

Effects of 3D Displays on Mental Workload, Situation Awareness, Trust, and Performance Assessment in Automated Vehicles

MARK COLLEY, Institute of Media Informatics, Ulm University, Germany

ANNIKA STAMPF, Institute of Media Informatics, Ulm University, Germany

WILLIAM FISCHER, Institute of Media Informatics, Ulm University, Germany

ENRICO RUKZIO, Institute of Media Informatics, Ulm University, Germany

Novel display technologies, such as lightfield displays, have become increasingly available. In the automotive domain, these are already in use or are to be used in the future. However, their effect on the user is yet to be explored. Therefore, we conducted a within-subject study (N=15) comparing a baseline visualization of information about automated vehicle functionality with it being visualized either via a LumePad or a LookingGlass display. Interestingly, we found almost no significant differences, thus, indicating that the display technology is less relevant for conveying automated functionality.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; *Empirical studies in visualization*.

Additional Key Words and Phrases: automated vehicles, 3D display, trust, holographic

ACM Reference Format:

Mark Colley, Annika Stampf, William Fischer, and Enrico Rukzio. 2023. Effects of 3D Displays on Mental Workload, Situation Awareness, Trust, and Performance Assessment in Automated Vehicles. In *International Conference on Mobile and Ubiquitous Multimedia (MUM '23)*, December 3–6, 2023, Vienna, Austria. ACM, New York, NY, USA, 16 pages. <https://doi.org/10.1145/3626705.3627786>

1 INTRODUCTION

Novel display technologies focus on improved color-gamut performance [6], reduced power consumption [22], better contrast ratio [50], display thickness, size, and visualization capabilities, including 3D display technology [31]. While the first 3D displays are commercially available (e.g., [21, 28, 37]), there has not been widespread adoption. For example, the 3D display market in the USA in 2021 is estimated to be around 162 million \$ [47] while the general display market in the USA in 2021 is approximately 43.4 billion \$ [48], resulting in a share of 0.37% for 3D displays. Displays are used to visualize various information, including spatially distributed data such as maps. 3D displays could provide unique benefits for visualizing such spatially distributed data.

Previous work using 3D displays has shown benefits for surgeries [5, 53], air traffic control [35], spatial ability [42], and target search [54]. In the automotive context, Broy et al. showed that 3D displays do not increase mental workload compared to 2D displays [4] and can even improve task completion times and interface attractiveness [3]. While these task-focused measurements were important in the era of manual driving, the focus of relevant dependent variables shifts with fully automated vehicles (AVs).

With AVs, users will not need to take over the vehicle. However, it is crucial that the technology is accepted and trusted for AVs to be able to positively impact mobility. Previous studies have shown overtrust or distrust in this technology. Both overtrust or distrust could lead to an undesired adoption of automation. Therefore, calibrated trust [41] should be

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2023 Copyright held by the owner/author(s).

Manuscript submitted to ACM

the main goal [11, 15]. Previous work has shown that visualizing uncertainty can lead to a better understanding of AV capabilities [11, 15]. These visualizations were either tested in (simulated) 2D displays using virtual reality (VR) [10], or simulated augmented reality windshield displays in remote online monitor-based studies [11, 15]. However, current high-class vehicles such as the Mercedes Benz S-Class already integrate 3D displays¹. With advanced displays, the so-called Halo Effect [43], i.e., attributing capabilities based on unrelated characteristics, could come into play. This means that these displays, with their “shiny” effects, could lead to higher trust in the AV despite not contributing to the actual relevant capabilities. Previous research has explored the Halo Effect [43] in automotive user interface design and found correlations with perceived vehicle capability and trust [9, 15, 23]. However, some results were inconsequential [10]. One way trust and other capability assessments are influenced, according to the Halo Effect [43], could be via unrelated characteristics. In the automotive domain, this could be, for example, the design of the chassis but also the visualization technology within the AV.

To investigate this phenomenon, we conducted a within-subject study with N=15 participants, evaluating the effect of three displays on mental workload, situation awareness, trust, and performance assessment (e.g., recognizing other vehicles or pedestrians) in AVs. The displays included (1) a baseline, where the LumePad acts as a standard display, and lightfield technology as implemented by the (2) LumePad 1 or (3) the LookingGlass LKG-2K-05780 (LookingGlass 8.9 Gen1) display. We visualized the prediction for other vehicles and the planning of the own trajectory in Unity on three different display types. We hypothesized that the display type has an effect on mental workload, situation awareness, trust, and performance assessment based on the Halo Effect [43]. This work is only possible with actual 3D displays and cannot be done, for example, in VR due to missing 3D display implementations for this device type. These displays have only recently become widely available. Interestingly, we only found a significant difference in assessing whether AV recognized other vehicles. We employed Bayes factor analysis to test whether these non-significant differences actually stem from a lack of difference between the conditions. We found moderate evidence for this regarding trust and the AV’s capability to drive.

Contribution Statement: Our main contributions are (1) the definition and implementation of visualizations regarding vehicle and pedestrian recognition and intention on two 3D displays and (2) the results of a user study with N=15 participants regarding the effects of 3D displays on mental workload, situation awareness, trust, and performance assessment in AVs.

2 RELATED WORK

This work builds on prior work on general 3D display technology and its advantages, especially with a focus on automotive appliances. Additionally, we report the results of previous work on general information visualization in the automotive context.

2.1 Information Visualization in Automated Vehicles

While the display technology is particularly relevant in this work, the displayed information is also relevant. Therefore, we introduce prior work on information visualization with a particular focus on mental workload, trust, and situation awareness.

Previous research has evaluated various methods for communicating decisions, detections, destinations, regulations, and navigation information to users in the context of AVs. For example, Löcken et al. [36] proposed using ambient light

¹<https://www.caranddriver.com/news/a33239124/2021-mercedes-s-class-mbox-3d-display/>; Accessed: 22.06.2023

to inform users of AV decisions, and Wilbrink et al. [59] suggested using light strips to indicate the AV's intentions or perceptions. Lindemann et al. [34] employed an augmented reality (AR) heads-up display (HUD) to highlight potential hazards such as pedestrians, which they found to result in higher situation awareness in low and high visibility scenarios than only having the basic elements of speed and navigation information.

The concept of "calibrated trust" [41] refers to a state where the user's trust in an automated system is appropriate to the system's capabilities, which can help prevent associated issues with over- and under-trust.

Koo et al. [32] investigated the effect of providing "how" and "why" information on trust in semi-autonomous vehicles and found that explanatory information led to the highest trust, although providing information on the vehicle's behavior could lead to cognitive overload [32]. However, combining both types of information resulted in the safest driving behavior. Häuslschmid et al. [25] evaluated the impact of visualizing the vehicle's current situation interpretation using a world in miniature or a simulated chauffeur avatar and found that the world in miniature led to the greatest increase in trust, although participants had varying opinions on whether such a visualization was necessary. In VR, Colley et al. [10] compared the cognitive load of a tablet-based with an AR version to display pedestrian intention and found that the tablet version was rated significantly higher in cognitive load. Currano et al. [16] also tested an AR HUD and found that the results varied depending on the dynamic nature of the scene and the reported driving style of the participants, and therefore concluded that the HUD should be adaptable to these factors. Schneider et al. [49] evaluated explanations given via an AR HUD and an LED strip and found that user experience increased with the explanations. However, the combination with a post-explanation via a smartphone app did not increase the user experience. Finally, Colley et al. [13] evaluated an abstract representation of the perceived objects and showed that no futuristic visualization is necessary but that already current HUDs could provide appropriate information.

Recent research by Colley et al. [11] examined the impact of visualizations of the semantic segmentation task. The authors argued that prior studies in this domain have relied on abstract representations that do not facilitate the user's ability to comprehend the sources of uncertainty. Through their study, they discovered that the simulated AR Windshield Display (AR WSD) did not have any significant effect on trust or cognitive load. The subjective measure of situation awareness showed an improvement, and the users rated the recognition-related attributes to be better.

Finally, Colley et al. [15] evaluated the visualization of scene detection, prediction, and maneuver planning. This combined previous work [10, 11] and added visualization of predicting other vehicles' trajectory and the own maneuver via trajectories. The study evaluated the impact on trust, mental workload, situation awareness, and perceived safety. The results showed that visualizations related to Situation Prediction were perceived as inferior to the other levels.

However, what is currently missing is the investigation of whether such visualizations on Situation Understanding and Prediction could benefit from novel display types. We thus evaluate how display technologies affect trust, mental workload and situation awareness. We visualize other road users and their and one's own intention to indicate what the AV recognizes, predicts, and plans (see Figure 3b).

2.2 Effects of 3D Displays in the Automotive Context

3D displays can be categorized based on the technologies of lightfield display, integral imaging, and volumetric 3D or holographic 3D displays [46]. These differ in various performance metrics such as color presentation, parallax effects, and interaction possibilities (for a review, see Ren et al. [46]).

3D displays were investigated in various contexts including surgeries [5, 53], air traffic control [35], spatial ability [42], target search [54], and navigation [7]. In their survey, McIntire et al. [39] found that "41 (58%) of the experiments

showed 3D to be better than 2D regarding the performance measurements of interest; ten (14%) showed mixed results; and 20 (28%) of the experiments showed no benefit to using 3D” [39, p. 4].

The use of 3D dashboards in the automotive industry has been the subject of considerable research to assess their potential benefits and drawbacks. While it has been established that excessive parallaxes in 3D displays can cause discomfort to users, as reported by Broy et al. [2] in their study, they do not appear to increase visual load or driver distraction [1]. Weidner and Broll [55] evaluated the zone of comfort of stereoscopic 3D dashboards for vehicles, finding a “budget ranging from 23.8 ± 2.9 cm in front of the dashboard to 22.7 ± 5.5 cm behind it” [55, p. 1].

In a driving simulation study, Weidner and Broll [56] instructed participants to perform two simplified cognitive tasks: a change detection task and a list selection task. The study’s results indicate that the presentation of stereoscopic 3D content does not have a statistically significant impact on the performance of secondary cognitive tasks or driving performance, as determined by analysis of the entire dashboard and its individual subareas.

However, there were also positive effects. Studies such as that conducted by Pitts et al. [45] have demonstrated that 3D displays can enable shorter glance times and faster object recognition for users. Furthermore, in the context of partially autonomous vehicles, where drivers must take over control of the vehicle occasionally, quickly gaining awareness of the situation is crucial. In such scenarios, stereoscopic 3D dashboards have been shown to be beneficial for situation assessment and performance, as demonstrated by Weidner and Broll [57, 58].

Broy [1] have conducted extensive research on 3D displays in automotive contexts. Their studies have aimed to evaluate the potential advantages and disadvantages of these displays and have investigated the discomfort that they can induce in users, as well as the risks they may pose in in-car environments. Furthermore, in their study [4], they found that there was no significant difference in workload between 2D and 3D displays and that the latter did not have a negative impact on the driving scenario. Additionally, they showed that 3D displays could increase accuracy for expected events, decrease task completion times, and make the interface more attractive [3]. In take-over scenarios, 3D displays also have been shown to improve driver re-engagement [1]. However, Broy et al. [4] only visualized an abstract representation of a street and upcoming events such as navigation cues, traffic signs, and traffic information (such as traffic jams) but not spatially more complex information such as maneuver planning trajectories what we look at in our study.

Another effect is the so-called Halo Effect [43], i.e., attributing capabilities based on other (unrelated) characteristics, was shown in the general HCI context [8, 30]. However, literature in the transportation and, more specifically, in the automotive domain, is inconsequential.

Frison et al. [23] have shown that how a vehicle agent interacts with its users influences their expectations of the vehicle’s capabilities. Colley et al. [9] also could show that for AVs communicating and answering pedestrians were perceived as more intelligent and evoked higher trust. Work that also visualized different parts of the inner workings of AVs such as scene detection, situation prediction, and maneuver planning [15] also discusses whether the relatively error-prone assessment of intentions led to the significantly lower ratings of lateral control, thereby representing a Halo Effect.

Despite these positive showings of the Halo Effect, Colley et al. [10] report that while the trust was increased by the AR visualization, the assessment of longitudinal and lateral control was not increased. While lateral and longitudinal control remained on the same level, Colley et al. [13] showed that any visualization led to better critical situation perception significantly higher perceived intelligence and intention of developers. Finally, in urban air mobility, Colley, Meinhardt et al. [14] also evaluated different visualization with the assumption that better visual aesthetics could influence trust in urban air mobility. However, they do not discuss this effect further.

Based on these results, we were interested in whether the display technology would lead to altered trust and capability assessments.

3 EXPERIMENT

3.1 Research Question and Hypotheses

The following research question (RQ) guided our study:

RQ *How do different display technologies influence (1) mental workload, (2) trust, (3) situation awareness, and (4) the attributed capabilities of the AV?*

Based on previous work on information visualizations, 3D displays, and the Halo Effect [43], we hypothesized that the display technology could lead to higher trust, situation awareness, and attributed capabilities and lower mental workload.

3.2 Materials



Fig. 1. A participant sitting in front of the study setup with the LumePad as an onboard computer (a).

We used the Windridge City asset for Unity and recreated the ego trajectory and the predicted trajectories of the other vehicles by Colley et al. [15]. Unlike in previous work, these visualizations were shown not in a simulated AR WSD but on an additional display (see Figure 1). While this was shown to increase mental workload [10], this arrangement is both currently feasible to study and could be implemented in vehicles easily. The additional display did not show the entire scene but a reduced visualization only including the road, its users, pedestrians, and predicted vehicle trajectories both for the own vehicle and others (see Figure 3b). The detection of pedestrians and vehicles was included implicitly as the additional display can only show what the simulated AV has detected [11]. For the different displays, which are shown in Figure 2, we used the official SDKs.

The driving scene was shown to the participants on a Samsung QLED 4K Q60A 75" TV. The distance was chosen to be both appropriate for immersing in the scene and not noticing pixels.



Fig. 2. The two devices used in the study up close. The distortion in the figure stems from the 3D effect.

The used 3D displays are both lightfield displays but differ in some characteristics. The Lumepad 1 is a 10.8-inch tablet. As a front sensor array tracks a user's face, it is only viewable from one angle at a time. The Looking Glass Gen1 has a size of 8.9 Inches with a viewing angle of 50° viewable from different angles simultaneously.

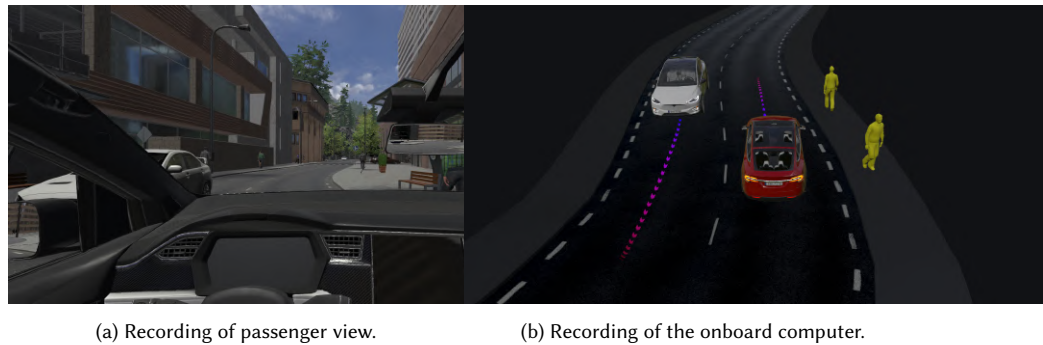


Fig. 3. The participants' scene view (left) and the onboard computer view (right).

3.3 Measurements

Based on our research question, we employed the mental workload subscale of the raw NASA-TLX [24]. This uses a 20-point scale ("How much mental and perceptual activity was required? Was the task easy or demanding, simple

or complex?"; 1=Very Low to 20=Very High). Furthermore, we used the subscales *Predictability/Understandability* (*Understandability* from here) and *Trust of the Trust in Automation* questionnaire by Körber [33]. This questionnaire measures Understandability via agreement on four statements ("The system state was always clear to me.", "I was able to understand why things happened."; two inverse: "The system reacts unpredictably.", "It's difficult to identify what the system will do next.") using 5-point Likert scales (1=*Strongly disagree* to 5=*Strongly agree*). Trust is measured via agreement on equal 5-point Likert scales on two statements ("I trust the system." and "I can rely on the system."). Also, participants stated the level of perceived safety via four 7-point semantic differentials from -3 (anxious/agitated/unsafe/timid) to +3 (relaxed/calm/safe/confident) [20]. We additionally measured the perceived quality of situation awareness with the situation awareness rating technique (SART) [52]. This situation awareness [19] may be a predictor of "how a person will choose to act on that SA" [19, p. 86]. Also, we employed the subscales Performance ("I would have performed better than the automated vehicle in this situation"), Judgement ("The automated vehicle made an unsafe judgment in this situation"), and Reaction ("The automated vehicle reacted appropriately to the environment") from the Situational Trust Scale for Automated Driving [27].

Furthermore, the following 7-point Likert scale questions were asked based on visualizations of object recognition, prediction, and maneuver planning (in brackets, we write the abbreviation for reporting purposes):

- "The automated vehicle recognizes all vehicles in every situation perfectly" (recognized vehicle)
- "The automated vehicle predicts all vehicle paths in every scene perfectly" (predict paths)
- "My vehicle recognized where the other vehicles will drive to very well." (certain drive to)
- "The automated vehicle has perfect longitudinal guidance (braking, acceleration, ...)" (longitudinal guidance)
- "The automated vehicle has perfect lateral guidance (keeping track, staying on the road)" (lateral guidance)
- "It was always clear what the automated vehicle will do next" (clear AV will do next)
- "The board computer has helped me to better recognize road users" (perceive road users)
- "The board computer helped me perceive distances better" (perceive distances better)
- "The board computer helped me to perceive my surroundings better" (perceive surroundings better)
- "How would you rate the driving style of the automated vehicle?" (1 = completely safe to 7 = completely dangerous) (driving style)

Finally, general questions about 3D displays and their usefulness and feedback regarding concerns, improvement proposals, and other thoughts were asked qualitatively (with open-ended text answers).

3.4 Procedure

Prior to the initiation of each session, participants were required to sign a consent form and receive an explanation of the simulated scenario. Each participant underwent the simulated ride with the AV three times (constituting a within-subjects design). Each iteration of the course lasted approximately 4 minutes and 30 seconds. To control for order effects, the displays were presented in a counterbalanced order across the three rides. The scenario was introduced as follows:

You will take a ride in a highly automated vehicle through a town. The vehicle automatically steers, brakes, and accelerates (lateral and longitudinal guidance). The automated vehicle will try to predict the other vehicles' future maneuvers (=intention) via its onboard sensors and determine its trajectory accordingly. Additionally, there will be a board computer on which all recognized traffic participants will be displayed. While participating, you are supposed to imagine sitting in such an automated vehicle, following the entire journey attentively, and then assessing it.

Therefore, the task of the participant was solely to observe the simulated automated ride. The difference in the conditions was solely **how** the additional visualization of the AV was presented (i.e., conventional display, LookingGlass, LumePad in Lightfield mode). The duration of the study, including all three rides, was approximately 45 minutes on average for each participant. Participants were provided with a monetary compensation of 8€.

4 RESULTS

4.1 Data Analysis

Prior to every statistical test, we checked the required assumptions (normality distribution and homogeneity of variance assumption). Due to the within-subject design, we employed either Friedman tests (non-parametric data) or one-way ANOVAs (parametric data).

For post-hoc tests, we used Holm correction and employed Dunn's test if not stated otherwise. Additionally, we used the BayesFactor package [40], using Jeffreys-Zellner-Siow (JZS) priors, to compute Bayes factors. Interpretations were made according to Jeffreys [29]. R in version 4.3.1 and RStudio in version 2023.06.1 was employed. All packages were up to date in August 2023.

4.2 Participants

The age ranged from 22 years to 60 years, with an average age of $M=28.73$ years ($SD=9.03$). Five participants reported their gender as female and ten participants as male. Of those 15 participants, nine were still studying, four were employed, and one did not specify their employment status. The participants have held a driving license for an average of $M=10.10$ ($SD=6.90$) years. Seven participants indicated they drive "daily", one participant chose "on working days", three picked "3-4 times a week", one was "one time a week", two picked "1-3 times a month", and the last participant was driving "less than 1 time a month". One participant drove "25000 - 32999 km", seven participants drove "15000 - 24999 km", four participants drove "7000 - 14999 km", and three chose "less than 7000 km" per year.

4.3 Statistical Analysis

Descriptive data can be found in [Table 2](#).

First, we evaluated every dependent variable with the appropriate inferential test (Friedman or repeated measures one-way ANOVA). We only found one significant difference. A Friedman test found that there was a significant difference in the assessment of recognized vehicles (question: "The automated vehicle perfectly recognizes all vehicles in any situation") ($\chi^2(2)=6.06$, $p=0.05$). However, post-hoc tests showed no significant differences between the conditions.

Then, we used the BayesFactor package [40] to determine whether there was evidence for no differences between the conditions (see [Table 1](#)).

We found that for most dependent variables, there is anecdotal evidence supporting no difference. For Supply, Trust, and longitudinal and later guidance, there is moderate evidence (see [Table 1](#)).

Participants ranked the display technology ("Please rate the display technologies from best to worst"). A Friedman's ANOVA found a significant difference (see [Figure 4](#)). The post-hoc test showed that the LookingGlass was significantly better ranked compared to the LumePad both in 2D and using it in 3D mode. The qualitative feedback showed that this is mostly due to the lower eye strain induced by the LookingGlass.

Table 1. BayesFactor analysis to evaluate whether there is evidence to support that there is no difference between the conditions. If the User model is favored, this implies evidence against an influence of the conditions. If the User + Condition model is favored, this implies that the conditions have an influence. Bayes factors provide a continuous quantification of **relative evidence**, where a Bayes factor exceeding 1 indicates evidence supporting one of the models (usually referred to as the numerator), while a Bayes factor less than 1 indicates evidence in favor of the other model (the denominator). For example a value of 2 “**2 times more probable under the null compared to the alternative hypothesis**” (see website of [38]).

Dependent Variable	Evidence	BF	Favoured Model
Mental Workload [24]	Anecdotal	1/2.99 = 0.33	User model
Understanding [52]	Anecdotal	1/1.55 = 0.65	User model
Demand [52]	Anecdotal	1.27	User + Condition model
Supply [52]	Moderate	1/5.87 = 0.17	User model
Situation Awareness [52]	Anecdotal	1/1.52 = 0.66	User model
Trust [33]	Moderate	1/5.07 = 0.20	User model
Understandability [33]	Anecdotal	1/1.43 = 0.70	User model
Perceived Safety [20]	Anecdotal	1.83	User + Condition model
Driving Style	Anecdotal	1/1.11 = 0.90	User model
Longitudinal Guidance	Moderate	1/4.25 = 0.24	User model
Lateral Guidance	Moderate	1/4.03 = 0.25	User model
Judgement [27]	Anecdotal	1/2.03 = 0.49	User model
Reaction [27]	Anecdotal	1/1.46 = 0.68	User model
Performance [27]	Anecdotal	1/2.19 = 0.46	User model
Recognized vehicle	Anecdotal	1.16	User + Condition model
Predict paths	Anecdotal	1/2.17 = 0.46	User model
Clear AV will do next	Anecdotal	1/1.51 = 0.66	User model
Recognize road users	Anecdotal	1/1.43 = 0.70	User model
Perceive distances better	Anecdotal	1.47	User + Condition model
Perceive surroundings better	Anecdotal	1/2.17 = 0.46	User model
Certain drive to	Anecdotal	1/2.71 = 0.37	User model

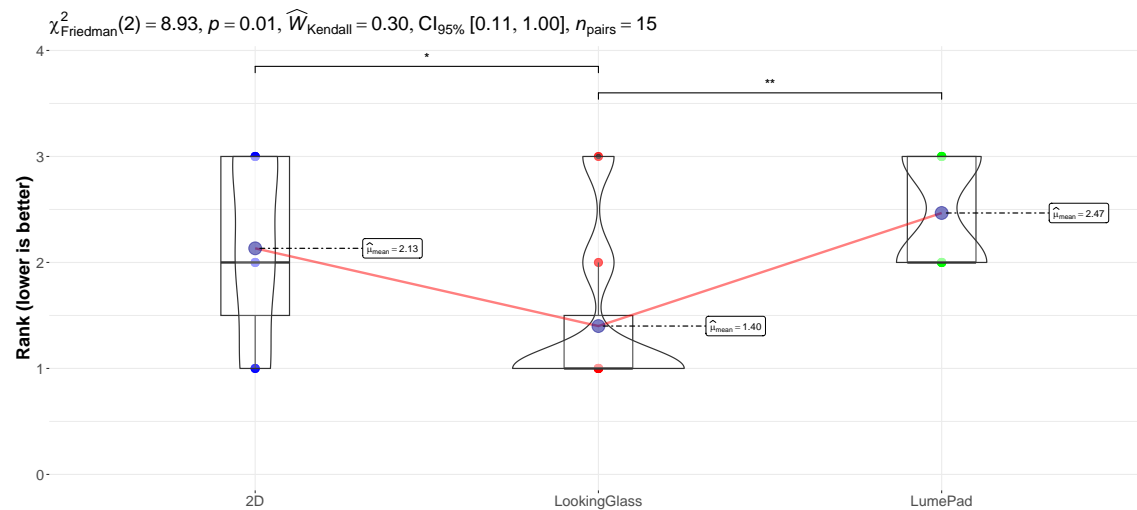


Fig. 4. Ranking of the devices. The plot shows the individual data points, the bar charts, the density curve, the mean values, and the significant post-hoc tests.

4.4 Qualitative Feedback

The qualitative feedback, on the one hand, showed that two participants did not find the 3D display to offer relevant advantages (e.g., “Not useful! 2D would be enough for me because you don’t necessarily have a better view through 3D.” or “Interesting but I’m not sure if they are useful. A great feature or addition in a car to stand out, but not really needed IMO [in my opinion].”). On the other hand, two other participants stated that the 3D displays enhanced their spatial understanding: “Unlike 2D, 3D gives you a slightly better feel for ranges and the spatial environment. The Looking Glass was more pleasant to look at and less strenuous on the eyes, unlike the LumePad.” and “The[y] are very useful. Its easier to predict distances on the 3D displays.” Only two participants provided feedback regarding concerns, improvements, or other thoughts, with one respondent requesting higher compensation and another suggesting visualization of navigation information.

5 DISCUSSION

This work investigated the effect of display technology on mental workload, situation awareness, trust, and performance assessment in AVs. Using Bayes factors, we were able to show moderate evidence that trust and subjective assessment of lateral and longitudinal guidance are not affected by the display technology. In line with prior work [4], we found no significant differences in the mental workload when comparing 2D and 3D displays. In the following, we discuss our results with regards to the larger literature.

5.1 3D Displays — Necessity, Gimmick, Nice-to-Have, or Distraction?

There has been some media hype about 3D displays. While most articles tend to lean on the positive side (e.g.,^{2,3}), there are also negative commentaries (e.g., “Mercedes-Benz doesn’t say what this technology might be used for, or what kinds of benefits it could provide — which is disconcerting considering it’s an idea that failed to provide any usefulness the case of those two doomed smartphones”⁴). While the 3D display received qualitative feedback suggesting positive effects on user acceptance and the appearance, our data suggests that there is little to no impact of this on users in AVs. The question arises whether such displays, as they are already available to consumers⁵, are a necessity, a gimmick, nice-to-have, or even a distraction. Previous work on scenarios unlike automated driving confirmed that they have no significant effect on mental workload [1]. Other positive effects such as situation assessment and performance [57, 58] or shorter glance times [45] become less important with the proliferation of automation. According to our data, we suggest that 3D displays in the context of AVs and visualizing information about its inner workings could be considered a nice-to-have feature. However, they offer a sense of pleasantness and hint at advanced technological capabilities. In addition to emphasizing calibrated trust and acceptance [44], the automotive industry is increasingly focusing on user experience [51] due to its influence on users’ engagement with vehicle interfaces and interiors [23], and its recognized value as a unique selling feature [17]. Thus, it remains to be investigated how 3D display technologies can enhance user experience and assist users in non-driving related tasks.

Our data suggests that there are either no significant differences or even evidence for no difference. Therefore, we suggest that the development of 3D displays has to be properly adjusted to the specific use case. For the automotive use

²https://uk-motor1-com.translate.google/news/433025/2021-mercedes-s-class-interior-preview/?_x_tr_sl=en&_x_tr_tl=de&_x_tr_hl=de&_x_tr_pto=sc; Accessed: 06.10.2023

³<https://www.autoblog.com/2020/07/09/2021-mercedes-benz-s-class-technology-interview/>; Accessed: 06.10.2023

⁴<https://www.theverge.com/2020/7/8/21316700/mercedes-benz-mbox-second-generation-infotainment-system-touchscreen-hey-mercedes>; Accessed: 06.10.2023

⁵<https://group-media.mercedes-benz.com/marsMediaSite/en/instance/ko/Meet-the-S-Class-DIGITAL-My-MBUX-Mercedes-Benz-User-Experience-At-home-on-the-road-luxurious-and-digital.xhtml?oid=46757031>; Accessed: 06.10.2023

case, we assume that there is negligible effect for information visualizations regarding the AV's situation understanding and prediction, therefore, rendering current advances in the industry mute (e.g.,⁶). The currently used 3D displays are, however, expectedly safe to use.

5.2 Methodical Reflections

Due to the unavailability of consumer-grade vehicles capable of automated driving and having incorporated 3D displays, we conducted a lab-based study. The driving scene was visualized using a 75" TV, therefore, there was no actual depth there. This might have altered the mapping between the driving scene and the 3D display. While a real driving scene as captured via dashcams could have been used and post-processed to display relevant information (e.g., where pedestrians are), to the best of our knowledge, there is no solely RGB-based prediction of other vehicles' intention. Therefore, a Unity scene was required in which the future position of all objects was known.

Based on these methodical constraints, we believe that our approach was valid to elicit the first results regarding 3D displays in AVs. Nonetheless, this should be re-evaluated in a real-world scenario once possible.

We also applied Bayes factors to draw conclusions about the absence of effects. This is a "coherent approach to determining whether non-significant results support a null hypothesis over a theory, or whether the data are just insensitive." [18, p. 1].

5.3 Limitations

There are two main limitations: The sample size for this study was rather small, with only 15 participants. This limits the generalizability of the findings to a larger population. The study was also conducted in a laboratory setting, which does not fully capture the complexities of real-world environments. Specifically, the setting was not in a vehicle, which could have a significant impact on the results [12, 26]. Therefore, these findings may not be directly applicable to the in-vehicle setting. However, this was due to the unavailability of both AVs and varying display technologies in vehicles. Additionally, as we did not want to tarnish the results by **not** employing state-of-the-art displays, these varied in size, aspect ratio, and color space.

6 CONCLUSION

Overall, this work provides additional insights into the effects of 3D displays in an automotive context. In a user study with N=15 participants, we found evidence that 3D displays possibly do not alter trust and perceived AV capabilities. Nonetheless, participants highlighted that they find it interesting and sometimes even useful. They also clearly ranked the LookingGlass as their favorite technology representative. This work helps to focus on relevant novel technologies other than 3D displays and their use cases in AVs.

OPEN SCIENCE

The implementation will be made available to interested 3rd parties.

ACKNOWLEDGMENTS

We thank all study participants. We also thank the Human Factors group of Ulm University for allowing us to use their LookingGlass device. This research was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German

⁶<https://www.continental.com/en/products-and-innovation/innovation/infotainment-in-the-car/natural-3d-display/>; Accessed: 06.10.2023

Research Foundation) through the project "Empirical Assessment of Presence and Immersion in Augmented and Virtual Realities" (project number: 425365472).

REFERENCES

- [1] Nora Broy. 2016. Stereoscopic 3D User Interfaces: Exploring the Potentials and Risks of 3D Displays in Cars.
- [2] Nora Broy, Florian Alt, Stefan Schneegass, Niels Henze, and Albrecht Schmidt. 2013. Perceiving Layered Information on 3D Displays Using Binocular Disparity. In *Proceedings of the 2nd ACM International Symposium on Pervasive Displays* (Mountain View, California) (*PerDis '13*). Association for Computing Machinery, New York, NY, USA, 61–66. <https://doi.org/10.1145/2491568.2491582>
- [3] Nora Broy, Florian Alt, Stefan Schneegass, and Bastian Pfleging. 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* (Seattle, WA, USA) (*AutomotiveUI '14*). Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/2667317.2667319>
- [4] Nora Broy, Elisabeth André, and Albrecht Schmidt. 2012. Is Stereoscopic 3D a Better Choice for Information Representation in the Car?. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Portsmouth, New Hampshire) (*AutomotiveUI '12*). Association for Computing Machinery, New York, NY, USA, 93–100. <https://doi.org/10.1145/2390256.2390270>
- [5] Jana Busshoff, Rabi R Datta, Thomas Bruns, Robert Kleinert, Bernd Morgenstern, David Pfister, Costanza Chiapponi, Hans F Fuchs, Michael Thomas, Caroline Gietzelt, et al. 2022. Gender benefit in laparoscopic surgical performance using a 3D-display system: data from a randomized cross-over trial. *Surgical Endoscopy* 36, 6 (2022), 4376–4385.
- [6] Jian Chen, Veeral Hardev, Jason Hartlove, Jurgen Hofler, and Ernie Lee. 2012. 66.1: Distinguished Paper: A High-Efficiency Wide-Color-Gamut Solid-State Backlight System for LCDs Using Quantum Dot Enhancement Film. *SID Symposium Digest of Technical Papers* 43, 1 (2012), 895–896. <https://doi.org/10.1002/j.2168-0159.2012.tb05931.x> arXiv:<https://sid.onlinelibrary.wiley.com/doi/pdf/10.1002/j.2168-0159.2012.tb05931.x>
- [7] Jessie Y Chen, Razia V Oden, Caitlin Kenny, and John O Merritt. 2010. *Effectiveness of stereoscopic displays for indirect-vision driving and robot teleoperation*. Technical Report. ARMY RESEARCH LAB ABERDEEN PROVING GROUND MD.
- [8] Jong Kyu Choi and Yong Gu Ji. 2015. Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction* 31, 10 (2015), 692–702.
- [9] Mark Colley, Jan Henry Belz, and Enrico Rukzio. 2021. Investigating the Effects of Feedback Communication of Autonomous Vehicles. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Leeds, United Kingdom) (*AutomotiveUI '21*). Association for Computing Machinery, New York, NY, USA, 263–273. <https://doi.org/10.1145/3409118.3475133>
- [10] Mark Colley, Christian Bräuner, Mirjam Lanzer, Marcel Walch, Martin Baumann, and Enrico Rukzio. 2020. Effect of Visualization of Pedestrian Intention Recognition on Trust and Cognitive Load. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Virtual Event, DC, USA) (*AutomotiveUI '20*). Association for Computing Machinery, New York, NY, USA, 181–191. <https://doi.org/10.1145/3409120.3410648>
- [11] Mark Colley, Benjamin Eder, Jan Ole Rixen, and Enrico Rukzio. 2021. Effects of Semantic Segmentation Visualization on Trust, Situation Awareness, and Cognitive Load in Highly Automated Vehicles. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 155, 11 pages. <https://doi.org/10.1145/3411764.3445351>
- [12] Mark Colley, Pascal Jansen, Enrico Rukzio, and Jan Gugenheimer. 2022. SwiVR-Car-Seat: Exploring Vehicle Motion Effects on Interaction Quality in Virtual Reality Automated Driving Using a Motorized Swivel Seat. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 150 (dec 2022), 26 pages. <https://doi.org/10.1145/3494968>
- [13] Mark Colley, Svenja Krauss, Mirjam Lanzer, and Enrico Rukzio. 2021. How Should Automated Vehicles Communicate Critical Situations? A Comparative Analysis of Visualization Concepts. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 94 (sep 2021), 23 pages. <https://doi.org/10.1145/3478111>
- [14] Mark Colley, Luca-Maxim Meinhardt, Alexander Fassbender, Michael Rietzler, and Enrico Rukzio. 2023. Come Fly With Me: Investigating the Effects of Path Visualizations in Automated Urban Air Mobility. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 2, Article 52 (jun 2023), 23 pages. <https://doi.org/10.1145/3596249>
- [15] Mark Colley, Max Rädler, Jonas Glimmann, and Enrico Rukzio. 2022. Effects of Scene Detection, Scene Prediction, and Maneuver Planning Visualizations on Trust, Situation Awareness, and Cognitive Load in Highly Automated Vehicles. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 2, Article 49 (jul 2022), 21 pages. <https://doi.org/10.1145/3534609>
- [16] Rebecca Currano, So Yeon Park, Dylan James Moore, Kent Lyons, and David Sirkin. 2021. Little Road Driving HUD: Heads-Up Display Complexity Influences Drivers' Perceptions of Automated Vehicles. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 511, 15 pages. <https://doi.org/10.1145/3411764.3445575>
- [17] Henrik Detjen, Sarah Faltalous, Bastian Pfleging, Stefan Geisler, and Stefan Schneegass. 2021. How to Increase Automated Vehicles' Acceptance through In-Vehicle Interaction Design: A Review. *International Journal of Human-Computer Interaction* 37, 4 (Feb. 2021), 308–330. <https://doi.org/10.1080/10447318.2020.1860517>
- [18] Zoltan Dienes. 2014. Using Bayes to get the most out of non-significant results. *Frontiers in psychology* 5 (2014), 781.

- [19] Mica R Endsley, Stephen J Selcon, Thomas D Hardiman, and Darryl G Croft. 1998. A comparative analysis of SAGAT and SART for evaluations of situation awareness. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 42. SAGE Publications Sage CA: Los Angeles, CA, SAGE Publications, Los Angeles, CA, USA, 82–86.
- [20] Stefanie M. Faas, Andrea C. Kao, and Martin Baumann. 2020. A Longitudinal Video Study on Communicating Status and Intent for Self-Driving Vehicle – Pedestrian Interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '20*). Association for Computing Machinery, New York, NY, USA, 1–14. <https://doi.org/10.1145/3313831.3376484>
- [21] David Fattal. 2021. LEIA: The Lume Pad: 3D Lightfield at your fingertips. <https://doi.org/10.1117/12.2597482>
- [22] María Rodríguez Fernández, Eduardo Zalama Casanova, and Ignacio González Alonso. 2015. Review of display technologies focusing on power consumption. *Sustainability* 7, 8 (2015), 10854–10875.
- [23] Anna-Katharina Frison, Philipp Wintersberger, Andreas Riener, Clemens Schartmüller, Linda Ng Boyle, Erika Miller, and Klemens Weigl. 2019. In UX We Trust: Investigation of Aesthetics and Usability of Driver-Vehicle Interfaces and Their Impact on the Perception of Automated Driving. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (*CHI '19*). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3290605.3300374>
- [24] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, Amsterdam, The Netherlands, 139–183.
- [25] Renate Häuslschmid, Max von Bülow, Bastian Pflöging, and Andreas Butz. 2017. Supporting Trust in Autonomous Driving. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces* (Limassol, Cyprus) (*IUI '17*). Association for Computing Machinery, New York, NY, USA, 319–329. <https://doi.org/10.1145/3025171.3025198>
- [26] Philipp Hock, Mark Colley, Ali Askari, Tobias Wagner, Martin Baumann, and Enrico Rukzio. 2022. Introducing VAMPIRE - Using Kinaesthetic Feedback in Virtual Reality for Automated Driving Experiments. In *14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (*AutomotiveUI '22*). Association for Computing Machinery, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3543174.3545252>
- [27] Brittany E. Holthausen, Philipp Wintersberger, Bruce N. Walker, and Andreas Riener. 2020. Situational Trust Scale for Automated Driving (STS-AD): Development and Initial Validation. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Virtual Event, DC, USA) (*AutomotiveUI '20*). Association for Computing Machinery, New York, NY, USA, 40–47. <https://doi.org/10.1145/3409120.3410637>
- [28] Sony Electronics Inc. 2022. Spatial Reality Display. <https://electronics.sony.com/more/spatial-reality-display/p/elfsr1>. [Online; accessed 12-JANUARY-2022].
- [29] Harold Jeffreys. 1998. *The theory of probability*. OUP Oxford.
- [30] Hornbæk Kasper and Hertzum Morten. 2017. Technology Acceptance and User Experience. *ACM Transactions on Computer-Human Interaction* 24, 5 (2017), 1–30.
- [31] Nam Kim, Anh-Hoang Phan, Munkh-Uchral Erdenebat, Ashrafal Alam, Ki-Chul Kwon, Mei-Lan Piao, and Jeong-Hyeon Lee. 2014. 3D display technology. *Disp. Imag* 1, 1 (2014), 73–95.
- [32] Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. 2015. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 9, 4 (2015), 269–275.
- [33] Moritz Körber. 2019. Theoretical Considerations and Development of a Questionnaire to Measure Trust in Automation. In *Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018)*, Sebastiano Bagnara, Riccardo Tartaglia, Sara Albolino, Thomas Alexander, and Yushi Fujita (Eds.). Springer International Publishing, Cham, 13–30.
- [34] Patrick Lindemann, Tae-Young Lee, and Gerhard Rigoll. 2018. Catch my drift: Elevating situation awareness for highly automated driving with an explanatory windshield display user interface. *Multimodal Technologies and Interaction* 2, 4 (2018), 71.
- [35] Yisi Liu, Fitri Trapsilawati, Zirui Lan, Olga Sourina, Henry Johan, Fan Li, Chun-Hsien Chen, and Wolfgang Mueller-Wittig. 2019. Human factors evaluation of ATC operational procedures in relation to use of 3D display. In *International Conference on Applied Human Factors and Ergonomics*. Springer, Springer, Cham, 715–726.
- [36] Andreas Löcken, Wilko Heuten, and Susanne Boll. 2016. AutoAmbiCar: Using Ambient Light to Inform Drivers About Intentions of Their Automated Cars. In *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Ann Arbor, MI, USA) (*AutomotiveUI '16 Adjunct*). Association for Computing Machinery, New York, NY, USA, 57–62. <https://doi.org/10.1145/3004323.3004329>
- [37] Inc. Looking Glass Factory. 2022. Holograms for your home, office, & home office. <https://lookingglassfactory.com/>. [Online; accessed 12-JANUARY-2022].
- [38] Dominique Makowski, Mattan S Ben-Shachar, SH Annabel Chen, and Daniel Lüdecke. 2019. Indices of effect existence and significance in the Bayesian framework. *Frontiers in psychology* 10 (2019), 2767.
- [39] John P McIntire, Paul R Havig, and Eric E Geiselman. 2012. What is 3D good for? A review of human performance on stereoscopic 3D displays. *Head-and-Helmet-Mounted Displays XVII; and Display Technologies and Applications for Defense, Security, and Avionics VI* 8383 (2012), 280–292.
- [40] Richard Morey. 2022. *Richarddmorey/BayesFactor*. Retrieved 10 January, 2023 from <https://github.com/richarddmorey/BayesFactor>
- [41] Bonnie M Muir and Neville Moray. 1996. Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics* 39, 3 (1996), 429–460.
- [42] Aljoscha C. Neubauer, Sabine Bergner, and Martina Schatz. 2010. Two- vs. three-dimensional presentation of mental rotation tasks: Sex differences and effects of training on performance and brain activation. *Intelligence* 38, 5 (2010), 529–539. <https://doi.org/10.1016/j.intell.2010.06.001>

- [43] Richard E Nisbett and Timothy D Wilson. 1977. The halo effect: evidence for unconscious alteration of judgments. *Journal of personality and social psychology* 35, 4 (1977), 250.
- [44] Bastian Pfleging, Shadan Sadeghian, and Debargha Dey. 2021. User Interfaces for Automated Vehicles. *it - Information Technology* 63, 2 (June 2021), 73–75. <https://doi.org/10.1515/itit-2021-0020>
- [45] Matthew Pitts, Elvir Hasedzic, Lee Skrypchuk, Alex Attridge, and Mark Williams. 2015. Adding Depth: Establishing 3D Display Fundamentals for Automotive Applications. *SAE Technical Papers* 2015 (04 2015). <https://doi.org/10.4271/2015-01-0147>
- [46] Hui Ren, Li-Xia Ni, Hai-Feng Li, Xin-Zhu Sang, Xin Gao, and Qiong-Hua Wang. 2020. Review on tabletop true 3D display. *Journal of the Society for Information Display* 28, 1 (2020), 75–91.
- [47] Daedal Research. 2021. US 3D Display Market: Size, Trend & Forecast with Impact of COVID-19 (2021-2025). <https://www.researchandmarkets.com/reports/5459155/us-3d-display-market-size-trend-and-forecast>. Accessed: 2022-04-14.
- [48] Precedence Research. 2021. Commercial Display Market Size Worth Around US\$ 72.4 Bn by 2030. <https://www.globenewswire.com/news-release/2021/12/28/2358387/0/en/Commercial-Display-Market-Size-Worth-Around-US-72-4-Bn-by-2030.html>. Accessed: 2022-04-14.
- [49] Tobias Schneider, Joana Hois, Alischa Rosenstein, Sabiha Ghellal, Dimitra Theofanou-Fülbier, and Ansgar R.S. Gerlicher. 2021. ExplAIIn Yourself! Transparency for Positive UX in Autonomous Driving. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 161, 12 pages. <https://doi.org/10.1145/3411764.3446647>
- [50] Ranbir Singh, KN Narayanan Unni, Ankur Solanki, et al. 2012. Improving the contrast ratio of OLED displays: An analysis of various techniques. *Optical Materials* 34, 4 (2012), 716–723.
- [51] Annika Stampf, Mark Colley, and Enrico Rukzio. 2022. Towards Implicit Interaction in Highly Automated Vehicles - A Systematic Literature Review. *Proceedings of the ACM on Human-Computer Interaction* 6, MHCI (Sept. 2022), 191:1–191:21. <https://doi.org/10.1145/3546726>
- [52] Richard M Taylor. 2017. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational awareness*. Routledge, Abingdon, UK, 111–128.
- [53] Roger Wahba, Rabi Datta, Jana Bußhoff, Thomas Bruns, Andrea Hedergott, Caroline Gietzelt, Georg Dieplinger, Hans Fuchs, Bernd Morgenstern, Desdemona Möller, et al. 2020. 3D versus 4K display system—Influence of “State-of-the-art”-display technique on surgical performance (IDOSP-study) in minimally invasive surgery: A randomized cross-over trial. *Annals of surgery* 272, 5 (2020), 709.
- [54] Wendell R Watkins, Garrett D Heath, Michael D Phillips, J Mathieu Valetton, and Alexander Toet. 2001. Search and target acquisition: single line of sight versus wide baseline stereo. *Optical Engineering* 40, 9 (2001), 1914–1927.
- [55] Florian Weidner and Wolfgang Broll. 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 (Oldenburg, Germany) (AutomotiveUI '17)*. Association for Computing Machinery, New York, NY, USA, 212–216. <https://doi.org/10.1145/3131726.3131763>
- [56] Florian Weidner and Wolfgang Broll. 2019. Exploring Large Stereoscopic 3D Dashboards for Future Automotive User Interfaces. In *Advances in Human Aspects of Transportation*, Neville Stanton (Ed.). Springer International Publishing, Cham, 502–513.
- [57] Florian Weidner and Wolfgang Broll. 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 (Palermo, Italy) (PerDis '19)*. Association for Computing Machinery, New York, NY, USA, Article 5, 7 pages. <https://doi.org/10.1145/3321335.3324937>
- [58] Florian Weidner and Wolfgang Broll. 2020. Stereoscopic 3D dashboards. *Personal and Ubiquitous Computing* 26, 3 (27 Jul 2020), 697–719. <https://doi.org/10.1007/s00779-020-01438-8>
- [59] Marc Wilbrink, Anna Schieben, and Michael Oehl. 2020. Reflecting the Automated Vehicle’s Perception and Intention: Light-Based Interaction Approaches for on-Board HMI in Highly Automated Vehicles. In *Proceedings of the 25th International Conference on Intelligent User Interfaces Companion (Cagliari, Italy) (IUI '20)*. Association for Computing Machinery, New York, NY, USA, 105–107. <https://doi.org/10.1145/3379336.3381502>

A DESCRIPTIVE RESULTS

Variable	Levels	Min	q ₁	\tilde{x}	\bar{x}	q ₃	Max	s	IQR
Mental Workload [24]	2D	1.00	3.00	6.00	6.80	9.50	17.00	4.52	6.50
	LookingGlass	1.00	3.50	6.00	7.00	9.50	17.00	5.36	6.00
	LumePad	1.00	5.00	7.00	8.40	13.50	18.00	5.08	8.50
	all	1.00	3.00	6.00	7.40	12.00	18.00	4.94	9.00
Perceived safety [20]	2D	-0.25	0.88	2.00	1.68	2.62	3.00	1.05	1.75
	LookingGlass	-0.75	0.75	2.25	1.82	2.88	3.00	1.28	2.12
	LumePad	-2.00	0.25	1.00	1.13	2.38	3.00	1.48	2.12
Understanding [52]	2D	8.00	13.00	13.00	13.93	16.00	19.00	2.79	3.00
	LookingGlass	4.00	14.00	17.00	15.27	17.00	19.00	3.71	3.00
	LumePad	5.00	12.00	14.00	13.67	15.50	19.00	3.66	3.50
	all	4.00	13.00	14.00	14.29	17.00	19.00	3.41	4.00

Demand [52]	2D	4.00	7.50	11.00	10.53	13.00	17.00	3.85	5.50
	LookingGlass	3.00	6.50	9.00	8.93	11.50	16.00	3.77	5.00
	LumePad	3.00	8.00	10.00	10.67	13.50	18.00	4.40	5.50
	all	3.00	7.00	9.00	10.04	12.00	18.00	4.01	5.00
Supply [52]	2D	12.00	14.50	16.00	16.80	19.00	23.00	3.14	4.50
	LookingGlass	11.00	15.00	17.00	16.47	18.50	20.00	2.70	3.50
	LumePad	9.00	15.00	16.00	16.60	19.00	23.00	3.70	4.00
	all	9.00	15.00	17.00	16.62	19.00	23.00	3.14	4.00
Situation Awareness [52]	2D	10.00	17.00	19.00	20.20	22.00	30.00	5.14	5.00
	LookingGlass	11.00	18.00	22.00	22.80	28.00	32.00	6.06	10.00
	LumePad	11.00	14.50	22.00	19.60	23.50	28.00	5.58	9.00
	all	10.00	17.00	21.00	20.87	24.00	32.00	5.66	7.00
Trust [33]	2D	2.00	3.25	4.00	3.80	4.00	5.00	0.80	0.75
	LookingGlass	1.00	3.25	4.00	3.77	4.75	5.00	1.19	1.50
	LumePad	2.00	3.00	4.00	3.67	4.25	5.00	0.94	1.25
	all	1.00	3.00	4.00	3.74	4.50	5.00	0.97	1.50
Understandability [33]	2D	2.25	3.38	4.00	3.80	4.12	4.75	0.70	0.75
	LookingGlass	2.75	3.62	4.25	4.07	4.50	5.00	0.64	0.88
	LumePad	1.00	3.50	3.75	3.55	4.00	4.75	1.03	0.50
	all	1.00	3.50	4.00	3.81	4.50	5.00	0.82	1.00
Driving style	2D	2.00	2.00	3.00	3.53	5.00	7.00	1.81	3.00
	LookingGlass	2.00	2.00	2.00	3.20	4.00	7.00	1.70	2.00
	LumePad	1.00	2.00	3.00	3.80	5.00	7.00	1.93	3.00
	all	1.00	2.00	3.00	3.51	5.00	7.00	1.79	3.00
Longitudinal	2D	1.00	4.50	5.00	5.07	6.00	7.00	1.49	1.50
	LookingGlass	1.00	4.00	5.00	4.87	6.00	7.00	1.55	2.00
	LumePad	1.00	4.00	5.00	4.73	5.50	7.00	1.62	1.50
	all	1.00	4.00	5.00	4.89	6.00	7.00	1.53	2.00
Lateral	2D	3.00	5.50	6.00	5.93	7.00	7.00	1.10	1.50
	LookingGlass	2.00	5.00	6.00	5.53	7.00	7.00	1.46	2.00
	LumePad	3.00	5.00	6.00	5.67	7.00	7.00	1.29	2.00
	all	2.00	5.00	6.00	5.71	7.00	7.00	1.27	2.00
Judgement [27]	2D	1.00	2.00	3.00	3.27	4.50	6.00	1.62	2.50
	LookingGlass	1.00	2.00	2.00	2.60	3.00	6.00	1.35	1.00
	LumePad	1.00	2.00	2.00	2.73	3.00	6.00	1.49	1.00
	all	1.00	2.00	3.00	2.87	3.00	6.00	1.49	1.00
Reaction [27]	2D	4.00	4.50	5.00	5.40	6.00	7.00	1.12	1.50
	LookingGlass	2.00	5.00	5.00	5.13	6.00	7.00	1.25	1.00
	LumePad	3.00	4.00	4.00	4.80	6.00	7.00	1.15	2.00
	all	2.00	4.00	5.00	5.11	6.00	7.00	1.17	2.00
Performance [27]	2D	1.00	4.00	4.00	4.67	5.50	7.00	1.63	1.50
	LookingGlass	1.00	3.00	4.00	4.20	5.00	7.00	1.61	2.00
	LumePad	1.00	4.00	4.00	4.27	5.00	7.00	1.49	1.00
	all	1.00	4.00	4.00	4.38	5.00	7.00	1.56	1.00
Recognized vehicle	2D	4.00	5.00	5.00	5.47	6.00	7.00	0.83	1.00
	LookingGlass	3.00	5.50	6.00	5.93	7.00	7.00	1.22	1.50
	LumePad	3.00	4.50	5.00	5.20	6.00	7.00	1.42	1.50
	all	3.00	5.00	6.00	5.53	6.00	7.00	1.20	1.00
Predict paths	2D	4.00	5.00	5.00	5.47	6.00	7.00	0.92	1.00
	LookingGlass	3.00	5.00	6.00	5.67	6.00	7.00	1.11	1.00
	LumePad	3.00	4.00	5.00	5.13	6.50	7.00	1.46	2.50
	all	3.00	5.00	6.00	5.42	6.00	7.00	1.18	1.00
Clear AV will do next	2D	3.00	4.50	5.00	4.93	5.50	7.00	1.28	1.00
	LookingGlass	3.00	4.00	5.00	5.07	6.00	7.00	1.22	2.00

	LumePad	1.00	3.00	5.00	4.20	6.00	7.00	1.93	3.00
	all	1.00	4.00	5.00	4.73	6.00	7.00	1.53	2.00
Recognize road users	2D	3.00	4.00	5.00	5.20	6.00	7.00	1.26	2.00
	LookingGlass	1.00	5.00	6.00	5.73	7.00	7.00	1.62	2.00
	LumePad	1.00	4.00	5.00	4.80	6.00	7.00	1.70	2.00
	all	1.00	4.00	6.00	5.24	6.00	7.00	1.55	2.00
Perceive distances better	2D	1.00	2.50	4.00	3.87	5.00	7.00	1.85	2.50
	LookingGlass	1.00	4.00	5.00	4.80	6.50	7.00	2.11	2.50
	LumePad	1.00	2.00	3.00	3.67	5.50	6.00	1.95	3.50
	all	1.00	2.00	4.00	4.11	6.00	7.00	1.99	4.00
Perceive surroundings better	2D	3.00	5.00	5.00	5.07	5.50	7.00	1.16	0.50
	LookingGlass	1.00	5.50	6.00	5.67	7.00	7.00	1.76	1.50
	LumePad	1.00	5.00	5.00	5.00	6.00	6.00	1.41	1.00
	all	1.00	5.00	5.00	5.24	6.00	7.00	1.46	1.00
Certain drive to	2D	4.00	5.00	6.00	5.87	6.50	7.00	0.92	1.50
	LookingGlass	4.00	5.50	6.00	6.13	7.00	7.00	0.99	1.50
	LumePad	2.00	5.00	6.00	5.60	7.00	7.00	1.55	2.00
	all	2.00	5.00	6.00	5.87	7.00	7.00	1.18	2.00

Table 2. Table of scores.