations than the instrument cluster and the upper center stack. Nevertheless, these results motivate the evaluation of S3D take-over notifications that show spatial information.

2.4 (Directional) take-over requests

There is a large body of research on in-vehicle warnings in general and take-over requests in detail. In 2015, Bazilinsky et al. [2] performed a crowd-survey with 1692 respondents. They conclude that visual take-over requests on the dashboard are among the most preferred ones. Hence, many researchers have been using visual take-over requests during studies, either as main research objective or in multimodal user interfaces. Visual take-over requests have been used as icons on the center stack [48, 66], on the instrument panel [35, 36, 42, 43, 46], or on head-up displays (HUD) [29]. These studies used simple icons

to indicate a take-over request (TOR) without providing any additional information that might support situation assessment. Because these signs have been heavily used in research on take-over requests, we use an icon paired with an audio signal as a baseline.

Going beyond simple warning icons, directional TORs provide navigational cues or guide attention and try to support drivers during take-overs. It has been shown that such directional cues can improve take-over performance by means of various modalities. Besides spatial audio [76], especially haptic and visual TORs have been explored. For example, shape-changing steering wheels [6] and vibro-tactile seats [52, 66] have been shown to positively influence workload and take-over performance. However, they were less suited to communicate directions. Also, both approaches, the steering wheel and the vibro-tactile seat, are only able to encode a relatively low level of information, which might not be enough for more complex situations. Nevertheless, the provided directional information improved performance and safety, which acts as a motivation for our study. In the domain of visual directional take-over requests, ambient light cues presented via simple LEDs or LED strips have been shown to improve take-over performance but are also limited regarding information density due to low pixel resolution of LED arrangements [10, 33]. Similar to our approach but in 2D and only on the HUD, Rezvani et al. [61] provided 2D illustrations of the surroundings during a control transfer with the intent to communicate the internal and external awareness of the vehicle to the driver. Their results indicate that such illustrations improve drivers' trust, performance, and SA.

Most of these take-over requests have in common that they only, but often successfully, provide a direction to the driver. This direction highlights either the hazard or a potentially safe route whereas this work tries to communicate information about the location of the hazard as well as the surroundings to improve SA (similar to Rezvani et al. [61]). In addition to that and building on the promising results of in-vehicle S3D, we also want to explore, if potential benefits are reinforced by S3D. It is important to note that most of the mentioned take-over requests are in fact multimodal and accompanied by an auditory signal. Further information on control transitions in (semi-)automated vehicles is provided by Lu et al. [37] and Mirnig et al. [45].

3 Research questions

Based on the related work, the research questions were formulated as follows:

- RQ1 Does a visual TOR that provides information about the surroundings increase take-over performance compared to a simple warning symbol?
- RQ2 Does the location of the TOR visualization on the dashboard influence take-over performance?
- RQ3 Does the presentation of a visual TOR that provides information about the surroundings in S3D lead to better take-over performance than P3D?
- RQ4 Does the added visual information of a visual TOR that provides information about the surroundings increase mental workload during take-over?

RQ1 is based on research that indicates that directional TORs can support the take-over process but research on visual directional TORs is scarce. RQ2 is motivated by the fact that non-directional visual TORs have been applied in several locations such as the instrument cluster or the center stack but no other, possibly more suitable, location. RQ3 is based on the fact that S3D has been shown to be beneficial for understanding of spatial information but has not been evaluated in time-critical take-over situations. RQ4 controls for mental workload as the additional visual information of a visual directional TOR that contains information about surroundings could lead to a resource conflict in information processing.

We first performed a pre-study to create an appropriate design for the directional take-over requests that provides information about the surroundings. Based on the results of this preliminary study, we ran the final study to answer the research questions.

4 Apparatus

Source code of the user interface software and driving simulation are available for free upon request.

4.1 Driving simulation and laboratory environment

The simulation environment is shown in Fig. 1. It consists of a front screen, a rear-screen visible via a rear-view mirror, and a car mock-up. The driving simulation software was built with Unreal Engine 4.18.3¹ (UE4). A steering wheel and pedals are integrated into UE4 using plugins. For this study, using another commercial or free driving simulation software (e.g., STISIM [65]) would have also been possible.

A screen (3.6 m × 2.25 m, 2560 × 1600 pixel at 120 Hz) displayed the driving environment. A rear-view mirror was positioned next to the mock-up and a 1920 × 1080, 30-Hz projector projected the rear-view. Participants were

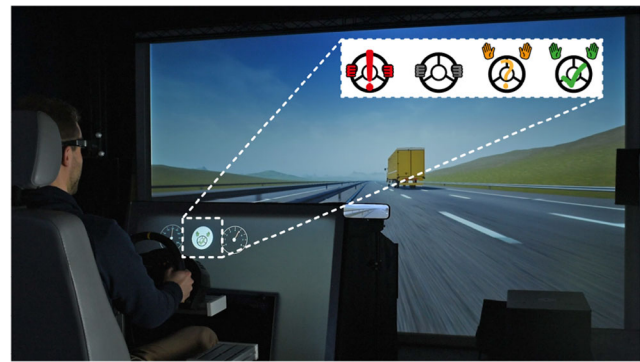


Fig. 1 Driving simulator environment with front screen, car mock-up, and rear-view mirror. Highlighted are the four states of the automation indicator: take-over request (red), automation cannot be enabled (gray), automation disabled but can be turned on (yellow), automation enabled (green)

positioned approximately 2.5 m in front of the screen. This led to an approximate horizontal field of view of 72°. The laboratory has blacked-out windows to control the lighting conditions during the experiment.

4.2 Car mock-up

The car mock-up is shown in Fig. 2. Its design is inspired by the interior of a Mercedes A-class A200 [15]. It provides a large L-shaped S3D dashboard (appr. 90 cm × 60 cm) via rear-projection. The projector provides S3D images with a resolution of 2560 × 1600 pixels at 120 Hz (in S3D mode: 60 Hz per eye). A first-surface mirror redirects the image from the projector onto the projection surface. Stereo vision has been realized using volfoni Edge RF glasses² and a volfoni ActivHub RF50³. For head-tracking, we used an Optitrack Motive system⁴. Primary input devices of the mock-up are a Thrustmaster TX Racing Wheel Leather Edition and pedals⁵.

The UI software responsible for displaying content on the car mock-up consists of two basic components: the tracking system (in our case Optitrack Motive) and the user interface software. The latter was built with Unreal Engine 4.22.3. An off-axis projection is used to correct shear-distortion introduced by stereoscopic 3D rendering. This leads to an improved visualization of virtual objects [1]. Communication between driving simulation and user interface software (c.f. Section 4.2) was realized via local area network (UDP).

²<http://volfoni.com/en/edge-rf/>, 2020-05-16

³<http://volfoni.com/en/activhub-rf50/>, 2020-05-16

⁴<http://optitrack.com/>, 2020-05-16

⁵http://thrustmaster.com/de_DE/produkte/tx-racing-wheel-leather-edition, 2020-05-16

¹<http://unrealengine.com/>, 2020-05-16

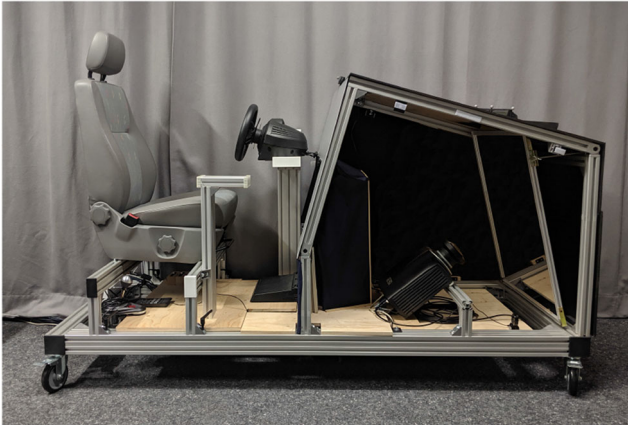


Fig. 2 Side-view of the mock-up showing the projector, first-surface mirror, black-out fabric, and primary control elements

The time it takes for a single frame of motion data from Optitrack to reach Unreal Engine is $L_{Motive} = 5.18 \text{ ms}$ ⁶. In our setup, Unreal Engine 4 has a between-frame latency of $L_{UE4} = 8.3 \text{ ms}$ (120 Hz). These measurements were acquired on a workstation equipped with two Quadro K6000 (one powers UI software and one the driving simulation), an Intel Xeon CPU E5-2670 v3 @ 2.30 GHz, and Gigabit LAN. The projector displays images with 120 Hz per second ($L_{Projector} = 8.3 \text{ ms}$ per image). In S3D, the projector displays the images with 60 Hz per eye (16.6 ms). Together, this results in a latency estimation of $L = 30.08 \text{ ms}$. These calculations do not include additional delays introduced by transmission between components. Hence, the absolute latency L is slightly higher. While motion-to-photon latency needs to be measured with special devices (c.f. Choi et al. [14]), this calculation provides an informed estimation of our system's latency.

4.3 Car and automation capabilities

For our experiment, the car was in automatic mode—no manual shifting was required. We designed the following automation system: The car is able to drive with conditional automation of SAE level 3 on a highway (partial automation, Society of Automotive Engineers, [62]). An icon located between the speedometer and the tachometer indicated the state of the automation system (c.f. Fig. 1). This indicator was gray when the vehicle was outside its operational design domain, orange if automation is off but could be enabled, green when automation was enabled, and red during certain take-over requests. Partial automation could only be activated when the vehicle was on the center lane (to enforce that participants are located

at the correct lane when the TOR will be issued). To turn on automation, participants had to push a green button on the steering wheel. With automation enabled, the car could hold the center lane, maintain distance to leading vehicles, and overtake cars on the right lane. It maintained a speed of 120 km/h which is a common speed on a German autobahn. Automation turned automatically off if participants turned the steering wheel more than 10° in either direction or pushed a pedal more than 10% [25]. Note that the steering wheel was calibrated and centered automatically for each participant using the control panel application of the manufacturer⁷. The influence of 10% steering on the vehicle's velocity is $\Delta_v = -1.420 \text{ km/h}$ per 10 s (SD = 0.014 km/h; averaged over $N = 5$ trials). Initially, the value was set to 2° but our steering wheel has some backlash (or play). Hence, we had to increase the steering wheel angle from 2° to 10° . The operational design domain of the car does not include evasive maneuvers requiring lateral movements. Participants always drove with partial automation enabled except during the take-over maneuvers.

4.4 TOR locations

While driving an automated car, the driver-passenger does not necessarily monitor the road or the instrument cluster. However, most warnings are displayed on the instrument cluster. This might lead to situations where users miss critical information. Hence, we also evaluate locations that might be more suitable during time-critical situations and might better support the driver-passenger in handling such situations. We defined three locations for displaying the directional TOR. They are shown in Fig. 3:

- *TOR-IC*: The area behind the steering wheel at the instrument cluster where warnings are usually displayed.
- *TOR-FoA*: The location on the dashboard where the user's attention is during automated driving (FoA: focus of attention).
- *TOR-Large*: The large area at the center stack. In this condition, all other elements of the dashboard are hidden to mitigate visual distraction.

TOR-IC is the location behind the steering wheel where today's vehicles often show information about advanced driving assistance systems like lane keeping assist or adaptive cruise control. This location has led to promising results in previous studies on multimodal take-over requests that used visual warnings as drivers are used to check there for updates and warnings. The *TOR-FoA* location (FoA: focus of attention) was chosen to assess the potential of future systems that could track the users gaze and display

⁶<https://v21.wiki.optitrack.com/index.php?title=Latency-Measurements> Client Latency, 2020-05-16

⁷<https://support.thrustmaster.com/en/product/txracingwheelleather-en/>, Thrustmaster - Technical support, 2020-07-14

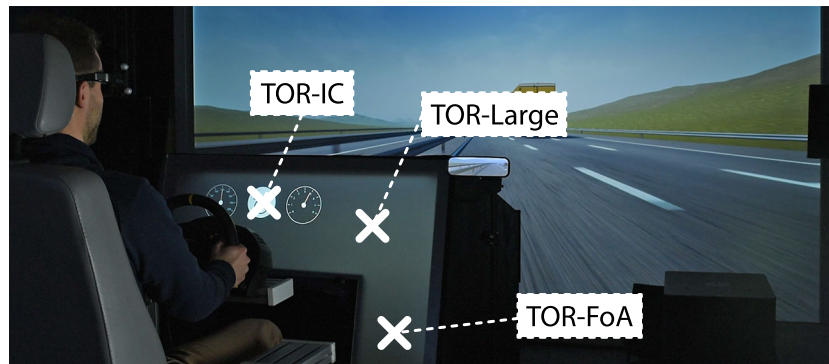


Fig. 3 Locations of the visualization of the take-over requests: TOR-IC at the instrument cluster, TOR-FoA at the focus of attention, TOR-Large at the large upper area of the center stack

time-critical warnings at the current fixation location. In the final study, we used a non-driving-related task (c.f. Section 6.3) to make sure that users had focused on a certain location which we then considered as the focus of attention. With TOR-Large, we want to assess if prominently displaying the take-over request supports the driver. For that, we remove all other visual elements on the dashboard to reduce visual clutter and display the take over request on the large empty area at the upper center stack. The baseline condition uses the same location as TOR-IC but a different visualization: a simple icon in the design language of the automation indicator (c.f. Fig. 3). Here, symbols in today's cars inspired the design. Note, contrary to the other three conditions, the baseline provides no additional information about the surroundings.

5 Pre-study: design of the TOR visualization

Due to the absence of existing literature for visual directional take-over requests, we designed a preliminary TOR visualization for a smart S3D take-over request. The objective was to make an easy-to-understand visual representation of the current traffic situation that can be perceived and processed in a short period of time. It should further indicate potentially dangerous objects. The design is intended for the driving situation of our study where a car in front of the ego car on the center lane performs a full stop and also hides a vehicle that is in front of it (c.f. Section 6.2). We decided to run an informal pre-study with few participants to guide design of the visualization.

For the initial design, we took inspiration from visualizations of navigational systems and visualizations of advanced driving assistance systems (e.g., semi-automated vehicles, lane keeping assist systems, adaptive cruise control) that provide not only status information but also information about the surroundings (e.g., Tesla Model S). The visualizations were 2D graphics and varied in level of detail (c.f. Fig. 4). During the design phase, we realized that distances

had to be modified to only represent relative distances due to the available design space and the zone of comfort [69]. All visualizations presented similar amount of information: the position of the ego vehicle, the position of dangerous actors (red), and the current environment (road etc.). Further, for each pair, one instance presented navigational information in the form of arrows.

In the *abstract TOR* (Fig. 4a), cars are stylized as cubes, the road is a plain surface, and colors indicate the ego vehicle and a potentially dangerous object. The objective

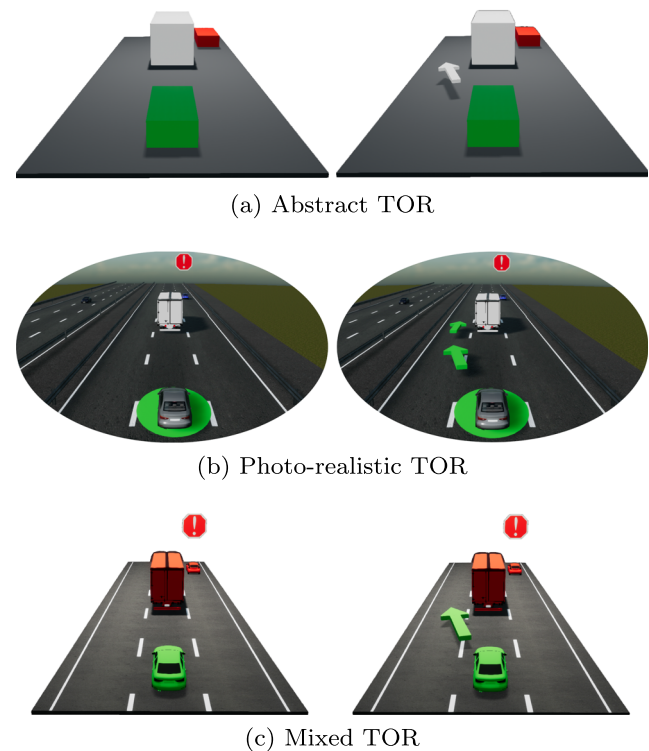


Fig. 4 Three pairs of take-over requests — **a** abstract, **b** photo-realistic, and **c** mixed — presented during pre-study showing the ego car, other vehicles, and a warning sign indicating potentially dangerous objects. Each pair consist of one visualization with arrows indicating a safe direction and one without arrows

of this version was to show a very basic visualization of the surroundings, containing only the most important information (the location of objects and the potential danger they pose). We thought that the absence of many visual elements would make information processing faster and hence, increase reaction time and take-over quality. In the *photo-realistic TOR* (Fig. 4b), colors in the notification match colors of the driving simulation. The guardrail, detailed road markings, and terrain were displayed. The ego-car is a vehicle and a warning sign indicates potentially dangerous objects. Primary intention of this version was to replicate the road scenario as closely as possible to make it easier for the participant to map the provided information to the real world. In the *mixed TOR* (Fig. 4c), colors of cars indicate potential danger (red), the ego car is green (inspired by navigation devices), it shows only the own lane with road markings, and no environment is displayed. In this compromise between *photo-realistic* and *abstract*, we intended to keep the design pseudo-realistic. By that, we hoped to keep the visualization aesthetic without overloading it with unnecessary details.

5.1 Sample and procedure

We invited $N = 5$ (2 female, 3 male) people from university campus (24–36 years old; $M_{age} = 30$ years, $SD = 6.52$). All possessed a valid drivers license and drove their car every day during the week.

The aim of this participatory design study (getting insights on the design of visual warnings for time-critical take-overs) was explained to participants. After that, participants took a seat in the simulator and enabled automation. They were repeatedly encouraged to think-aloud during the drive and to pay attention to the dashboard. In regular intervals, they saw one of the six visualizations, had to take over control of the vehicle, and avoid the obstacles on the road. After each take-over, we paused the simulation and asked participants about their opinion on the visualization. After that, they were told to enable partial automation again. The TOR was only a visual warning without an audio signal. They did not execute a non-driving-related task. All participants experienced all visualizations on all locations in randomized order.

5.2 Results

All participants commented that the visualization showing the surroundings could be a good idea. Three out of 5 participants mentioned that the *abstract* design is too plain, especially the road. One participant was unsure if the gray area depicts only one lane or the complete side of the road. The *photo-realistic* version was perceived as the most visually pleasing one but all participants mentioned that it

might be very overloaded and distracting when there is a crowded road or an environment that contains many objects (e.g., a highway rest area). All participants commented that it is not a good idea to depict the ego vehicle as a car because it was not easily recognizable. Being in a car mock-up, they probably lacked a visual representation of the “car” they are sitting in. Four out of 5 participants preferred the mixed version because it showed them the most important objects with basic information of the environment. Regarding the navigational markers, 4 out of 5 participants stated that they liked the idea, but they were also very skeptic about the accuracy of such recommendations, especially in critical situations. One participant stated that he would “always and ever double check such recommendations - Google Maps makes errors, too.” Regarding the warning sign, participants mentioned that the general idea is good but that it could be displayed more prominently because it is actually the most important item. Four participants also mentioned that the visualization on the center stack (TOR-Large) could be larger to better utilize the available display space.

Key takeaways of the pretest are:

- Recommendations during critical situations require a high level of trust in automation.
- The visualization of the TOR should show only necessary information and basic environmental objects.
- The ego car must be easily identifiable.
- Sources of potential danger must be marked very prominently.

5.3 Final TOR visualization

Based on the pre-study, we redesigned the visualization of the TOR. The final version is shown in Fig. 5 and has the following properties and design elements:

- Other potentially dangerous cars are red.

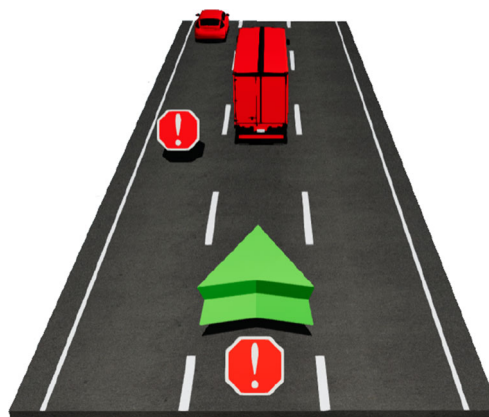


Fig. 5 Final TOR visualization redesigned based on results of the pre-study. It shows the ego car in green; other objects are highlighted in red. Red signs with an exclamation mark indicate locations that require special attention

- The road is shown with lane markings.
- A green arrow indicates the position of the ego car.
- Warning signs indicate dangerous objects or areas.

The green triangular shape was chosen as a representation of the ego vehicle because similar versions are used in today's navigational systems. Other cars are shown in their relative location with the appropriate car type (e.g., truck and sedan). The visualization shows the lane of the ego vehicle with markings. We refrained from arrows that recommend a driving maneuver to prevent potential slow downs introduced by confirming the additional information based on participants comments and also previous results from Endsley et al. [21]. The warning sign was re-positioned so that it is better visible. Another sign was added, if there is potential danger behind the ego vehicle. The baseline condition as well as the final visualization applied to the locations is shown in Fig. 6.

5.4 Stereoscopic 3D configuration

In the S3D condition, inter-pupillary distance was constant across all participants and set to $IPD = 63$ mm (c.f. Howard et al. [27]). Disparity was calculated according to McIntire et al. [40] for a viewing distance of 80 cm. Following participants' feedback from the pretest, the size of TOR-Large was scaled up with a factor of 1.36. The visualizations' sizes and disparities are listed in Table 1.

The user interface was designed so that all objects are displayed within the zone of comfort [69]. Gauges for speed and RPM were displayed with a binocular disparity of 0° . We further provide visual reference points at screen level and displayed the main content of the take-over requests behind the projection plane [7]. Also, we applied a head-coupled perspective to provide the best possible projection and counteract unnatural motion parallax as well as shear-distortion [68]. Simultaneously to the visual TOR, we provide an auditory warning that increases in pitch for 1 s.

6 Study: evaluating the TOR visualizations

6.1 Conditions

The study had a between-within 4×2 design yielding 8 different conditions. Between-subject factor was type of TOR with four levels: TOR-IC, TOR-FoA, TOR-Large, and baseline. Within-subject factor was dimensionality of the presentation with two levels: S3D and P3D. Table 2 shows the conditions.

Our study and the scenario are intended to resemble a hypothetical situation where a leading vehicle on a highway has to perform a full stop due to some (unknown) hazard.

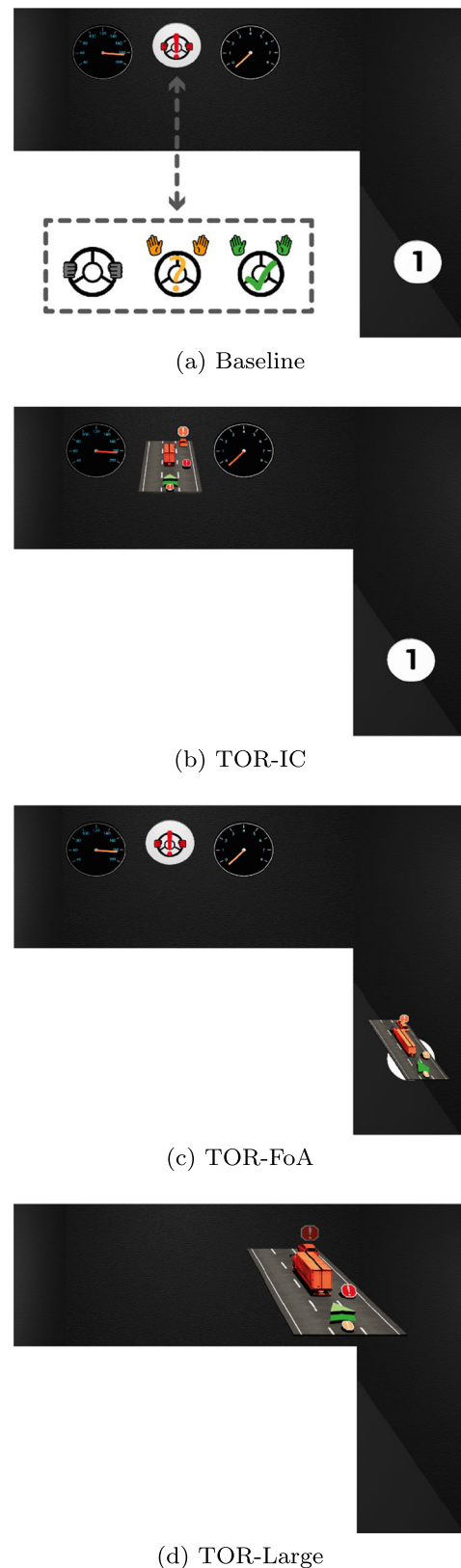


Fig. 6 The four TOR conditions Baseline (a), TOR-IC (b), TOR-FoA (c), and TOR-Large (d) with a tachometer, speedometer (both black), the n-back task (white circle in the lower right area), and the take-over requests. Baseline provides no additional information

Table 1 Disparities D and size of the TORs

	TOR-IC	TOR-FoA	TOR-Large	Baseline	n-back
$D_{NearestPoint}$ (°)	-.005	-.005	-.005	0	.002
$D_{FarestPoint}$ (°)	.016	.016	.024	.002	.003
Width (cm)	13	13	17.7	12	12
Height (cm)	4.5	4.5	6.1	12	12
Depth (cm)	33	33	45	2	1

Negative disparities indicate positions in front of the screen

In this case, the driver has to perform an evasive maneuver. In our case, simply hitting the break is not possible due to a tailing vehicle. Participants were told that the conditional automation system receives data of other vehicles via V2V communication and that this data is integrated into the visualization of the take-over request. We want to assess if these TOR visualizations perform better or worse than traditional take-over requests and if the presentation in S3D leads to improvements compared with P3D while not impeding workload.

6.2 Driving task

Participants started in the center lane of a German autobahn with three lanes in each direction and light traffic (about 5 cars every 1 km). They started driving and turned on the automation system when they felt comfortable. The dashboard showed a speedometer, tachometer, the automation indicator, and the n-back task (c.f. Fig. 6a). After 3 min of driving with partial automation enabled, the car encountered a truck in the center lane and followed it. For this, it decelerated to a speed of 80 km/h and maintained distance for a time between 20 and 35 s. At some time during this interval, the truck performed a full stop. Five seconds before collision with the truck, the car issued a take-over request [10, 25]. With that, the dashboard changed to one of the four TOR conditions shown in Fig. 6. Participants were then required to take over control, evade the truck

Table 2 Conditions of the user study with a 4×2 design (between: TOR, within: dimension)

		Within: Dimension	
		S3D	P3D
Between: TOR	Baseline at instrument cluster	S3D:Baseline	P3D:Baseline
	TOR at instrument cluster	S3D:TOR-IC	P3D:TOR-IC
	TOR at focus of attention	S3D:TOR-FoA	P3D:TOR-FoA
	TOR at center stack	S3D:TOR-Large	P3D:TOR-Large

and also a vehicle that was hidden either on the left or right lane, 25 m in front of it. Figure 7 illustrates the take-over situation. After that, participants turned on automation. The dashboard returned to its initial layout (speedometer, tachometer, automation indicator, n-back task).

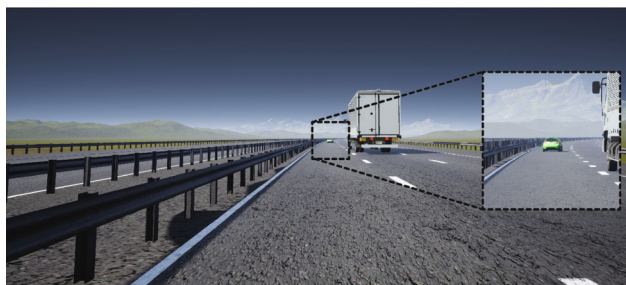
In general, participants could perform a full stop without an evasive maneuver. However, they were instructed to avoid any crashes and in our scenario, a car behind the ego vehicle was close (40 to 60 m) and would have crashed into it during a full stop. The vehicle was visible in the rear-view mirror. The critical event happened four times during each drive. Figure 8 exemplary illustrates a critical event.

6.3 Non-driving-related task

While driving in automated mode, participants were required to perform the n-back task ($n = 2$) [60]. Participants saw a number that changed every 1.5 s to another number between 1 and 10. They had to remember the second last number and hit the enter button of a keypad if they matched. The keypad was positioned on an armrest to their right. The n-back task was chosen because it constantly requires attention. By that, we force the driver to look away from the screen and enforce engagement in a non-driving-related task (similar to reading something on a display in the center



(a) View of the driver when the TOR was issued. The green car is hidden.



(b) A few moments after the TOR was issued. The green car makes one of the two possible evasive maneuvers more dangerous (here, left is more dangerous than right).

Fig. 7 Driving simulation during a TOR: When the truck suddenly brakes, the driver has to evade it (a). A vehicle is hidden either on the left or right side behind the truck (b)

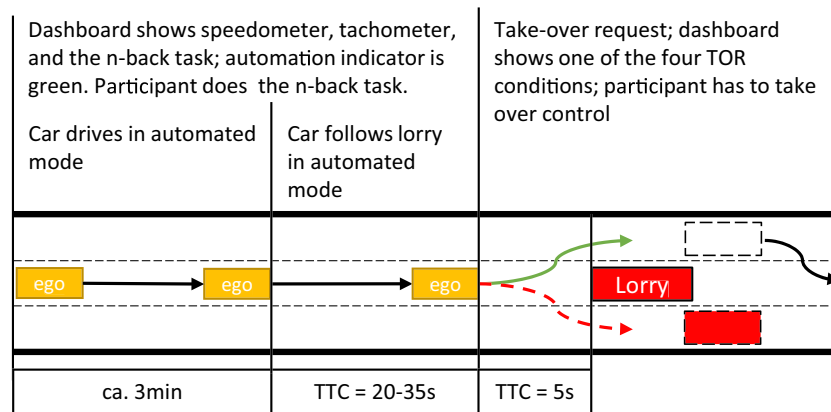


Fig. 8 Exemplary illustration of a take-over event with timings of the three phases. In total, there were four such take-over requests

stack). Participants were instructed to score as high as possible in the n-back task and by that, encouraged to do their best.

6.4 Procedure

When participants arrived, they were asked to sign a consent form. After that, they took a seat in the driving simulator and were handed information about the experiment, simulator, automation capabilities, and take-over requests. Subsequently, they filled out a demographics questionnaire and the Motion Sickness Susceptibility Questionnaire (MSSQ). Each participant had to pass a stereo vision test by pointing out a rectangle on a random dot stereogram [64]. Continuing with the experiment, participants put on the eye tracking and shutter glasses. The experimenter guided the participants through a 5-min training drive to get them accustomed to vehicle physics, non-driving-related task, and take-over requests. After this introductory phase, the first experiment condition started (either P3D or S3D). According to a counterbalanced procedure, participants were assigned to a TOR-group and started either with S3D or P3D visualization. Each of the two drives took about 12 min and participants encountered four events each where they had to take over control. After the first drive, visualization was switched to either P3D or S3D. After each drive, participants had to fill out the User Experience Questionnaire (UEQ) and the Simulator Sickness Questionnaire (SSQ). That concluded the experiment. Approximate experiment duration was 40 min. Participants had the chance to win 50 Euros.

6.5 Measures

6.5.1 Subjective measures

We check for simulator sickness using the Simulator Sickness Questionnaire (SSQ) [28]. User experience is

evaluated with the User Experience Questionnaire (UEQ, [32]). Both questionnaires were provided using a tablet pc with the participant sitting in the simulator.

6.5.2 Glance behavior

We measured glance behavior using an Ergoneers DIKABILIS Professional eyetracker and D-Lab 3.51.⁸ We defined four area of interests (AOIs): one for each TOR, covering the area on the dashboard where it is displayed (TOR-IC and baseline are sharing one area), one area for the road covering the area in front of the participant above the dashboard, and one area covering the rear-view mirror. We extracted the following data:

- Number of glances at the TOR visualizations
- Mean glance duration at the TOR visualizations
- Time to first glance at the road
- Pupil diameter
- Time-series data on glances at the AOIs

We started measuring gaze behavior when participants started automation first and stopped when they reached the end of the track.

We further use scarf plots [4] to analyze the glance behavior of participants regarding the AOIs. By that, it is possible to get insights on how participants used the visualizations of the TORs during the take-over process. We use this data to highlight the gaze behavior of participants and perform a qualitative analysis.

As scarf plots are good for illustrating gaze patterns and their duration, they make comparisons across the sample hard due to their cluttered nature [74]. Hence, we also report participant's fixation transition patterns regarding the AOIs to evaluate the gaze behavior during take-over maneuvers. In addition to that, we analyze the single transitions from one AOI to another. Both methods allow us to investigate

⁸<http://ergoneers.com>, 2020-05-16

differences of the gaze behavior and to understand the effect of the dimension (P3D/S3D) and the TOR visualizations.

6.5.3 Workload

We measured workload of participants directly during the take-over process via ocular measures. This avoids post-rationalization and recall bias. It also ensures that the workload data is directly related to the TOR event and not, for example, to the secondary task or the overall driving task. We refrained from using additional subjective measures (e.g., NASA TLX, DALI) to not lengthen the duration of the experiment.

Mean pupil diameter change. We recorded pupil diameter and calculated the mean pupil diameter change *MPDC* following the procedure of Palinko et al. [49].

$$MPDC = \frac{\sum_{i=0}^{n_{TOR}} (MPD_n - MPD_{total})}{n_{TOR}} \tag{1}$$

with *MPD_n* being the mean pupil diameter per event, *MPD_{total}* being the mean pupil diameter per participant, and *n_{TOR}* representing the number of TORs. A low MPDC indicates low mental workload.

Mean pupil diameter change rate. Calculation of the mean pupil diameter change rate *MPDCR* is based on the procedure of Palinko et al. [49]:

$$MPDCR = \frac{\sum_{i=0}^{n_{TOR}} \frac{\sum_{j=1}^{n_T-1} F'(t)}{n_T-2}}{n_{TOR}} \tag{2}$$

with *t* representing the time, *n_T* the total amount of pupil diameter entries, *n_{TOR}* the number of TORs per participant, $F'(t) = \frac{F(t+h) - F(t-h)}{2h}$, *h* the time between measurement points, and *F(t)* returning the mean pupil diameter *MPD* at a time point *t*. If the MPDCR is positive, the pupil dilated during the event. This can represent increased mental workload and vice versa. According to Palinko et al. [49], this measure is especially suited for detecting changes in cognitive load during time intervals of several seconds.

Mean blink duration change. Similar to MPDC for the pupil diameter, we measure blink duration (MBD) and calculated mean blink duration change *MBDC*.

$$MBDC = \frac{\sum_{i=0}^{n_{TOR}} (MBD_n - MBD_{total})}{n_{TOR}} \tag{3}$$

Again, the smaller the *MBDC*, the lower the average blink duration during the events compared with the average blink duration of the whole drive. A positive MBDC means that the blink duration during events is larger than the duration during the rest of the drive. A decreased blink duration is an indicator for increased mental workload [30, 38].

Blink latency. Blink latency *BL* was measured as the time between the TOR and the first blink. Previous work states that blink inhibition in time-critical events is related to the fact that humans want to perceive as much visual information as possible before blinking. Hence, with increasing visual workload, blink latency has been shown to increase [3, 30, 59].

6.5.4 Driving-related measures

For each participant, we calculated the *number of safe take-overs*. Maneuvers were classified into two classes: correct and incorrect. The maneuver was correct if the evasive action did not go towards the lane with the hidden vehicle and there was no accident or full stop and vice versa. We further calculated the time from the TOR to the first interaction with the steering wheel or pedals that exceeded either 10° in wheel turn angle or 10% of pedal position (motor reaction time).

6.5.5 Non-driving-related task performance

We calculated the *nBackRate* for each participant which indicates how well s/he performed in the non-driving-related task. For this, we calculated the success rate in percent per participant. We counted a success if the participant hit the enter key at the correct number. If s/he missed, pressed the button on the wrong number, or pressed it too late, it counted as a failure. We use this measure to control if participants were out of the loop and focused on the non-driving-related task.

6.6 Sample

The final sample consisted of 52 participants (34 male, 18 female; aged 19–63 years, mean age *M* = 31.9 years, SD = 10.6 years). We used convenience sampling for recruitment via university mailing lists and posts in Facebook groups. All participants had a valid driving license, either normal or corrected vision, and had passed a stereo vision test [64]. Thirty-six had previous experience with stereoscopic displays. Thirty participants had no experience with driving simulators, 10 once, 7 more than once and less than 5 times, and 5 more than 5 times. Mean score of the Motion Sickness Susceptibility score is *M* = 8.45 (SD = 7.06).

7 Results

Data was analyzed using R 3.6.1 (afex 0.24.1, bestNormalize 1.4.0, emmeans 1.4, and fBasics 3042.89). An α -value of 0.05 was used as significance criterion were necessary

(significance codes: ***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$). Outliers were removed if they were above or below 2 times the inter-quartile range [19]. All data was normal distributed unless stated otherwise (tested with Shapiro-Wilk test for normality and QQ-plots [75]). All data showed homoscedasticity according to Levene's tests. Data was analyzed with mixed ANOVAs. For all mixed ANOVAs, no corrections were necessary (all p values for Mauchly's test $p > .05$). Post hoc analysis was performed using Tukey-corrected pairwise comparisons. If not stated otherwise, the data of $52 \times 2 \times 4 = 416$ take-overs was analyzed.

Of the 416 take-over maneuvers, 119 of them were classified as wrong. Fourteen of these wrong maneuvers were due to the participants making a full stop of the highway (not the emergency lane) and by that risking or actually being in a rear-end collision.

7.1 SSQ

A two-way mixed ANOVA yielded no significant main or interaction effects of the factors TOR and dimension ($p > .07$; nausea: $M = 16.05$, $SD = 14.76$, oculomotor: $M = 20.99$, $SD = 18.18$, disorientation: $M = 16.66$, $SD = 21.70$, total: $M = 21.54$, $SD = 18.75$).

7.2 n-back performance: nBackRate

Data on *nBackRate* was not normal distributed. Hence, we applied an ordered quantile normalizing transformation [53]. Overall, participants performed with an *nBackRate* = 81.75% ($SD = 12.91\%$). A mixed ANOVA did not indicate significant main or interaction effects in transformed n-back performance (TOR: $F(3,48) = 0.93$, $p = .434$, $\eta_p^2 = .05$; dimension: $F(1,48) = 1.42$, $p = .239$, $\eta_p^2 = .03$; interaction: $F(3,48) = 0.43$, $p = .731$, $\eta_p^2 = .03$). This means that the groups did not show any differences in performance.

7.3 User experience

Data of the UEQ scales did not show normal distribution. A ordered quantile normalizing transformation restored normal distribution. For the transformed *Attractiveness* scale, there is a significant interaction effect of TOR and dimension, $F(3,48) = 4.30$, $p = .009$, $\eta_p^2 = .21$. Post hoc analysis shows a significant difference between baseline-P3D ($M = 0.86$, $SD = 1.10$) and baseline-S3D ($M = 1.41$, $SD = 0.99$) in favor of the S3D condition, $t(48) = 3.270$, $p = .0386$, Cohen's $d = 0.52$. This result tells us that the baseline-P3D condition was less attractive than the baseline-S3D condition. All other factor combinations are not significantly different in attractiveness.

For the transformed stimulation scale, there was a significant main effect of TOR, $F(3,48) = 5.40$, $p = .003$, $\eta_p^2 = .25$. Post hoc analysis revealed significant differences between baseline and TOR-IC ($t(48) = 3.193$, $p = .0129$, Cohen's $d = 0.997$) as well as baseline and TOR-FoA ($t(48) = 3.299$, $p = 0.010$, Cohen's $d = 0.010$). That means that both TOR-IC ($M = 1.416$, $SD = 0.850$) and TOR-FoA ($M = 1.442$, $SD = 0.983$) were perceived as more exciting and motivating than the baseline ($M = 0.413$, $SD = 1.138$). There were no other significant main or interaction effects.

To put the results in context, we further classified the results of the UEQ according to the benchmark results of Schrepp et al. [63]. Figure 9 illustrates the result of this classification. It can be seen that, according to this classification, there are hardly any differences between S3D and P3D. Between the TOR conditions, TOR-IC and TOR-FoA fall in slightly better categories compared with TOR-Large and baseline.

7.4 Eye measures

We subdivided the total measurement into intervals starting with the TOR and ending 7 s after that. After 7 s, all participants had handled the situation. All participants

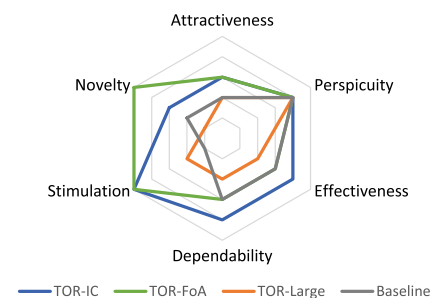
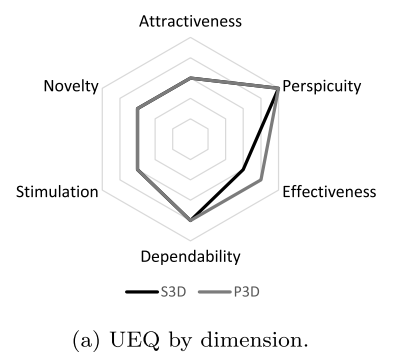


Fig. 9 Classified UEQ results by dimension (a) and TOR condition (b). The larger the better. From inner to outer line: bad - below average - above average - good - excellent

looked at the lower center stack when a TOR was issued and 5 s before that.

7.4.1 Number of glances

There was a statistically significant interaction effect of TOR and dimension on the number of glances at a TOR ($F(3,48) = 3.30, p = .028, \eta_p^2 = .17$) as indicated by Fig. 10. Post hoc tests revealed significant differences between S3D-TOR-IC and all other factor combinations except S3D-TOR-Large, $t(48) > 3.421, p < 0.05$, Cohen’s $d > 0.77$. That means that participants looked significantly more often at S3D-TOR-IC than at any other TOR visualizations except S3D-TOR-Large.

There was also a main effect of TOR ($F(3,48) = 4.88, p = .005, \eta_p^2 = .23$) and also of dimension ($F(1,48) = 5.88, p = .019, \eta_p^2 = .11$). The former indicates that TOR-IC attracts more glances than baseline and TOR-FoA. The latter indicates that P3D leads to fewer glances at any TOR than S3D. However, the interaction effect reflects both.

7.4.2 Mean glance duration

Figure 11 illustrates the mean glance duration for each factor combination. A two-way mixed ANOVA shows a significant main effect of TOR on glance duration, $F(3,48) = 19.30, p < .001, \eta_p^2 = .55$. The baseline TOR shows a significantly shorter mean glance duration than TOR-FoA, $t(48) = -7.394, p < .0001$, Cohen’s $d = 2.55$.

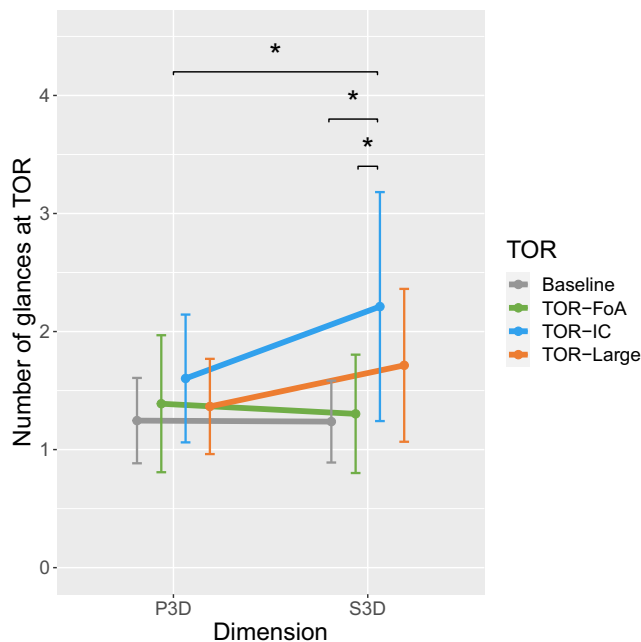


Fig. 10 Interaction effect of dimension and TOR on number of glances at a TOR visualization during a take-over. S3D-TOR-IC is significantly larger than any other value except S3D-TOR-Large

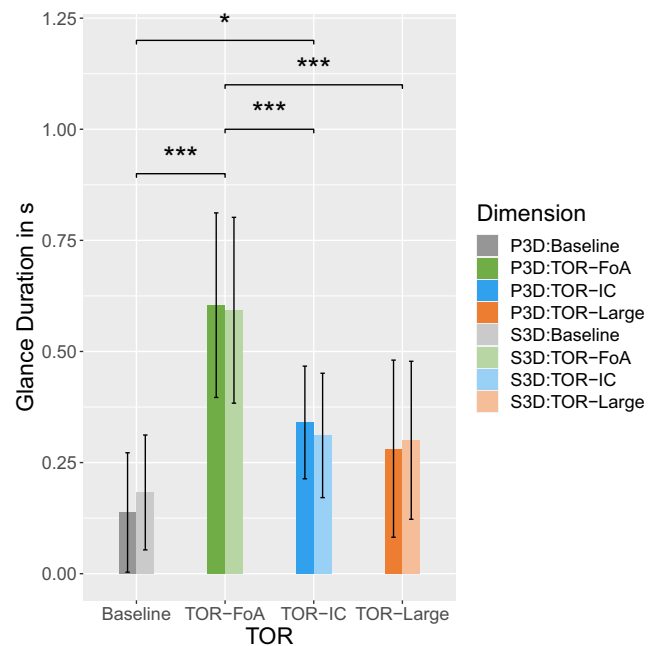


Fig. 11 Summary of mean glance duration at TOR visualizations (mean and standard deviation). In total, participants looked longer at TOR-FoA than at the other TORs. TOR-IC was looked at longer than the baseline TOR

Also, TOR-IC accumulated a significantly shorter mean glance duration than TOR-FoA, $t(48) = -4.603, p = 0.0002$, Cohen’s $d = 1.61$. In the same way, participants spent significantly less time looking at TOR-Large than TOR-FoA, $t(48) = -5.194, p < .0001$, Cohen’s $d = 1.58$. In summary, this suggests that the TOR-FoA was observed longer than all other TORs. In addition to that, the baseline TOR lead to a significantly shorter mean glance duration than TOR-IC, $t(48) = -2.791, p = 0.037$, Cohen’s $d = 1.26$.

Results did not indicate a significant main effect of dimension, $F(1,48) = 0.08, p = .783, \eta_p^2 < .01$. We can assume that neither S3D nor P3D did significantly influence glance duration during our experiment. We do also not observe a significant interaction effect of TOR and dimension, $F(3,48) = 0.61, p = .612, \eta_p^2 = .04$. That tells us that no combination of take-over request and dimension of presentation significantly differs from another one regarding mean glance duration.

7.4.3 Time to first glance at road

Figure 12 illustrates the times it took participants to look at the road for the first time. The type of TOR did have a significant effect, $F(3,48) = 4.06, p = .012, \eta_p^2 = .20$. Results of post hoc tests indicate that, compared with TOR-IC, participants needed less time to look at the road when confronted with the baseline condition, $t(48) = -3.341, p = .0085$, Cohen’s $d = 0.84$. We did not find any other

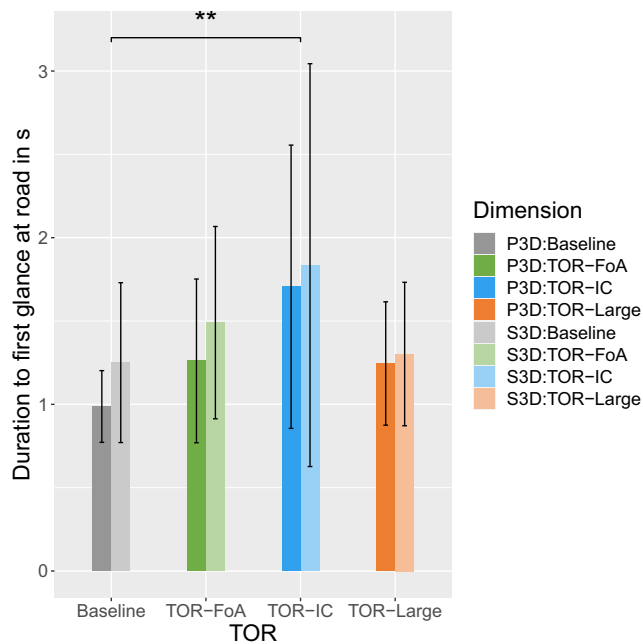


Fig. 12 Summary of eye-tracking data showing the time to the first glance at the road (mean and standard deviation): Compared with baseline, participants looked later at the road when TOR-IC was presented

main or interaction effects. The former tells us that the time to the first glance on the road did not significantly differ between TOR visualizations—regardless of dimensionality. The latter indicates that no factor combination led to significantly higher or lower time it took participants to look at the road.

7.4.4 Glance behavior regarding AOs

Gaze over time. Each scarf plot in Fig. 13 shows the glance behavior of the 52 TOR events for a single TOR. One line represents one single take-over event. For the baseline condition, Fig. 13a and e indicate that the gaze remained on the lower right location for a short time after the TOR was issued. After that, participants mostly looked either at the mirror or to the front. It is noticeable that participants in the P3D-Baseline condition looked more often at the rear-view mirror than participants in the S3D-Baseline condition.

In the TOR-IC condition (c.f. Fig. 13b and f), the gaze quickly shifts to the instrument cluster and the TOR visualization. It then shifts to the front or to the mirror. It is noteworthy that for many participants, there are several switches back to the instrument cluster. This is indicated by many short subsequent blue lines in one row. There are few glances at the rear-view mirror but overall, not as many as in the baseline condition. Opposite to the baseline condition, there is no salient difference between P3D and S3D.

The TOR-FoA visualization (c.f. Fig. 13c and g) attracted very few subsequent glances but rather long initial glances. The figures indicate that the total time participants looked at the TOR is often longer than the time it took participants to look from the n-back task to the IC and then to the road. Also, there were hardly any glances at the rear-view mirror or the IC but participants mostly looked directly at the road. Similarly to TOR-IC, there are no salient differences between S3D and P3D.

Finally, the large TOR visualization (c.f. Fig. 13d and h) often attracted one short glance per participant, only sometimes followed by another glance at the TOR. If such a glance happened, it usually happened close to the first one. There are very few glances at the instrument cluster or the rear-view mirror. Again, no obvious differences between S3D and P3D could be found.

Gaze transitions patterns. Figure 14 shows the distribution of the single gaze patterns or transitions, but without giving information about the duration of a gaze. Here, each single digit (or key) on the x-axis represents a dedicated area of interest. Several digits (or keys) form a pattern and indicate the gaze path across area of interests over time. To analyze possible differences between the gaze transition patterns in the P3D and S3D condition per take-over visualization, we performed Fisher’s exact tests. With them, we investigated if there is any link between dimension of the visualization and the measured gaze transition patterns. For TOR-IC and TOR-FoA, no significant differences could be found ($p > 0.0559$, Fig. 14b and c). That indicates that the distribution of measured gaze patterns is the same in the P3D and S3D condition for these visualizations. Nevertheless, it is noteworthy that the TOR-IC visualization led to almost twice as many gaze patterns as the other three TOR visualizations.

For TOR-Large ($p = 0.0493$, c.f. Fig. 14d) and the baseline visualization ($p = 0.0373$, c.f. Fig. 14a), Fisher’s exact tests indicated significantly different distributions. Salient differences are that, in the Baseline-S3D condition, participants looked more often directly from the n-back task (key 1) at the road (key 5, pattern “15”). In the Baseline-P3D condition, participants looked more often from the n-back task (key 1) to the road (key 5), than on the mirror (key 4) and finally back on the road (pattern “1545”). In the TOR-Large-S3D condition, it is noticeable that users looked more often from the n-back task (key 1) to the TOR-Large visualization (key 3) to the road (pattern “135”). In the TOR-Large-P3D condition, participants often looked back from the road to the TOR visualization (pattern “13535”).

Gaze transition matrices. To analyze gaze transitions in detail, we split the patterns listed in Fig. 14 into pairs. For

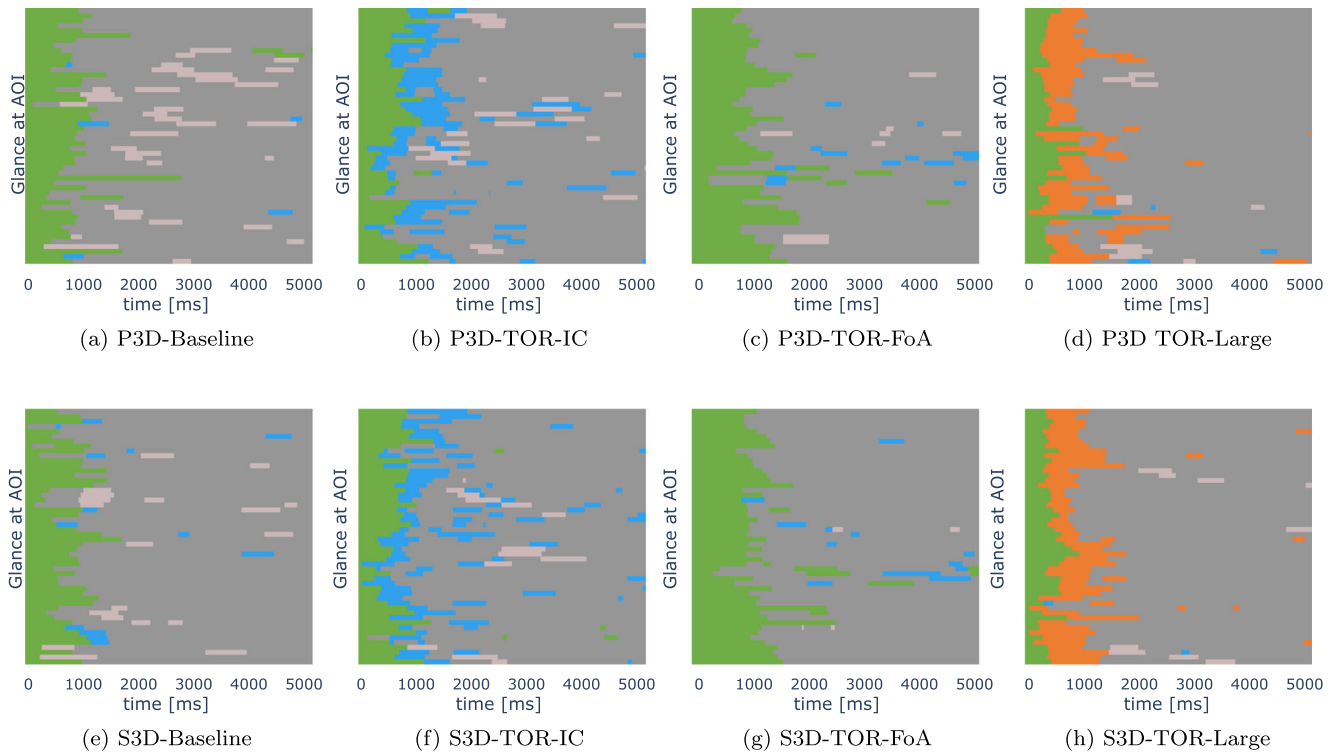


Fig. 13 Scarf plot of glances at TOR visualizations during take-over over a time interval of 5s (Baseline/TOR-IC=blue, TOR-FoA=green, TOR-Large=orange, rear-view mirror=beige, front=grey). Baseline (a, e) shows the default gaze behavior. TOR-IC (b, f) shows many

short glances at the visualization whereas TOR-FoA (c, g) shows few but long glances at the visualization. TOR-Large (d, h) shows relatively few and short glances at the visualization. TOR-FoA (c, g) and TOR-Large (g, h) show few glances at the rear-view mirror

example, “15” is a transition from the n-back task (key 1) to the road (key 5). This resulted in frequency tables telling us how often participants performed transitions. We tested the frequency tables with Fisher’s exact test. Results indicated no significant differences between P3D and S3D for TOR-FoA, TOR-Large, and TOR-IC ($p > 0.1849$). That suggests that dimension did not significantly influence the likelihood of a gaze transition between two AOIs.

For the baseline condition, the Fisher’s exact test indicated significant differences ($p = 0.0290$, c.f. Fig. 15). Salient differences are that participants in the P3D condition looked more often from the mirror (key 4) to the road (key 5) and vice versa. That indicates that participants in the Baseline-S3D condition checked the rear-view mirror more often during the take-over maneuver. Also, participants in the S3D condition looked more often from the instrument

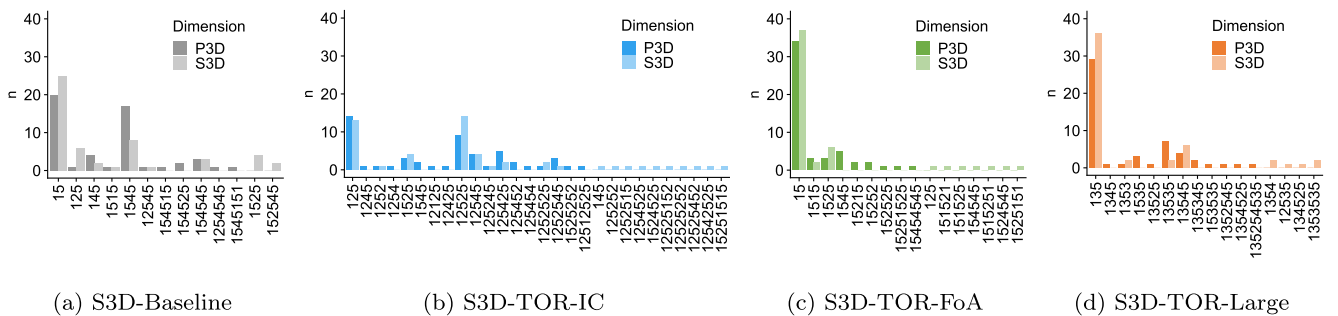


Fig. 14 Gaze transition patterns during the S3D take-over processes for Baseline (a), TOR-IC (b), TOR-FoA (c), and TOR-Large (d) (x-axis: 1 = TOR-FoA, 2 = TOR-IC, 3 = TOR-Large, 4 = rear-view mirror, 5 = front/road; example: “125”: gaze switched from TOR-FoA

to the area covering the TOR-IC to the road). Distributions of P3D and S3D in the conditions *Baseline* (a) and *TOR-Large* (d) differ significantly

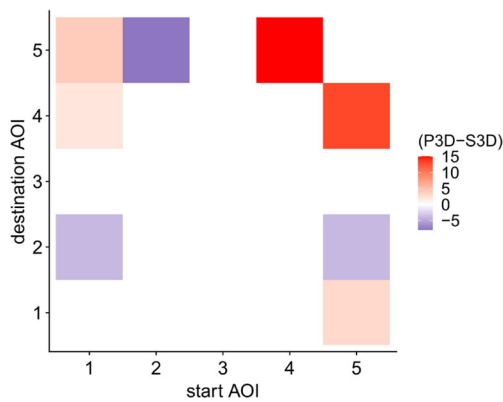


Fig. 15 Difference between gaze transition in the Baseline-P3D condition and the Baseline-S3D condition. Values smaller than zero indicate that the transition was more often measured in the S3D condition (and vice versa)

cluster (key 2) to the road (key 5) and vice versa. That suggests that they were checking if the baseline visualization provides additional information.

7.5 Workload

7.5.1 Blink latency

Blink latency describes the interval between the time the TOR was issued and the time the participant blinked first. Mean blink latency of the untransformed data is ($M = 992.27$ ms, $SE = 55.41$ ms). Data on blink latency was not normal distributed. Hence, we applied an ordered quantile normalizing transformation. Fifteen participants did not blink during the 7-s interval after issuing the TOR. For the analysis, we set their blink latency to the maximum of the observed interval of 7 s. We analyzed data using a mixed ANOVA. We did not uncover significant main or interaction effects, indicating that transformed blink latency did not differ significantly between conditions and factor combinations (TOR: $F(3,48) = 1.66$, $p = .187$, $\eta_p^2 = .09$; dimension: $F(1,48) = 0.18$, $p = .669$, $\eta_p^2 < .01$; interaction: $F(3, 48) = 0.52$, $p = .668$, $\eta_p^2 = .03$). Note, analyzing data without participants who did not blink did also not uncover significant differences. These results suggest that blink latency and in return, mental workload, did not differ between groups in the transformed data.

7.5.2 MBDC: mean blink duration change

We calculated MBDC without the 15 participants who did not blink as there was no logical replacement value. Because of unequal sample sizes, we used type III sum of squares and orthogonal contrasts [22]. A mixed ANOVA did not indicate

a significant main effect of TOR ($F(3,37) = 1.45$, $p = .245$, $\eta_p^2 = .10$) or an interaction effect of TOR visualization and dimension ($F(3, 37) = 0.28$, $p = .841$, $\eta_p^2 = .02$). However, results indicate a significant main effect of dimension on mean blink duration change ($F(1,37) = 4.37$, $p = .043$, $\eta_p^2 = .11$), with $M_{P3D} = -107.09$ ms ($SE = 7.94$ ms) being higher than $M_{S3D} = -114.82$ ms ($SE = 8.49$ ms). That means that in the S3D condition, mean duration of blinks was shorter compared with the P3D condition which indicates a higher workload in S3D.

7.5.3 MPDC: mean pupil diameter change

Data of the mean pupil diameter change was not normal distributed. We applied a ordered quantile normalizing transformation to gain normal distribution. A mixed ANOVA did not uncover any main effects of TOR or dimension on transformed mean pupil diameter change ($M = 6.32$ pixel, $SE = 0.42$ pixel; TOR: $F(3,48) = 1.13$, $p = .347$, $\eta_p^2 = .07$; dimension: $F(1, 48) = 0.00$, $p = .992$, $\eta_p^2 < .01$). Results also do not indicate a significant interaction effect, $F(3,48) = 1.22$, $p = .314$, $\eta_p^2 = .07$. That suggests that mental workload did not differ between groups.

7.5.4 MPDCR: mean pupil diameter change rate

Again, data was not normal distributed and an ordered quantile normalizing transformation was applied before calculating a mixed ANOVA. Analysis of the transformed mean pupil diameter change rate did not uncover a significant difference ($M = 0.011$ pixel, $SD = 0.008$ pixel). Neither the TOR visualization ($F(3,48) = 2.58$, $p = .064$, $\eta_p^2 = .14$) nor dimension ($F(1,48) = 1.30$, $p = .259$, $\eta_p^2 = .03$) lead to significant main effects. There was also no interaction effect between the factor combinations, $F(3, 48) = 0.19$, $p = .902$, $\eta_p^2 = .01$. By that, results suggest that there is no difference in transformed mean pupil diameter change rate and based on that, in mental workload.

7.6 Driving-related measures

7.6.1 Reaction times

There were no significant main or interaction effects on participants' reaction times ($M = 1749$ ms, $SD = 284.8$ ms; TOR: $F(3,48) = 0.89$, $p = .452$, $\eta_p^2 = .05$, dimension: $F(1,48) = 0.03$, $p = .875$, $\eta_p^2 < .01$, interaction: $F(3,48) = 1.77$, $p = .165$, $\eta_p^2 = .10$). By that, we can assume that neither the take-over request nor the dimension of the visualization affected reaction times of our participants.

7.6.2 Safe take overs

Figure 16 shows mean values and standard deviations for the number of safe take-overs per participant for each factor combination. We did not observe a significant interaction between dimension and TOR, $F(3,48) = 2.22$, $p = .098$, $\eta_p^2 = .12$. That means that no combination of TOR and dimension is significantly different to another one with respect to the number of safe take-overs. However, according to the results of a mixed ANOVA, there is a significant main effect of TOR on the number of safe take-overs, $F(3,48) = 7.75$, $p < .001$, $\eta_p^2 = .33$.

Post hoc tests showed that the conditions baseline and TOR-IC differ significantly, $t(48) = -4.384$, $p = .0004$, Cohen’s $d = 1.706$. Also, the groups TOR-FoA and baseline differ significantly, $t(48) = 3.468$, $p = .0059$, Cohen’s $d=1.605$. This suggests that TOR-IC and TOR-FoA lead to significantly more safe maneuvers, regardless of dimension. Also, results indicate a main effect of dimension, $F(1,48) = 4.51$, $p = .039$, $\eta_p^2 = .09$. It tells us that the S3D condition lead to significantly more safe take-overs than the P3D condition ($M_{S3D} = 2.98$, $SD_{S3D} = 0.960$; $M_{P3D} = 2.73$, $SD_{P3D} = 1.06$).

7.7 Learning effect

Each participant performed 8 take-overs — 4 in S3D and 4 in P3D. Hence, we checked if driving-related measures

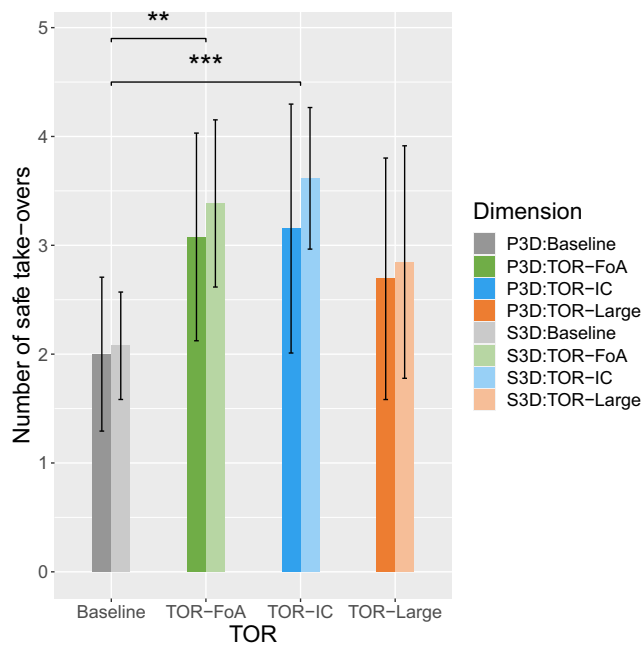


Fig. 16 Average number of safe take-over maneuvers with standard deviation (scale: 0–4; 2 being the value a participant can achieve when always evading either on the left or the right lane or choosing randomly). TOR-FoA and TOR-IC resulted in more safe take-over maneuvers than baseline

or workload changed over time. Here, we analyzed the differences between trials for all participants (8 trials × 52 participants). We also investigated if there are any differences between trials for the subgroups of people who experienced the same take-over visualization (8 trials × 13 participants × 4 TOR visualizations).

For the workload measures, results of a repeated-measures ANOVA with Greenhouse-Geisser correction did not show any significant difference between trials, $F(2.03,24.36) < 2.60$, $p > .094$, $\eta_p^2 < .18$. Similarly, there was also no significant difference in participants’ reaction times (repeated-measures ANOVA with Greenhouse-Geisser correction, $F(5.44, 277.61) < 1.42$, $p > .211$, $\eta_p^2 < .03$) or in the number of safe take-overs (Cochran’s Q-test, $\chi^2(7) = 8.0812$, $p = .3256$).

8 Discussion

We did not find any differences in simulator sickness; hence, we can assume that such differences do not influence our results. The results of the UEQ were relatively equal between groups. While the baseline-S3D condition was perceived more attractive than the baseline-P3D condition, all other factor combinations were not perceived more or less attractive. Compared with the baseline TOR, the conditions TOR-IC and TOR-FoA were perceived more stimulating. Analyzing the categorization in Fig. 9, it is noticeable that the conditions TOR-FoA and TOR-IC often perform better in the classification compared with baseline and TOR-Large. However, differences were largely insignificant. All participants were involved in the n-back task. Hence, we assume that they were sufficiently out of the loop.

8.1 RQ1 and RQ2: added visual information and location

We assumed that supporting situation assessment with TOR visualizations that show information about the surrounding improves situation awareness and results in more safe take-overs maneuvers.

Contrary to our assumption, we found no effect of the large TOR at the center stack on the number of safe take-over maneuvers. We assume that participants did not or not fully process the information provided by the visualization. The time participants looked at it is not significantly different than the baseline. Also, they did not need significantly more time to look at the road for the first time. Overall, there was also no significantly increased number of glances at the TOR. We argue that most of the participants moved the head to the road and tried to process the visualization of TOR-Large in a fly-by motion. The brief first glances and few additional glances in

Fig. 13d and h illustrate this. We believe that this time might not be sufficient to process the information. Additional measurements of the head movements might provide further insights. Also, a different design that is better conceivable in a fly-by motion might improve the performance of this type of TOR.

Participants looked longer at the visualization in TOR-FoA than at all other take-over requests. We think that this is because of two reasons: First, we need to consider that participants were already looking at the location where TOR-FoA appeared when the take-over request was issued: they were doing the n-back task. It is likely that their gaze remained there until they processed the information of the visualization completely and that it then quickly shifted to the road. The low number of glances at the visualization after the TOR was issued, supports this. Figure 13c and g illustrate this behavior. We can only see very few second glances per row and most participants performed only one glance at the visualization of TOR-FoA. All in all (and despite the longer glance duration), participants performed significantly more safe take-overs when the visualization was presented at the focus of attention than in the baseline condition.

In the TOR-IC condition, participants looked significantly more often at the visualization than at all other TORs except TOR-Large-S3D. Being at a dashboard location that is quite close to the road, it is likely that participants shifted their glances between road and the instrument cluster: when the TOR was issued, most participants switched a few times back and forth between road and the instrument cluster which results in the elevated number of glances. The many short glances in Fig. 13a (scattered blue lines) highlight this behavior. The time it took participants to look at the road for the first time is significantly higher than the mean time of the baseline condition. Because mean glance time at the visualization in the TOR-IC condition is not significantly higher than baseline, we assume that the additional gaze movement from the n-back task to the instrument cluster introduced this increase. Nevertheless, the TOR-IC condition lead to significantly more safe take-overs compared with baseline.

While we could not find a positive effect of the large TOR, there was also no detrimental effect in this condition. Confirming with our initial assumptions, the conditions TOR-IC and TOR-FoA lead to significantly more safe takeovers than the baseline condition. In addition to that, reaction time did not change significantly — regardless of take-over request and despite the high number of glances (TOR-IC) or long glance time (TOR-FoA). That is an indicator that TORs with added visual information about the surroundings can improve take-over performance.

8.2 RQ3: S3D and P3D

We assumed that by presenting spatial information in a more natural way and similar to the way humans perceive the real world — in stereoscopic 3D — participants can process the information of the smart TOR better and hence, perform more safe take-over maneuvers.

The dimension of the take-over request did not significantly affect reaction time.

There were some differences in gaze behavior between S3D and P3D. Interestingly, some of them were in the baseline condition where no TOR visualization showing the surroundings was displayed. Here, the gaze transition patterns indicate that participants in the P3D condition included the mirror more often during the take-over (pattern “1545”) whereas S3D lead to more patterns from the n-back task to the road without a mirror check (pattern “15”). Further analysis of this showed that participants in the Baseline-P3D condition looked more at the mirror after having looked at the road (transition “5 – 4” and “4 – 5” in Fig. 15) whereas participants in the Baseline-S3D conditions looked more often at the instrument cluster (transition “5 – 2” and “2 – 5”). Here, it is possible that the S3D display as an advanced display technique increased expectations in the driving automation system and that participants expected to get further helpful information to master the critical situation. P3D users might have refrained to more traditional methods to gather information like checking the rear-view mirrors.

The S3D-TOR-IC lead to significantly more glances than all P3D conditions and S3D-TOR-FoA. One reason for this phenomena might be that the glance strategy with quick glances which shift between road and TOR visualization does not necessarily work with stereoscopic images. Research indicates that the human visual system can need up to 200 ms to fuse stereoscopic images [27]. Short glances below the fusion threshold of an individual might require additional glances to process the information.

For the S3D-TOR-Large condition, the gaze transition patterns also indicate more glances from the road back at the TOR to confirm or gather information shown by the TOR visualization. Reason for these additional confirmatory glances could again be the fusion threshold of 200ms.

Confirming with our assumptions, there was an effect of the take-over requests’ dimension on the number of safe take-over maneuvers. Overall, the S3D condition lead to significantly more safe take-over maneuvers than the P3D condition. By that, our results are evidence that encoding spatial information in S3D can be superior to traditional 2D presentation of take-over requests.

8.3 RQ4: mental workload

We expected that the additional effort necessary to process the visual TORs would increase mental workload. Indeed, there was a significant difference in mean blink duration change *MBDC*: in the S3D condition, participants' *MBDC* was significantly lower compared with the P3D condition. This indicates that, compared with the P3D visualizations, the S3D condition lead to a shorter average blink duration during a TOR compared with the rest of the drive. Thus, this suggests a higher mental workload in the S3D condition. However, Howard et al. [27] mention that blinking rapidly can control and mitigate binocular rivalry. It could be that the sudden appearance of the S3D content and the related necessity for binocular fusion made shorter blinks necessary. It is important to note that there was no main effect of TOR. This would have suggested that the visual TOR and its location might influence mental workload. The other measures — mean pupil diameter change (MPDC), mean pupil diameter change rate (MPDCR), and blink latency — suggest that mental workload does not differ. Especially MPDCR is well known to predict workload over intervals of several seconds [49]. All in all, looking at the absence of any significant differences between baseline TOR and the other TORs in three of our four measures (MPDC, MPDCR, and blink latency), results suggest that the additional information does not significantly increase mental workload. Similar to our research, Petermeijer et al. [52] analyzed different multimodal and directional TORs. They also do not report significant differences.

9 Recommendations and implications

By showing that smart TORs have potential to increase safety while not increasing workload, we can derive some more general recommendations and design guidelines for warnings encoding spatial information.

During a critical situation, it is inherently important to only provide adequate and appropriate visual information. Humans should be able to process the information within a time-frame that does not pose a safety risk (e.g., by deteriorating motor readiness or maneuver performance). Our results argue for a window of opportunity during critical situations where the presentation of such visual warnings to humans is possible. Table 3 shows those values for the three locations of our study.

Those values can act as an initial orientation for designing visual warnings. If perception and processing of a warning is below those values, our results indicate that they do not affect workload, motor readiness, or driving performance at a speed of 80 km/h (the speed of the ego vehicle when the TOR was issued). However, the amount

Table 3 Average glance times in milliseconds by location that did not impair driving-related measures or increased mental workload

Location	Mean	SD
Instrument cluster (IC)	326	132
Lower center stack (FoA)	598	132
Upper center stack (large)	290	185

of information and the warning design need to be carefully considered. We deliberately encoded very few information in the visual warnings, kept the design language very simple, and used basic colors. More information might lead to longer glance times and might increase the risk of neutralizing any benefits or even introduce detrimental effects.

For the design of such warnings, it is important to understand how humans perceive the information. We can characterize three combinations of mean glance duration and number of glances by dashboard location as follows:

- *Switch*: For visual TORs displayed on the instrument cluster, participants showed a back-and-forth gaze pattern. This gaze pattern showed a high number of glances at the TOR visualization (c.f. Figs. 13b and f and 14b)
- *Observe*: On the lower center stack (or focus of attention), we identified a long-glance pattern. This pattern is characterized by a long glance duration at the TOR visualization and a similar number of glances compared with baseline (c.f. Fig. 13c and g).
- *fly-by*: For the upper-right center stack, participants most likely applied a pattern featuring a normal number of glances and a relatively normal glance duration — they did not fixate the TOR visualization very often or that long (c.f. Fig. 13d and h).

These gaze characteristics could guide the design of warning symbols. For example, for the switch-pattern, it is especially important that symbols do not require complex search patterns in order to work with short glances. The observe-pattern requires warnings that appear in a similar region (also depth region) than the previously observed visual element to support visual analysis and to avoid large eye adjustments in accommodation and vergence. For example, the position of the TOR-FoAs visualization confirmed with the location of the n-back task and we tried to keep it at a similar size. Designs for the fly-by pattern should show only very salient and obvious information which can be observed in a short period of time without many fixations. Naturally, for all applications and warnings, the design should be as simple as possible to improve performance and safety.

9.1 Limitations

We used convenience sampling for recruiting which limits the generalizability of results. We also did not ask for driving experience and can therefore not interpret the results within this context.

A major limitation of our study is the design of the traffic scenario. We took inspiration from related experiments that have implemented similar scenarios to evaluate take-over requests. However, the usage of events that are not necessarily occurring in the wild (e.g., suddenly appearing objects or pseudo-randomly stopping vehicles) might influence participants' driving behavior. While no participant commented on the traffic scenario, this lack of realism limits external validity of the study.

Results are also limited by the fact that we only had one critical event and visualization. While the location of the hidden vehicle was randomly assigned, a more diverse set of traffic and critical situations is necessary to draw a holistic picture of visual TORs. For eye-tracking, we used a head-mounted device which got heavy over time, probably influencing simulator sickness scores and by that, the overall results. While it was necessary to use this configuration (a remote eye-tracker would have been problematic due to the S3D glasses), further investigations with a more comfortable eye-tracker are necessary. The take-over requests, while being displayed in a controlled lighting environment, create slightly different brightness values and hence influence pupil diameter. This limits the interpretation and analysis of MPDC and MPDCR. The technical setup, while providing a front and rear-view mirror, lacked side-view mirrors. For further investigations of situation awareness, side-view mirrors should be included. With eye-tracking becoming a more prominent research tool for workload measurements, it is important to be sure that it is an appropriate measure for the experiment design. For our data, blink latency, MPDC, and MPDCR consistently indicated that workload did not increase. Only mean blink duration change lead to different results, suggesting an increased workload for the S3D condition. This argues for a deeper investigation of these measures and if they deliver correct results when used with active stereoscopic 3D technology or if this visualization technique influences blink behavior in a way that prohibits the usage of blink duration.

For the warning sign, we used a sign shape resembling a stop sign. This could have lead to participants intuitively reacting with a full stop (which then was classified as a wrong maneuver). However, the low number of full stops out of all maneuvers (14 out of 416) suggests a rather low impact. Nevertheless, future studies should avoid this design flaw and use a triangular sign shape as used in many warning signs.

10 Conclusion

Inspired and motivated by novel display technologies and previous research, we assessed an S3D dashboard and smart take-over notifications. Different to previous notifications, our system applies a smart approach: it integrates data from simulated vehicle-to-vehicle communication and by that, displays a simplified view of the surroundings with the objective to increase situation awareness and safety during take-over while not increasing mental workload.

To test this approach, we outlined our design process towards an effective design of such TOR visualizations which ultimately lead to an increased take-over performance. Our results further revealed that the location of the take-over notification on the dashboard influences take-over performance: displaying a smart TOR at the instrument cluster (TOR-IC) or at the focus of attention (TOR-FoA) significantly improved take-over performance. Also, our results indicate that displaying such visual TORs in stereoscopic 3D can improve take-over performance. We also showed that such visual warnings do not necessarily impact mental workload during take-over. We further showed that S3D does not necessarily lead to critically different gaze patterns or transitions if designed correctly. Finally, interpreting the acquired eye-tracking data, we identified three distinct patterns on how participants perceive and process the visual warnings on the dashboard.

Aimed at take-over requests, our positive and promising results enable further research on stereoscopic 3D visual warnings that communicate the current traffic situation to the driver. This may also apply for other time-critical warnings that require a certain level of situation awareness.

Future studies might explore the maximum complexity of such visual warnings in similar as well as in other scenarios, e.g., take-overs in inner cities or the involvement of pedestrians. It has also been shown that the type of non-driving-related task influences the take-over process (e.g., Radlmayr et al. [58]). Hence, research on visual TORs and different non-driving-related tasks is necessary. In addition to that, the underlying reasons on why the large TOR did not lead to positive results should be explored. Finally, a comparison of our smart TORs with plain 2D visualizations, AR-HMDs, and AR-HUDs would put our results in context to those alternative warning signs.

Acknowledgments The authors thank the reviewers for their helpful and encouraging feedback.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Funding Information Open Access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arsenault R, Ware C (2004) The importance of stereo and eye-coupled perspective for eye-hand coordination in fish tank VR. *Presence Teleop Virt Environ* 13(5):549–559. <https://doi.org/10.1162/1054746042545300>
- Bazilinskyy P, Petermeijer S, Petrovych V, Dodou D, de Winter J (2018) Take-over requests in highly automated driving: a crowdsourcing survey on auditory, vibrotactile, and visual displays. *Transport Res Part F Traff Psychol Behav* 56(February):82–98. <https://doi.org/10.1016/j.trf.2018.04.001>
- Benedetto S, Pedrotti M, Minin L, Baccino T, Re A, Montanari R (2011) Driver workload and eye blink duration. *Transport Res Part F Traff Psychol Behav* 14(3):199–208. <https://doi.org/10.1016/j.trf.2010.12.001>
- Blascheck T, Kurzhals K, Raschke M, Burch M, Weiskopf D, Ertl T (2014) State-of-the-art of visualization for eye tracking data. *Eurographics Conf Visual (EuroVis)* 1–20
- Borojeni SS, Boll S, Heuten W, Bühlhoff H, Chuang L (2018) Feel the movement: real motion influences responses to take-over requests in highly automated vehicles
- Borojeni SS, Wallbaum T, Heuten W, Boll S (2017) Comparing shape-changing and vibro-tactile steering wheels for take-over requests in highly automated driving. In: *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '17*. ACM Press, New York, pp 221–225. <http://dl.acm.org/citation.cfm?doid=3122986.3123003>
- Broy N, Alt F, Schneegass S, Henze N, Schmidt A (2013) Perceiving layered information on 3D displays using binocular disparity. In: *Proceedings of the 2nd ACM International Symposium on Pervasive Displays - PerDis '13*, p 61. <https://doi.org/10.1145/2491568.2491582>, <http://dl.acm.org/citation.cfm?doid=2491568.2491582>
- Broy N, Alt F, Schneegass S, Pflöging B (2014) 3D displays in cars: exploring the user performance for a stereoscopic instrument cluster. In: *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '14*, pp 1–9. <http://dl.acm.org/citation.cfm?doid=2667317.2667319>
- Broy N, Guo M, Schneegass S, Pflöging B, Alt F (2015) Introducing novel technologies in the car - conducting a real-world study to test 3D dashboards. In: *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15*. ACM Press, New York, pp 179–186. <http://dl.acm.org/citation.cfm?doid=2799250.2799280>
- Broy N, Schneegass S, Guo M, Alt F, Schmidt A (2015) Evaluating stereoscopic 3D for automotive user interfaces in a real-world driving study. In: *Extended Abstracts of the ACM CHI'15 Conference on Human Factors in Computing Systems*, <https://doi.org/10.1145/2702613.2732902>
- Bueno M, Dogan E, Hadj Selem F, Monacelli E, Boverie S, Guillaume A (2016) How different mental workload levels affect the take-over control after automated driving. In: *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp 2040–2045. <http://ieeexplore.ieee.org/document/7795886/>
- Buttme T, Lansdown TC, Walker GH (2019) Mental workload and performance measurements in driving task: a review literature, vol 818. Springer International Publishing, New York. https://link.springer.com/chapter/10.1007%2F978-3-319-96074-6_31
- Charles RL, Nixon J (2019) Measuring mental workload using physiological measures: a systematic review. *Appl Ergon* 74:221–232. <https://doi.org/10.1016/j.apergo.2018.08.028>
- Choi S-W, Lee S, Seo M-W, Kang S-J (2018) Time sequential motion-to-photon latency measurement system for virtual reality head-mounted displays. *Electronics* 7(9):171. <https://doi.org/10.3390/electronics7090171>, <https://www.mdpi.com/2079-9292/7/9/171>
- Classic (2019) MaRS the digital archives of Mercedes-Benz 169 series A-class coupe, 2004 - 2008, <https://mercedes-benz-publicarchive.com/marsClassic/en/instance/ko.xhtml?oid=453317>
- Continental AG (2019) Continental and Leia's new 3D lightfield display bring the third dimension into automotive vehicles, <https://www.continental.com/en/press/press-releases/2019-06-11-3d-instrument-cluster-174836>
- Dettmann A, Bullinger AC (2019) Autostereoscopic displays for in-vehicle applications. In: Bagnara S, Tartaglia R, Albolino S, Alexander T, Fujita Y (eds) *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018), Advances in Intelligent Systems and Computing*, Springer International Publishing, Cham, vol 821, pp 457–466. https://link.springer.com/chapter/10.1007%2F978-3-319-96074-6_48
- Duchowski AT, Krejtz K, Krejtz I, Biele C, Niedzielska A, Kiefer P, Raubal M, Giannopoulos I (2018) The index of pupillary activity. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press, New York, pp 1–13. <http://dl.acm.org/citation.cfm?doid=3173574.3173856>
- Elliott A, Woodward W (2007) *Statistical analysis quick reference guidebook*. SAGE Publications, Inc, 2455 Teller Road, Thousand Oaks California 91320 United States of America. <http://methods.sagepub.com/book/statistical-analysis-quick-reference-guidebook>
- Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. *Human Factors: J Human Fact Ergonom Soc* 37(1):32–64. <https://doi.org/10.1518/001872095779049543>
- Endsley MR (2017) From here to autonomy: lessons learned from human-automation research. *Hum Factors* 59(1):5–27. <https://doi.org/10.1177/0018720816681350>
- Field AP, Miles J, Field Z (2014) *Discovering statistics using R*. Sage, London. <https://uk.sagepub.com/en-gb/eur/discovering-statistics-using-r/book236067>
- Galante F, Bracco F, Chiorri C, Pariota L, Biggero L, Bifulco GN (2018) Validity of mental workload measures in a driving simulation environment. *J Adv Transp* 2018:1–11. <https://doi.org/10.1155/2018/5679151>. <https://www.hindawi.com/journals/jat/2018/5679151/>
- Gold C, Berisha I, Bengler K (2015) Utilization of drivetime - performing non-driving related tasks while driving highly

- automated. In: Proceedings of the Human Factors and Ergonomics Society, <https://doi.org/10.1177/1541931215591360>
25. Gold C, Damböck D, Lorenz L, Bengler K (2013) Take over how long does it take to get the driver back into the loop? Proc Human Fact Ergonom Soc Annual Meeting 57(1):1938–1942. <https://doi.org/10.1177/1541931213571433>
 26. Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock PA, Meshkati N (eds) Human mental workload, advances in psychology, North-Holland, vol 52, pp 139–183. <http://www.sciencedirect.com/science/article/pii/S0166411508623869>
 27. Howard IP, Rogers BJ (2008) Binocular vision and stereopsis. Oxford University Press, New York. <https://doi.org/10.1093/acprof:oso/9780195084764.001.0001>
 28. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG (1993) Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psychol 3(3):203–220. https://doi.org/10.1207/s15327108ijap0303_3
 29. Kim N, Jeong K, Yang M, Oh Y, Kim J (2017) Are you ready to take-over? In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17. ACM Press, pp 1771–1778, <https://doi.org/10.1145/3027063.3053155>
 30. Kramer AF (1990) Physiological metrics of mental workload: a review of recent progress, University of Illinois Urbana-Champaign, <https://apps.dtic.mil/docs/citations/ADA223701>
 31. Krejtz K, Duchowski AT, Niedzielska A, Biele C, Krejtz I (2018) Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. Plos One 13(9):e0203629. <https://doi.org/10.1371/journal.pone.0203629>
 32. Laugwitz B, Held T, Schrepp M (2008) Construction and evaluation of a user experience questionnaire. In: Symposium of the Austrian HCI and Usability Engineering Group. Springer, Berlin, pp 63–76, https://doi.org/10.1007/978-3-540-89350-9_6
 33. Locken A, Muller H, Heuten W, Boll S (2015) An experiment on ambient light patterns to support lane change decisions. In: 2015 IEEE Intelligent Vehicles Symposium (IV), IEEE, pp 505–510. <http://ieeexplore.ieee.org/document/7225735/>
 34. Lohani M, Payne BR, Strayer DL (2019) A review of psychophysiological measures to assess cognitive states in real-world driving. Front Hum Neurosci 13(March):1–27. <https://doi.org/10.3389/fnhum.2019.00057>
 35. Lorenz L, Kerschbaum P, Schumann J (2014) Designing take over scenarios for automated driving. Proc Human Fact Ergonom Soc Annual Meeting 58(1):1681–1685. <https://doi.org/10.1177/1541931214581351>
 36. Louw T, Merat N (2017) Are you in the loop? Using gaze dispersion to understand driver visual attention during vehicle automation. Transport Res Part C Emerg Technol 76:35–50. <https://doi.org/10.1016/j.trc.2017.01.001>
 37. Lu Z, Happee R, Cabral CDD, Kyriakidis M, de Winter JCF (2016) Human factors of transitions in automated driving: a general framework and literature survey. Transport Res Part F Traf Psychol Behav 43:183–198. <https://doi.org/10.1016/j.trf.2016.10.007>
 38. Marquart G, Cabral C, de Winter J (2015) Review of eye-related measures of drivers' mental workload. In: Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015. The Authors, pp 2854–2861. <https://linkinghub.elsevier.com/retrieve/pii/S2351978915007842>
 39. Marshall SP (2003) The index of cognitive activity: measuring cognitive workload. In: Proceedings of the IEEE 7th Conference on Human Factors and Power Plants, IEEE, pp 7–5–7–9. <http://ieeexplore.ieee.org/document/1042860/>
 40. McIntire JP, Havig PR, Pinkus AR (2015) A guide for human factors research with stereoscopic 3D displays. SPIE Defense + Security 9470:94700A. <https://doi.org/10.1117/12.2176997>, <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=2299395>
 41. McKendrick RD, Cherry E (2018) A deeper look at the NASA TLX and where it falls short. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp 44–48, <https://doi.org/10.1177/1541931218621010>
 42. Melcher V, Rauh S, Diederichs F, Widlroither H, Bauer W (2015) Take-over requests for automated driving. Procedia Manufact 3(Ahfe):2867–2873. <https://doi.org/10.1016/j.promfg.2015.07.788>, <https://linkinghub.elsevier.com/retrieve/pii/S2351978915007891>
 43. Meschtscherjakov A, Ju W, Tscheligi M, Szostak D, Krome S, Pflöging B, Ratan R, Politis I, Baltodano S, Miller D (2016) HCI and autonomous vehicles. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16. ACM Press, New York, pp 3542–3549. <http://dl.acm.org/citation.cfm?doid=2851581.2856489>
 44. Miller D, Sun A, Ju W (2014) Situation awareness with different levels of automation. In: Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics, pp 688–693, <https://doi.org/10.1109/SMC.2014.6973989>
 45. Mirmig AG, Gärtner M, Laminger A, Meschtscherjakov A, Trösterer S, McCall R, McGee F, Tscheligi M (2017) Control transition interfaces in semiautonomous vehicles: a categorization framework and literature analysis. In: Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '17)
 46. Mok B, Johns M, Miller D, Ju W (2017) Tunneled. In: Drivers with active secondary tasks need more time to transition from automation. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17, pp 2840–2844. <https://doi.org/10.1145/3025453.3025713>, <http://dl.acm.org/citation.cfm?doid=3025453.3025713>
 47. Naikar N (1998) Perspective displays : a review of human factors. Tech rep. Defence science and technology organisation Canberra (Australia), <https://catalogue.nla.gov.au/Record/2840016>
 48. Naujoks F, Mai C, Neukum A (2014) The effect of urgency of take-over requests during highly automated driving under distraction conditions. In: Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE, pp 2099–2106
 49. Palinko O, Kun AL, Cook Z, Downey A, Lecomte A, Swanson M, Tomaszewski T (2013) Towards augmented reality navigation using affordable technology. In: Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13, pp 238–241. <https://doi.org/10.1145/2516540.2516569>, <http://dl.acm.org/citation.cfm?doid=2516540.2516569>
 50. Pautz A (2008) A method to assess the driver mental workload: the driving activity load index (DALI). IET Intell Transp Syst 2(4):315. <https://doi.org/10.1049/iet-its:20080023>
 51. Paxion J, Galy E, Berthelon C (2014) Mental workload and driving. Front Psychol 5, <https://doi.org/10.3389/fpsyg.2014.01344>
 52. Petermeijer S, Bazilinskyy P, Bengler K, de Winter J (2017) Take-over again: investigating multimodal and directional TORs to get the driver back into the loop. Appl Ergon 62:204–215. <https://doi.org/10.1016/j.apergo.2017.02.023>
 53. Peterson RA, Cavanaugh JE (2019) Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. J Appl Stat <https://doi.org/10.1080/02664763.2019.1630372>
 54. Peugeot Deutschland GmbH (2019) Der neue PEUGEOT 208 <https://www.peugeot.de/modelle/alle-modelle/neuer-peugeot-208.html>

55. Pflęging B, Fekety DK, Schmidt A, Kun AL (2016) A model relating pupil diameter to mental workload and lighting conditions. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, New York, NY, USA, pp 5776–5788, <https://doi.org/10.1145/2858036.2858117>
56. Pitts MJ, Hasedžić E, Skrypchuk L, Attridge A, Williams M (2015) Adding depth: establishing 3D display fundamentals for automotive applications. In: SAE Technical Paper. SAE International. <http://papers.sae.org/2015-01-0147/>
57. Qiu H, Ahmad F, Bai F, Gruteser M, Govindan R (2018) AVR: augmented vehicular reality. In: Proceedings of the 16th Annual International Conference on Mobile Systems, applications, and services. ACM, New York, pp 81–95, <https://doi.org/10.1145/3210240.3210319>
58. Radlmayr J, Gold C, Lorenz L, Farid M, Bengler K (2014) How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving. Proc Human Fact Ergonom Soc Annual Meeting 58(1):2063–2067. <https://doi.org/10.1177/1541931214581434>
59. Recarte MA, Pérez E, Conchillo A, Nunes LM (2008) Mental workload and visual impairment: differences between pupil, blink, and subjective rating. Spanish J Psychol 11(2):374–385. <https://doi.org/10.1017/S1138741600004406>. https://www.cambridge.org/core/product/identifier/S1138741600004406/type/journal_article
60. Reimer B, Gulash C, Mehler B, Foley JP, Arredondo S, Waldmann A (2014) The MIT AgeLab n-back. In: Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '14. ACM Press, New York, pp 1–6. <http://dl.acm.org/citation.cfm?doid=2667239.2667293>
61. Rezvani T, Driggs-Campbell K, Sadigh D, Sastry SS, Seshia SA, Bajcsy R (2016) Towards trustworthy automation: user interfaces that convey internal and external awareness. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), IEEE, pp 682–688. <http://ieeexplore.ieee.org/document/7795627/>
62. SAE International (2015) Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles (SAE J 3016). Tech. rep. SAE International, https://www.sae.org/standards/content/j3016_201806/
63. Schrepp M, Hinderks A, Thomaschewski J (2017) Construction of a benchmark for the user experience questionnaire (UEQ). Int J Interact Multimed Artif Intell 4(4):40. <https://doi.org/10.9781/ijimai.2017.445>, <http://www.ijimai.org/journal/node/1514>
64. Schwartz BL, Krantz JH (2018) Random dot stereograms. <https://isle.hanover.edu/Ch07DepthSize/Ch07RandomDotStereograms.html>
65. Systems Technology Inc (2020) STISIM drive, <https://stisimdrive.com/>
66. Telpaz A, Rhindress B, Zelman I, Tsimhoni O (2015) Haptic seat for automated driving. In: Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15. ACM Press, New York, pp 23–30. <http://dl.acm.org/citation.cfm?doid=2799250.2799267>
67. Walch M, Lange K, Baumann M, Weber M (2015) Autonomous driving: investigating the feasibility of car-driver handover assistance. In: Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp 11–18, <https://doi.org/10.1145/2799250.2799268>
68. Ware C, Arthur K, Booth KS (1993) Fish tank virtual reality. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '93. ACM Press, New York, pp 37–42. <http://portal.acm.org/citation.cfm?doid=169059.169066>
69. Weidner F, Broll W (2017) Establishing design parameters for large stereoscopic 3D dashboards. In: Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct - AutomotiveUI '17. ACM Press, New York, pp 212–216. <http://dl.acm.org/citation.cfm?doid=3131726.3131763>
70. Weidner F, Broll W (2019) Exploring large stereoscopic 3D dashboards for future automotive user interfaces. In: Proceedings of the 9th International Conference on Applied Human Factors and Ergonomics AHFE. Springer, Orlando, pp 502–513, https://doi.org/10.1007/978-3-319-93885-1_45
71. Weidner F, Broll W (2019) Smart S3D TOR: intelligent warnings on large stereoscopic 3D dashboards during take-overs. In: Proceedings of the 8th ACM International Symposium on Pervasive Displays - PerDis '19. ACM Press, New York, pp 1–7. <http://dl.acm.org/citation.cfm?doid=3321335.3324937>
72. Wickens CD (2002) Multiple resources and performance prediction. Theor Issues Ergonom Sci 3(2):159–177. <https://doi.org/10.1080/14639220210123806>
73. Wickens CD, McCarley JS (2012) Applied attention theory. Taylor & Francis, Abingdon
74. Yang C-K, Wacharamanotham C (2018) Alpscarf: augmenting scarf plots for exploring temporal gaze patterns 1–6 <https://doi.org/10.1145/3170427.3188490>
75. Yap BW, Sim CH (2011) Comparisons of various types of normality tests. J Stat Comput Simul 81(12):2141–2155. <https://doi.org/10.1080/00949655.2010.520163>
76. Zarife R (2014) Integrative warning concept for multiple driver assistance systems, Ph.D. Thesis, Julius-Maximilians-Universität Würzburg, <https://nbn-resolving.org/urn:nbn:de:bvb:20-opus-101118>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.