Introducing Novel Technologies in the Car – Conducting a Real-World Study to Test 3D Dashboards

Nora Broy[1,2,3], Mengbing Guo[1], Stefan Schneegass[3], Bastian Pfleging[1,3], Florian Alt[1]

[1]University of Munich (LMU)
Media Informatics Group
Amalienstraße 17
80333 Munich, Germany
firstname.lastname@ifi.lmu.de

[2]BMW Group
Research & Technology
Hanauer Straße 46
80992 Munich, Germany
Nora.NB.Broy@bmw.de

[3]University of Stuttgart
Institute for Visualization and Interactive Systems
Pfaffenwaldring 5a
70569 Stuttgart, Germany
firstname.lastname@vis.uni-stuttgart.de

ABSTRACT

Today, the vast majority of research on novel automotive user interface technologies is conducted in the lab, often using driving simulation. While such studies are important in early stages of the design process, we argue that ultimately studies need to be conducted in the real-world in order to investigate all aspects crucial for adoption of novel user interface technologies in commercial vehicles. In this paper, we present a case study that investigates introducing autostereoscopic 3D dashboards into cars. We report on studying this novel technology in the real world, validating and extending findings of prior simulator studies. Furthermore, we provide guidelines for practitioners and researchers to design and conduct real-world studies that minimize the risk for participants while at the same time yielding ecologically valid findings.

Keywords

Automotive UIs; real world study; stereoscopic 3D

Categories and Subject Descriptors

H.5.2 [**Information interfaces and presentation (e.g., HCI)**]: User Interfaces

1. INTRODUCTION

Advances in engineering and computer science allow novel technologies to be introduced in cars at a rapidly accelerating pace. This includes input devices, such as touch screens, eye trackers, or sensors that enable mid-air gestures as well as output devices, such as head-mounted or 3D displays. Such technologies require careful investigation, primarily with regard to performance, driver distraction and behavior but also with regard to acceptance and UX.

To not put the driver at risk, the vast majority of researchers in the automotive domain conducts research in the lab, often using driving simulators. While we believe this to be important and valuable in early stages of the design process, we argue that there is a clear need to conduct more research in the real world to unveil the true benefits and pitfalls of new technology and interfaces built upon.

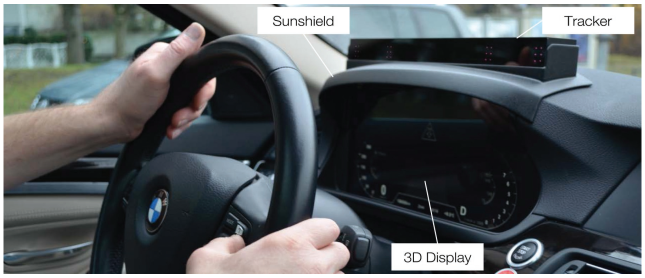

Figure 1: We replaced the instrument cluster of a car with an autostereoscopic display technology using eye tracking to adjust the sweet spot.

In particular, we show how real-world studies can blend in the development process and complement findings from the lab. We present important lessons learned that can be adopted by researchers and practitioners to conduct safe real-world studies while at the same time obtaining valuable, ecologically highly valid results.

To showcase the potential and benefits of complementing research in the lab with real-world driving studies, we report on our work on introducing autostereoscopic 3D displays into the vehicle. The driver behind our research is the question of whether or not visual cues presented at different depth positions are being perceived as more urgent by the user compared to color cues. We designed a 3D dashboard and integrated an autostereoscopic display into a car. We then conducted a real world study with 32 participants that had to rate the urgency of visual cues presented on one of three different depth levels in either white or red color. In addition we gathered qualitative feedback about the acceptance of a stereoscopic (S3D) visualization of informative content while driving. Our results show that color has a greater impact on the perceived urgency as S3D while their combination maximizes urgency ratings. Moreover, this study validates a strong increase of user experience due to a S3D visualization and the gain of UI elements that present temporal and spatial relations. Moreover, we gained a more detailed insight into the in-car use of this technology through conducting the study in the real world, e.g. an increase of motion sickness due to S3D and a functional benefit of S3D navigation cues.

The contribution of this work is twofold: We present the first real-world driving study with an autostereoscopic display validating and extending findings of prior laboratory studies. Also, we derive guidelines for safely conducting real-world studies with the aim to evaluate novel user interfaces and technologies in an ecologically valid setting, compared to a driving simulation.

AutomotiveUI'15, September 1–3, 2015, Nottingham, UK

ACM 978-1-4503-3736-6/15/09 ... $15.00.
http://dx.doi.org/10.1145/2799250.2799280

Reference	Test Environment	Subject of Research	Primary Task / Driving Performance	Secondary Task Performance	Gaze Behavior	Usability	User Experience
Broy et al. [5]	Laboratory	Infotainment System	No difference between 2D and S3D	No difference between 2D and S3D	–	S3D highlights interaction focus	S3D increases UX
Broy et al. [7]	Laboratory	Instrument Cluster	–	–	–	–	S3D increases UX
Szczerba et al. [22]	Laboratory	Instrument Cluster	–	S3D increases user performance for visual search tasks for small set sizes	–	–	–
Broy et al. [4]	Simulator	Instrument Cluster	No difference between 2D and S3D	S3D increases task completion times and task accuracy	No difference between 2D and S3D	S3D clarifies information structure	S3D increases UX
Pitts et al. [18]	Simulator	Instrument Cluster	–	S3D increases user performance for identifying depth	S3D decreases eyes-off the road time	–	–

Table 1: Results of former studies investigating in-car S3D displays.

2. BACKGROUND AND RELATED WORK

This work draws upon different strands of research from the domain of automotive user interfaces, most notably methodology, 3D displays, and communicating urgency of information.

2.1 Automotive UI Research Methodology

Evaluations of automotive user interfaces to assess safety and usability can be conducted in different environments, ranging from laboratory studies, via simulator studies, to real-world studies on test tracks, (short-term) road trials or (longer-term) field trials on regular roads. The selection of a certain environment affects the ecologic validity of the results: real world road tests have the highest degree of realism and therefore the highest ecologic validity but the environment cannot be controlled (road conditions, other vehicles, pedestrians, weather, etc.), which impacts reproducibility [8]. Additionally, potentially hazardous situations cannot be prevented completely. In contrast, simulator or even simpler lab studies have a lower validity regarding aspects such as driving behavior but provide a much better reproducibility and comparability.

In general, driving simulator studies are more suitable to identify effects of non-driving-related activities [17] on driving performance compared to on-road investigations [3, 19]. Hence, the increased costs and effort of real world studies due to the considerably higher safety requirements are not appropriate for early investigations of in-vehicle devices [20, 25]. Instead, initial evaluations are usually done in driving simulators, one of the few exceptions being [2], where data was gathered early in the project.

Although high-fidelity driving simulators have a close to 360° view and may provide kinesthetic feedback, a driving simulation can not fully replicate real-world environments, also regarding workload, risk tolerance, and realism. Thus, when the development of in-vehicle devices has reached a certain level, the validation of effects initially found in a driving simulator is necessary in the real world to fully understand parameters of novel interfaces [20, 25].

2.2 S3D Displays in Cars

Research on stereoscopic displays in the car is still in its infancy. To the best of our knowledge user studies regarding the use of stereoscopic 3D displays so far were conducted exclusively in the laboratory (e.g., [5, 22]) or in the simulator e.g., [18]. These studies attribute potential to using stereoscopic 3D compared to monoscopic representations rather than a negative impact on driver distraction. Table 1 outlines the findings of the prior studies.

A key finding of several studies about 3D presentation is the increase of attractiveness of the shown content [10, 21]. Nevertheless, even if stereoscopy offers a pleasant experience it induces simulator sickness [11, 21]. Beside these symptoms, the processing of the stereoscopic content can increase visual and cognitive workload and in turn decreases the driving performance. However, stereoscopic presentations foster an understanding of the 3D scene since the processing of depth cues is highly coupled [14].

As a result, we deliberately opted to validate and extend former findings from the lab (e.g., [4]) in the real world. We already conducted a first test in a real-world driving situation with domain experts to gather qualitative feedback on the usefulness of 3D displays [6]. In contrast, the experiment presented in this paper was conducted with ordinary drivers to evaluate everyday use, focusing on quantitative and qualitative data. As one outcome of the expert evaluation was that S3D may be helpful to express the urgency of UI elements, this serves as an objective of research in this study.

2.3 Communicating Urgency of Information

For time-critical tasks such as driving a car, it is important to prioritize information and encode / communicate urgency. Colors play a major role to encode the significance of information. Red is the color reserved for danger messages [9] and is well associated with risk [15]. Nevertheless, the salience of visual warnings need to be maximized to attract attention in the competition of various visual stimuli [23]. Former research showed that the combination of color and stereoscopic depth improves search times and task completion times [1, 16]. For traditional 2D representations as set of standards has been established to ensure safety and usability ISO 2575 defines symbols and colors that describe a system status (e.g., correct operation or malfunction). DIN EN ISO 15005 provides principles on dialogue management and presents compliance criteria. In particular for warnings and assistance systems, standards have been defined throughout the last years. ISO/TS 16951 and SAE J2395 provide methods to prioritize (simultaneous) messages and warnings and thus complement (DIN EN) ISO 15005.

3. PROTOTYPE

3.1 Test Vehicle

We replaced the instrument cluster (IC) of a BMW 5 series with a 13.3" autostereoscopic display with a native resolution of 1920× 1080 pixels (pixel pitch: 0.153 mm). The car had automatic transmission, hence considerably facilitating the driving task. The car was equipped with Active Cruise Control (ACC), a multifunctional steering wheel, a head-up display (HUD), and a central information display (CID). The displays ensured the correct representation of relevant information, e.g., speed and check controls, in case of a malfunction of the embedded hard- and software of our S3D IC.

The autostereoscopic display consisted of a display unit, an built-in eye tracker, and a Lubuntu PC. The display used lenticular lenses to create the autostereoscopic effect. On top of the display unit an

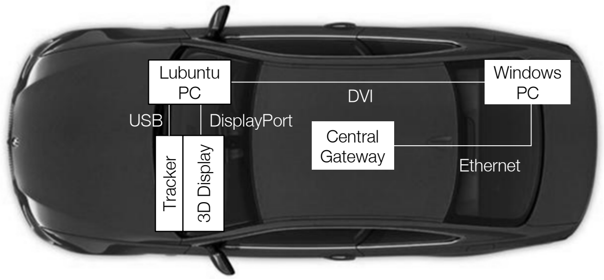

Figure 2: The test vehicle was equipped with several components to supply a stereoscopic display as IC. In particular, two PCs were used – one to render the output on the display and one to process vehicle data.

eye tracker was located. It enhanced the 3D effect by adjusting the sweet spot according to the viewer's eye positions. We positioned the display in a way such that the tracker could detect the viewer by tracking the area above the steering wheel (Figure 1). The viewing distance from the driver to the display was between 600 and 900 mm, depending on the height of the driver. We used a 3D printed sun shield to integrate the display and tracking unit into the car's interior.

A Windows PC by CarTFT[1] was mounted in the trunk. It created the simulation of the instrument cluster and passed a side-by-side image via DVI to the display. The Lububuntu PC interlaced the left and right image with regard to the tracker data. We placed the Lubuntu PC in the footwell of the front passenger side. The Windows PC was connected via Ethernet with the central gateway of the car. In this way, the instrument cluster application received real-time vehicle data such as speed and revolutions per minute (rpm), etc. We used Unity[2] with C# as scripting language to build the interactive instrument cluster application. Figure 2 depicts the arrangement of the integrated components in the vehicle.

3.2 User Interface

We developed an S3D instrument cluster that optimally exploits the 3D space by applying the shape of a tunnel, ranging from screen depth to the rearmost depth plane (Figure 3). The rearmost depth plane is at 44 pixels parallax and corresponds to 30.8 arc-min angular disparity for a viewing distance of 750 mm and an interocular distance of 63.5 mm. Current speed and revolutions per minute (rpm) are displayed on a scale at the outer edge of the IC tunnel at screen depth (zero parallax). The scales for the fuel level and the oil temperature are aligned inside the tunnel behind the speed and rpm scales at 20 pixels parallax to encode their lower priority. At the bottom, status information is displayed on screen depth. Check controls (tell-tales) are located on two dedicated panels on top of the IC tunnel. Thus, we provided a visual anchor for those elements.

The space between the two top panels is dedicated to notifications, instructions, and warnings, which can be displayed on various depth levels and colors in accordance with their level of urgency. We deliberately placed the check controls as well as urgent information at the top of the IC to minimize the distance to the driver's line of sight. All these elements are positioned at outer locations of the IC. This clears space for the flexible presentation of spatial information inside the IC tunnel. We implemented three content elements between which the driver can toggle using a steering wheel button:

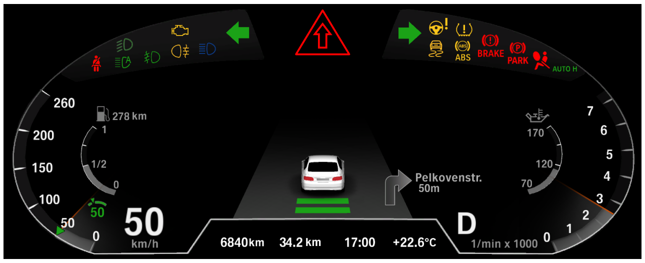

Figure 3: The instrument cluster places speed, RPMs, check controls, and status information at the outer edge to clear space for elements which can profit from a 3D representation as, for example, Active Cruise Control.

Abstract Driving Space (ADS): An abstract visualization of the road is visible which reaches from the screen to the rearmost depth plane. In the 3D space we visualize distance information of the ACC. If the driver activates ACC, green bars on the street represent the distance the car maintains to a preceding vehicle. A 3D car model located behind the green bars indicates if the system detects a vehicle ahead. ACC speed and distance can be adjusted by using buttons on the steering wheel. Besides the ACC, the abstract representation of the street can be used to encode spatial and timely relations of upcoming events, in our case navigation cues. Navigation cues include the turn instruction in the form of an arrow, the street name, and the distance to the turn maneuver. Its depth position adapts in accordance to the vehicle position in relation to the target junction.

3D map: The 3D map visualizes a 3D representation of the current driving scene to enhance navigation tasks. For our IC prototype we implemented a static representation of an example map rather than a fully functional navigation in a 3D map (cf., Figure 4). In this way, we can simply demonstrate the stereoscopic effect.

Infotainment menu: We included a small infotainment menu that allows the driver to choose between various audio sources (cf., Figure 4). The menu displays two lists on two different depth planes. A steering wheel button allows switching between the two lists while the currently selected list is placed further to the front. The driver can scroll through the selected list and select items using steering wheel buttons.

The three available content elements (ADS, map, menu) are visualized by icons on a turntable which appears when toggling between the three elements using a steering wheel button. It turns clockwise as well as the three content elements to visualize a reasonable appearance and disappearance of the elements in 3D space.

4. HYPOTHESES

Based on the aforementioned lab and simulator studies we derived five hypothesis that we investigate in this real world approach:

Hypothesis 1: A stereoscopic presentation increases the user experience and the attractiveness of the instrument cluster.

Hypothesis 2: A stereoscopic presentation has an influence on the cognitive and visual workload.

Hypothesis 3: A stereoscopic visualization increases symptoms of motion sickness.

Hypothesis 4: The use of stereoscopy impacts the perceived urgency of content elements.

Hypothesis 5: A stereoscopic visualization improves the information structure of the user interface.

---

[1] `http://www.cartft.com`

[2] Unity 3D: `www.unity.com`

Figure 4: The driver can toggle between different views in the center of the IC. The left picture depicts a 3D map and the right the infotainment menu.

5. REAL WORLD STUDY

In the following, we report on the evaluation of a stereoscopic display in a real world investigation with 32 participants.

5.1 Study Design and Tasks

During the study, participants drove with a monoscopic and a stereoscopic visualization of the IC. Their primary task was to safely maneuver the car and observe the traffic rules. While driving, participants used the IC view of the ADS and had to follow the navigation cues. Navigation announcements were not supported by auditory cues. As secondary task, we instructed participants to react on instructions placed between the two check control panels. The instructions were triangles with an arrow pointing up- or downward (cf. Broy et al. [4]). Once participants noticed the instruction they had to react by pressing the toggle button on the right side of the steering wheel up- or downwards. If participants correctly reacted by pressing the button in the displayed direction, the arrow inside the triangle switched to a star icon, staying visible for three seconds. Directly after reacting to the instruction, participants had to rate the perceived urgency of the visualization on a four-point Likert scale (1=not urgent at all; 4=very urgent).

We used abstract content for this task (arrows and stars) to avoid any influence on the perceived urgency due to the displayed content. While participants reacted on the instructions they did not use ACC. Nevertheless, they experienced driving ACC for two short segments of each drive. Overall, we varied three independent variables using a repeated measured design:

Display Mode: We present participants a monoscopic (*2D*) and a stereoscopic (*3D*) visualization of the IC.

Depth Layer: The instructions appearing in the IC are displayed either in front of the screen (-14 px parallax), on screen level (0 px parallax), or behind the screen (14 px parallax). We applied these parallaxes of the instructions regardless of the display mode. Note that the displayed elements maintain their size on the screen for all three depth positions.

Color: The instructions are either colored in *white* or *red*.

We counterbalanced the presentation of the display modes over all participants. For each display mode, participants had to react on all $3 \times 2 = 6$ task conditions four times (twice on arrows pointing upwards and twice on arrows pointing downwards). This resulted in 24 instruction tasks per drive presented in randomized order. All tasks did not apply any auditory cues in order to investigate visual cues without any bias through other modalities.

We measured task completion times (TCT), error rates (ER), and the urgency rating for reacting on the instructions. After each test drive, participants filled in a mini AttrakDiff, a Simulator Sickness Questionnaire (SSQ), the Driver Activity Load Index (DALI) and rated statements about their gaze behavior and the information structure of the display on a five-point likert scale (1=strongly disagree; 5=strongly agree). At the end, we interviewed participants.

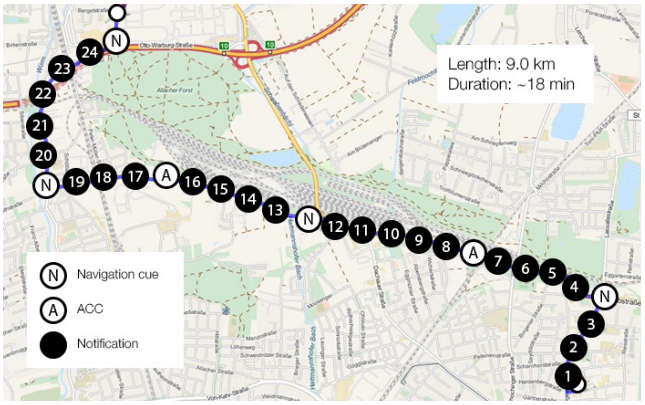

Figure 5: Participants drove the depicted route in each direction using a monoscopic and a stereoscopic presentation of the IC. The examiner triggered the tasks when the driving situations allowed for it.

5.2 Test Route

Before driving the test route, participants completed a short pre-drive to get used to the new IC. The pre-drive, conducted in a residential area with very low traffic density, took roughly 6 minutes and was 1.8 km long. The test route consisted of two drives, one for each display mode of the IC. It started and ended at the parking lot of our lab. The first drive also ended in a parking lot, which offered the opportunity for a break to gather subjective feedback. The routes are common urban roads with low to moderate traffic. Each drive took roughly 18 minutes for the 9 km on each direction. Speed limits along the routes ranged from 30 to 60 km/h.

The routes were carefully inspected before conducting the studies. Secondary tasks were designed to be triggered in suitable locations. Figure 5 shows the distribution of the tasks along the route.

5.3 Setup

During the study one examiner accompanied the participant as a co-driver. The examiner triggered all tasks in driving situations that were comparable but safe enough to allow an interaction with the display and recorded participants' urgency ratings. The test vehicle's interior was equipped with two GoPro cameras recording the whole study. The CID was disabled and covered by a label which showed the four-point Likert-scale to support the participants in rating the perceived urgency of the instructions while driving.

We configured the head-up display of the test vehicle to display the current speed and urgent information, such as warnings and check controls. Thus, we ensured the visualization of important information in the case of a malfunction or a crash of the IC.

5.4 Procedure

We met participants in the parking lot of our lab. First, participants adjusted the seat, mirrors, and the steering wheel to optimally operate all driving controls and to get the best view on the IC. After

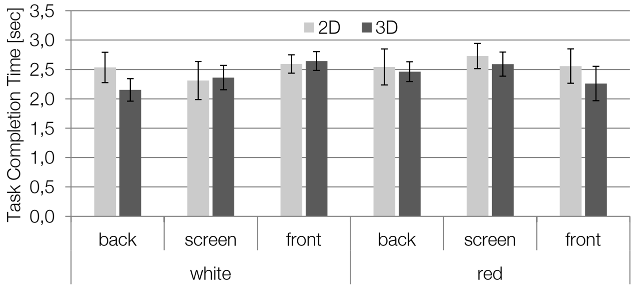

Figure 6: Means and standard errors as error bars for the TCT of confirming the instructions.

we provided participants some basic information on stereoscopic 3D, we tested their ability to view stereoscopic content with a Random Dot Stereogram [13] test and color blindness using an Ishihara test [12]. If the participants passed both tests they qualified for the study and filled in a demographic questionnaire and baseline SSQ. Then, the examiner explained the IC and its components in the respective vision mode (2D/3D) of the first drive as well as the navigation, ACC, and the instruction tasks. Participants were instructed to primarily focus on the driving task and to observe traffic rules but to react on displayed instructions accurately and fast.

Before driving, the participants practiced 12 instruction tasks to get acquainted with first using the steering wheel button and then to tell their urgency rating. After participants felt comfortable with the tasks they practiced the navigation and instruction tasks while driving a short test route in the residential area next to the parking lot before the two test drives started. After each test drive the participants filled in the questionnaires. Before driving with the second display mode, participants explored the IC and its components once again. At the end of the study, we conducted a semi-structured interview with participants about the IC and the S3D visualization.

5.5 Participants

We recruited 32 participants (5 female) aged between 22 and 51 years (M=34.63, SD=7.96). All participants are employees of the BMW Group and have already received special driving training that allows them to steer test vehicles in public. In contrast to the expert review reported by Broy et al. [6], the participants are no experts in UI development and cover rather technical backgrounds, ranging from mechanics over electrical and mechanical engineering to computer science. Three participants had no experience in viewing 3D content at all, while the remaining 29 participants know S3D from cinema and games. Eight participants reported on having experienced ghosting, headache, dizziness, and nausea due to S3D.

6. RESULTS

6.1 Secondary Task Performance

Based on the independent variables, *display mode*, *depth layer*, and *color*, we analyzed TCT and ER for reacting on instructions.

6.1.1 Task Completion Time

Figure 6 depicts the descriptive statistics for the TCT. A three-way ANOVA does not reveal significant different TCTs for the display mode, $F(1, 31)$=.678, p=.417, the depth layers, $F(2, 62)$=.307, p=.737, and the color, $F(1, 31)$=.543, p=.470. Moreover, we could not show statistically significant differences for the interactions of the tested factors, p>.102.

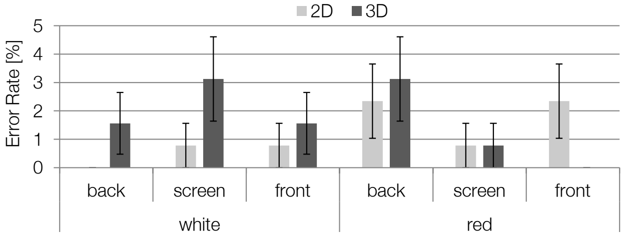

Figure 7: Means and standard errors as error bars for the error rates of confirming the instructions.

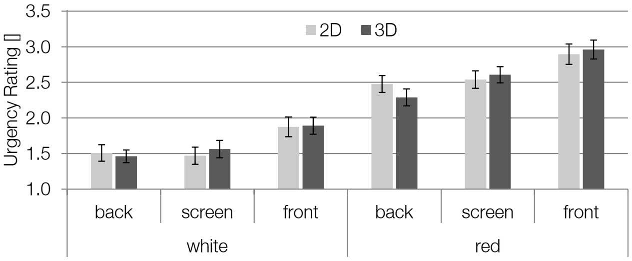

Figure 8: Means and standard errors as error bars for the urgency ratings.

6.1.2 Error Rate

In general, participants reacted correctly on the instructions. During all 1536 tasks, participants made in total 22 errors (Figure 7). Since the data of the error rates is not normally distributed, we used non-parametric tests for analyzing the main effects. Wilcoxon tests do not reveal significant results for display mode, Z=-1.155, p=.248, or color, Z=$-.484$, p=.629. Moreover, a Friedman test does not show a statistically significant difference for the three tested depth layers, X^2=.041, p=.980.

6.2 Perceived Urgency

For each instruction participants intuitively rated the urgency solely based on the visualization (4-Point Likert scale). Figure 8 shows that the rating increases for the red colored icons and for instructions with negative parallaxes while the display mode does not impact the rating (note that the instructions used the same stereo effect in 2D as in 3D). A three-way ANOVA shows significant differences for color, $F(1, 31)$=67.873, p<.001, η^2=.686, and the depth layers, $F(2, 62)$=16.884, p<.001, η^2.353. The display modes do not have a significant influence, $F(1, 31)$= .001, p<.972. Moreover, we found interaction effects for display mode × depth layer, $F(2, 62)$=6.436, p=.003, η^2= .172, and color × depth layer, $F(2, 62)$=3.706, p=.030, η^2=.107. The interactions display mode × color, $F(1, 31)$=.667, p=.420, and display mode × depth layer × color, $F(2, 62)$=2.172, p= .122, are not statistically significant.

6.3 Questionnaires

6.3.1 AttrakDiff

In the 3D version of the IC all dimensions of the AttrakDiff increase (Figure 9). Paired samples t-tests prove that these differences are significant for hedonic quality (HQ), $T(31)$=-5.015, p<.001, r=.448 and attractiveness (ATTR), $T(31)$=-4.425, p<.001, r=.387 but not for pragmatic quality (PQ), $T(31)$=$-.944$, p=.352.

6.3.2 DALI

The global DALI score is slightly higher for the 3D IC compared to its 2D variant, as Figure 9 shows. However, this tendency is not statistically significant, $T(31)$=2.239, p=.302.

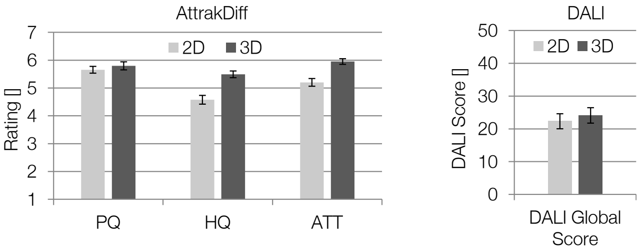

Figure 9: Means and standard errors as error bars for the dimensions of the AttrakDiff (left diagram) and the global score of the DALI (right diagram).

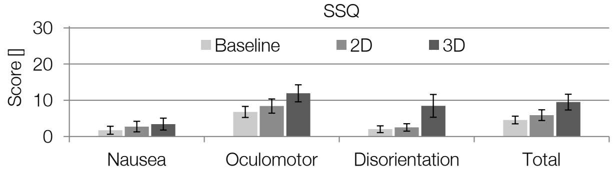

Figure 10: Means and standard errors as error bars of the ratings for the three dimensions and the global score of the SSQ.

6.3.3 SSQ

The participants filled in the SSQ questionnaire before the study started (serving as baseline) and after each of the two test drives. The scores of each dimension (Nausea, Oculomotor, Disorientation) and the total score are in general very low[3] (cf. Figure 10). All values are lowest for the baseline measurement while the 3D variant of the IC has a negative impact on all dimensions and the total score compared to its 2D counterpart. Since the data is not normally distributed we used non-parametric tests for a statistical analysis. Friedman tests show that the differences between the measures are not significant for Nausea, $X^2(2)=5.375$, $p=.068$, oculomotor, $X^2(2)=3.309$, $p=.191$, and the total score $X^2(2)=5.233$, $p=.073$. In contrast, testing the dimension disorientation reveals statistical significances, $X^2(2)=7.154$, $p=.028$. However, pairwise comparisons using Wilcoxon Tests with a Bonferroni corrected alpha level ($\alpha=.017$) do not reveal significant results for comparing the baseline with 3D, $Z=-2.226$, $p=.026$, the baseline with 2D, $Z=-.447$, $p=.655$, and 2D with 3D $Z=-2.047$, $p=.041$.

6.3.4 Gaze Behavior and Information Structure

Figure 11 shows that the statements about gaze behavior are in favor of the 3D version. However, Wilcoxon tests show that there are no statistical differences between the 2D and 3D IC version for the ratings about switching the gaze, $Z=-.599$, $p=.549$, the gaze duration, $Z=-1.574$, $p=.116$, as well as the gaze frequency, $Z=-1.048$, $p=.295$. Nevertheless, 3D significantly increases the rating for the statement about the clarity of the displays structure, $Z=-2.810$, $p=.005$, $r=.497$.

6.4 Qualitative Feedback

6.4.1 General Impressions

28 people out of 32 would rather use an S3D IC, while 2 participants (P32, P15) chose the 2D representation and 2 other persons (P17, P27) could not make an explicit decision. It has been 'faster' (P32) and 'easier' (P15) to read from the 2D representation. Furthermore, the 2D version provides a clearer functional overview (P27, P15). Most comments made on the IC are positive for the S3D representation. It is more attractive in an aesthetic way. In addition, the S3D IC provides more innovative and impressive features. Participants also preferred the S3D IC due to its naturalness (P20, P25), high quality (P11, P22) and creativity (P20, P25). The 2D IC representation was described as usual (P5, P17, P32, P15), but appears rather boring when compared to the S3D IC (P2, P25). 2 participants think that the difference between 2D and S3D is very small. P17 felt that only particular components were more suitable for S3D representation. All participants stated that they could imagine to use an S3D IC in their own car. Nevertheless, some participants reported to have problems with reading properly from the S3D IC. 4 persons needed more time to focus on elements in S3D mode, and 3 people found it cumbersome to read from a stereoscopic representation. As disturbing factor some participants mentioned the dynamic behavior of the S3D presentation due to motion parallax and tracking problems which induce jitter. As we did not implement motion parallax, the dynamic appearance arises due to shear distortions, which is a typical artifact of S3D displays [24].

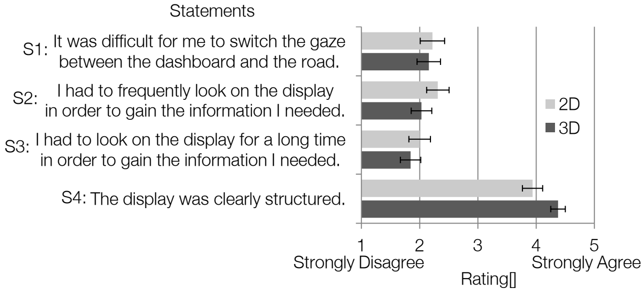

Figure 11: Means and standard errors as error bars of the ratings for the four statements (right diagram).

6.4.2 ACC and Navigation

Participants appreciated using S3D effects for the ACC (32/32). In particular, the S3D ACC was regarded as attractive, realistic and interesting. Spatial relation to the front vehicle has been effectively expressed through stereoscopic depth, thus enables intuitive understanding of the current distance and the driving situation. Nevertheless, 6 people verbalized that the S3D representation of ACC did not yield noticeable improvement of functionality.

31 out of 32 participants preferred the stereoscopic version of the navigation cues while one participant (P19) did not notice much difference. Only few comments were made on the aesthetic value of the S3D navigation, but people commented on the functional usability. Although nobody missed any navigation task, participants stated that the S3D representation of navigation cue has greatly contributed to the understanding of the real-world spatial relationship, thus enabling better estimation of the turning point.

6.4.3 3D Map and Infotainment Menu

Although not containing any real usable functionality, the map was described as a cool and attractive feature by 15 participants. Visualization depth- and spatial-related information is superior for the S3D display, as a better perspective and spatial perception is gained through its usage. The semi-structured interview also revealed high acceptance of the S3D map (31/32). One person (P15) stated that focusing on the stereoscopic 3D map takes more time and thus he preferred the 2D representation.

With regard to the infotainment list, the S3D version was favored by 24 participants. 7 persons would rather use the 2D list because it 'requires less effort for focusing' , 'is more familiar', 'simpler' and 'better readable'. Positive aspects attributed to the S3D list are a better interaction focus and improved item segregation.

---

[3] The maximum total score is 235.62

6.4.4 Information structuring

28 persons stated to have immediately perceived the structure of information shown within the IC. If it was not instantly recognized, the content structure within the IC context was explained to the participants. After clarification, 30 of 32 participants said that the information structure makes sense for them. 21 persons appreciated using depth to visually weight and communicate the importance level of information. For instance, placing the fuel and oil temperature gauge on a rear layer indicates their inferior importance to the speed display. Structuring by depth layers has relieved the display space, thus reduced information density. Still, the spatial structure of particular components could have been more sophisticated, for example, more detailed sub-structuring within the status bar, so that it gets decluttered or placing the speed gauge in a more central location instead of a low outside position.

7. DISCUSSION

In the following, we discuss the results of this real world experiment with regard to prior findings. In general, we obtained similar findings compared to previous simulator studies. Our results validate a significant increase of user experience and attractiveness due to S3D visualization which is consistent through lab [5, 7] and simulator studies [4]. Our data allows us to accept Hypothesis 1.

The DALI (driver activity load index) does not show a significant difference between the 2D and 3D representation of the IC. Moreover, participants felt that the S3D visualization does not have a significant influence on their gaze behavior. The findings regarding the gaze behavior as well as the primary driving task performance are in line with a former simulator study [4]. However, this simulator study revealed a significant increase of the DALI due to a S3D visualization. Our investigation in the real world could not verify this finding. The simulator study of Pitts et al. [18] showed that a stereoscopic visualization can decrease eyes-off-the-road times. Note that they applied a rather artificial task which benefits from a stereoscopic visualization. However, during the interviews some participants noted a decreased readability of 3D content. Further research needs to clarify if this decrease is a result of the used display technology or the S3D effect itself. We reject Hypothesis 2.

In contrast to a former simulator study [4], S3D shows a negative influence on SSQ ratings. Hence, we accept Hypothesis 3. Nevertheless, it is unclear which roles the used parallax settings as well as the used autostereoscopic display including the tracker performance play. Participants noted motion parallax to have a negative influence . Since the perceived motion parallax is a result of shear distortion [24] this stereoscopic artifact needs to be compensated. As this finding is not a result of any laboratory or simulator study, we claim that the dynamic motions resulting from driving through the real world are the reason for detecting this issue.

Although several lab and simulator studies demonstrated a significant increase in secondary task performance [4, 18, 22] due to stereoscopy our results show no difference in TCT or ER for the instruction tasks due to the display mode, color and depth layer. We claim that the uncontrolled real world situation hampers the sensitivity for identifying this effect. Nevertheless, the applied rating shows that the used depth layer as well as the color have an significant influence on the perceived urgency of the instructions. Thereby the color has a greater effect than the used depth layers. Nevertheless, positioning content in front of the screen significantly increases the perceived urgency. Hence, we accept Hypothesis 4.

The qualitative feedback of the participants yields an interesting insight on the navigation task. While participants emphasized the increased attractiveness of ACC due to S3D, their comments focused on the functional use of the navigation cues in S3D. One of the strengths of the S3D presentation is that it strongly matches to the real 3D environment. Since a simulator study can not reproduce the matching of a real 3D world with an artificial S3D interface this result is solely verifiable in a real world environment.

Qualitative feedback revealed that S3D strongly enhanced the usability of the infotainment list. Depth highlighted the focus of interaction, hence making use easier for the participants. This is in line with findings from a former lab study about an S3D infotainment menu [5]. In general, participants rated that S3D contributes to clarifying the IC structure. We hence accept Hypothesis 5.

8. REAL WORLD STUDY APPROACH

In the following, we summarize guidelines we gained throughout the development of our test vehicle and the planing and execution of the real world study. These simple guidelines allow to maximize safety for gathering data of high ecological validity.

Real World Studies for Validation The driving simulator is highly sensitive in identifying effects of secondary tasks on driver performance without putting the participants at risk. If new technologies reveal promising findings in simulator studies without a negative impact on driver distraction these technologies qualify for a validation in the real world. We claim that prior investigations in the lab as well as in the simulator are necessary before planing a study in the real world. We deliberately based our investigation on former studies conducted in the lab and the simulator.

Minimize Driver Workload Driving simulator environments allow for a safe evaluation of secondary tasks, even if inducing high cognitive workload. There are approaches that also integrate a tertiary task such as a peripheral detection task to measure workload. Since these tasks represent a further source of distraction we argue to neglect those methods for investigating novel in-car display or interaction technologies in real world studies. There is an increased need to design simple secondary tasks to minimize driver distraction. In consequence, we based our study on an instruction task already used in a former simulator study in combination with a depth judgment task [4]. We dropped the additional depth judgment task to decrease the complexity of the secondary tasks. It is also advisable to minimize the complexity of the driving task by choosing an appropriate route excluding difficult junctions and dense traffic as well as using a car with automatic transmission.

Focus on Subjective Data As real world environments do not allow for sufficient control of confounding factors, the sensitivity of objective data on primary and secondary task performance is rather low. The results of our study confirm this finding. Objective measures are highly valuable in fully controlled lab and simulator studies while real world studies offer the possibility of collecting subjective data of high ecologic validity. Hence, we applied a subjective rating while the participants conducted the driving task. As the rating refers to a single feature, in our case urgency, and uses a simple 4-Point scale, participants had no problems in attributing their rating while driving. We even had the impression that ratings are made very intuitively. From this we learn that our approach is meaningful for a real world study.

Provide backup for primary task-relevant information Evaluating systems that carry important driving information such as, for example, speed and warnings, need to be fully reliable. If this is not the case due to a prototypical implementation there need to be systems that reliably communicate this information. In our case we use a test vehicle equipped with a HUD display that shows important information as a fallback in the case of any errors.

Manually trigger tasks During the study one experimenter accompanied the participant. The experimenter took the seat of the co-passenger to optimally monitor the traffic situation. In this way, the experimenter could alert and support the driver in critical traffic situations. Moreover, the experimenter triggered the secondary tasks so that at no time safety-critical situations occurred.

Provide in-depth instructions Participants were extensively acquainted with the system and the study before starting the engine. In this way, participants got used to the new technology and the tasks. The participants practiced the tasks together with the urgency rating as long as they felt comfortable with this procedure. After that, they practiced the tasks once more while driving in a low traffic and at a low speed (30 km/h). As all participants had no problems in solving the tasks during the study this procedure optimally prepared the participants.

9. CONCLUSION

We presented a real-world driving study, investigating an in-car autostereoscopic display. Our research is based on former lab and simulator studies on in-car 3D displays. The findings validate that S3D increases the UX, clarifies the information structure of the display, and does not significantly affect driver workload. We found that S3D slightly increases discomfort in comparison to a 2D display. Moreover, we deliberately investigated the use of S3D to encode urgency. Our results show that color has a greater impact on communicating urgency than S3D, while the combination of both maximizes the perceived urgency. Our study provided detailed insights on the use of S3D in cars which are hard to find in the simulator. Besides shear distortions as a disturbing factor, S3D allows for an easy translation between the real and virtual 3D world.

Furthermore, we presented lessons learned on conducting real world driving studies for evaluating novel UI technologies in the car. First we suggest to use the real-world approach as a complementary validation method for exploring novel interaction technologies in cars. A necessary requirement is that prior laboratory evaluations ensures lower driver distraction compared to state of the art technologies. Moreover, we suggest to carefully choose the test track, the test vehicle, the secondary tasks, the study procedure as well as the required measures.

As a next step we plan to look into long term evaluation in the real world. In this way, more detailed insights about the daily use and novelty effects of new UI technologies can be obtained.

10. REFERENCES

[1] Alper, B., Hollerer, T., Kuchera-Morin, J., and Forbes, A. Stereoscopic highlighting: 2d graph visualization on stereo displays. *IEEE Trans. Vis. Com. Gr. 17*, 12 (2011).

[2] Alt, F., Kern, D., Schulte, F., Pfleging, B., Shirazi, A. S., and Schmidt, A. Enabling micro-entertainment in vehicles based on context information. In *Proc. AutomotieUI'10*, ACM (New York, NY, USA, 2010), 117–124.

[3] Blaauw, G. J. Driving experience and task demands in simulator and instrumented car: a validation study. *Human Factors 24*, 4 (1982), 473–486.

[4] Broy, N., Alt, F., Schneegass, S., and Pfleging, B. 3d displays in cars: Exploring the user performance for a stereoscopic instrument cluster. In *Proc. AutomotiveUI '14*, ACM (2014), 2:1–2:9.

[5] Broy, N., André, E., and Schmidt, A. Is stereoscopic 3d a better choice for information representation in the car? In *Proc. AutomotiveUI '12*, ACM (2012), 93–100.

[6] Broy, N., Schneegass, S., Guo, M., Alt, F., and Schmidt, A. Evaluating stereoscopic 3d for automotive user interfaces in a real-world driving study. In *CHI '15 EA*, ACM (2015), 1717–1722.

[7] Broy, N., Zierer, B. J., Schneegass, S., and Alt, F. Exploring virtual depth for automotive instrument cluster concepts. In *CHI '14 EA*, ACM (2014), 1783–1788.

[8] Burnett, G. E. On-the-move and in your car: An overview of hci issues for in-car computing. *IJMHCI 1*, 1 (2009), 60–78.

[9] Campbell, J., Richard, C., Brown, J., and McCallum, M. *Crash Warning System Interfaces: Human Factors Insights and Lessons Learned : Final Report*. NHTSA, 2007.

[10] Häkkilä, J., Posti, M., Ventä-Olkkonen, L., Koskenranta, O., and Colley, A. Design, implementation and evaluation of an autostereoscopic 3d mobile phonebook. In *Proc. MUM '14*, ACM (2014), 81–88.

[11] Häkkinen, J., Pölönen, M., Takatalo, J., and Nyman, G. Simulator sickness in virtual display gaming: A comparison of stereoscopic and non-stereoscopic situations. In *Proc. MobileHCI '06*, ACM (2006), 227–230.

[12] Ishihara, S. Test for color blindness. *Tokyo: Hongo Harukicho* (1917).

[13] Kuwayama, T. Random-dot stereogram. *Journ. Soc. of Phot. Sci. and Techn. Japan 72* (2011), 273–277.

[14] Landy, M. S., Maloney, L. T., Johnston, E. B., and Young, M. Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision research 35*, 3 (1995), 389–412.

[15] Leonard, S. D. Does color of warnings affect risk perception? *Int. Journ. of Ind. Erg. 23*, 5 (1999), 499–504.

[16] Nakayama, K., Silverman, G. H., et al. Serial and parallel processing of visual feature conjunctions. *Nature 320*, 6059 (1986), 264–265.

[17] Pfleging, B., and Schmidt, A. (non-) driving-related activities in the car: Defining driver activities for manual and automated driving. In *CHI WS on Exp. Aut. Veh.* (2015).

[18] Pitts, M. J., Hasedžić, E., Skrypchuk, L., Attridge, A., and Williams, M. Adding depth: Establishing 3d display fundamentals for automotive applications. Tech. Rep. 2015-01-0147, SAE International, 2015.

[19] Reed, M. P., and Green, P. A. Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task. *Ergonomics 42*, 8 (1999), 1015–1037.

[20] Santos, J., Merat, N., Mouta, S., Brookhuis, K., and De Waard, D. The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies. *Transportation Research 8*, 2 (2005), 135–146.

[21] Schild, J., LaViola, J., and Masuch, M. Understanding ux in stereoscopic 3D games. In *Proc. CHI '12*, ACM (2012).

[22] Szczerba, J., and Hersberger, R. The use of stereoscopic depth in an automotive instrument display evaluating user-performance in visual search and change detection. In *Proc. HFES*, vol. 58, SAGE Publications (2014), 1184–1188.

[23] Wogalter, M. S., Conzola, V. C., and Smith-Jackson, T. L. Research-based guidelines for warning design and evaluation. *Applied ergonomics 33*, 3 (2002), 219–230.

[24] Woods, A. J., Docherty, T., and Koch, R. Image distortions in stereoscopic video systems, 1993.

[25] Young, K., Regan, M., and Hammer, M. Driver distraction: A review of the literature. *Distracted driving. Sydney, NSW: Australasian College of Road Safety* (2007), 379–405.