DOI: 10.22032/dbt.39958

DOI (proceedings): 10.22032/dbt.39936

vdt Verband Deutscher Tonmeister e.V.

Full Reviewed Paper at ICSA 2019

Presented by VDT.

Exploring Audiovisual Support Systems for In-Car Multiparty Conferencing

F. Weidner[1], B. Fiedler[1,2], J. Redlich[1,2,3], and W. Broll[1]

[1] *TU Ilmenau, 98693 Ilmenau, Germany, Email: florian.weidner@tu-ilmenau.de*

[2] *Fraunhofer IDMT, 98693 Ilmenau, Germany, Email: bernhard.fiedler@idmt.fraunhofer.de*

[3] *ITA Weimar mbH, 99428 Weimar, Germany, Email: ita@ita-weimar.de*

Abstract

Calling while driving poses a severe safety risk. When more than two people are involved in a call - a conference call - this risk increases even more. Intelligent vehicles could offer support systems that ease the cognitive burden of such a multiparty calls. We explore the possibilities of such advanced driving assistant systems (ADAS) in two ways: first, we investigate object-based spatial audio where each remote caller is modeled as a distinct audio source. Second, we apply a non-intrusive ambient stereoscopic 3D (S3D) visualization that indicates the current speaker and its location. In a between-subject design driving simulator study (n=56), we assess workload, user experience and driving performance. Surprisingly, we found no positive effect of object-based audio. However, we present evidence how a supporting visualization might lower situational stress and increase the system's dependability. We conclude that a supportive and intelligent stereoscopic visualization is a promising candidate for enhancing multiparty conference calls while driving.

1. Introduction

People work in their cars while driving. They search through papers, make notes, check emails, or schedule meetings [18, 25]. It is likely and anticipated that advances in connectivity and automation will increase the time people work in cars and will make other tasks more likely to be performed [6, 21]. In addition to the mentioned work-related activities, people make business calls while driving [18, 25]. They call clients, secretaries, or colleagues to productively use the time they spent in their car. However, having a phone call while driving is, despite all technological advances and hands-free technology, a dangerous task. Leung et al. [19] concluded that making a cognitively demanding hands-free phone call while driving is as risky as driving with a blood alcohol concentration of 0.07 to 0.10%. Moreover, as soon as phone calls have multiple remote participants (multiparty conversation), cognitive load increases [29]. In particular, it gets harder to distinguish the speakers' voices and by that, to follow the conversation [8]. In a face-to-face meeting, it is easy to distinguish who is speaking: we see the people and hear their voices from where they are sitting or standing. In a remote multiparty conversation however, the sound sources that represent the communication partners usually come from the same direction. Additionally, there is no visible representation of the callers. Restoring both properties - spatial sound and visual representation - might help to lower the cognitive burden of such remote multiparty conversations and make them less risky.

Thus, this paper investigates two potential solutions to enhance the experience during multiparty conference calls while driving: an object-based spatial sound system and an adaptive visualization. The former enables the listener to locate each voice at the correct angle in the room. The latter provides a non-intrusive ambient stereoscopic 3D (S3D) dashboard visualization, adding visual representations to the voices, and enables the user to match the voice to a name.

The rest of the paper is structured as follows: We start with background information and related work (Section 2.1). Following, we present our experimental design and prototype in Section 3. Section 4 presents our results and Section 5 interprets and discusses them. Finally, we conclude and provide a brief outlook in Section 6.

2. Background & Related Work

2.1. Stereoscopic 3D

Traditional displays only show one image for the left and right eye. Depth is visible, but only via cues like occlusion or linear perspective. Stereoscopic 3D (S3D) visualizations display individual images for the left and right eye. The human visual system then fuses these images and calculates binocular depth cues for depth perception similar to real-life [9].

2.2. Channel- & Object-based Audio

Audio scenes usually consist of many channels mixed together by the audio engineer. For channel-based audio (CBA), the audio scene is down-mixed regarding standardized speaker layouts [8]. Since speaker layouts between common reproduction systems and the audio engineer's production or recording system usually do not match, the originally mixed sound field can not be reproduced exactly. This leads to a loss of spatial information (e.g. all sound sources come from one direction).

In contrast, object-based audio (OBA) does not store preproduced audio channels, but instead sound objects consisting of audio and metadata. Knowing the positions of loudspeakers, the sound field of the audio scene can be reproduced exactly by calculating each speaker signal independently in real time during playback [2, 22]. Hence, spatial information is retained (e.g. the direction of speakers during a multiparty conversation). Similar results can be achieved by applying a blind source separation algorithm that decomposes a multi-channel input stream into several output streams [23]. Although channel-based and object-based audio are both spatial audio formats, only the object-based approach enables the dynamic, real-time capable spatial location of sound sources. Therefore in this work only the object-based approach reproduces spatial audio.

2.3. Multiparty Conference Calls

Many research groups try to lessen the cognitive demand of calling while driving by exploring novel technologies like AR-glasses [15, 16] and using video-call software [14]. However, these systems have only been tested during one-to-one conversations.

For multiparty conference calls, Rajan et al. [26] try to solve this issue by using an intelligent user interface. It successfully performs, among other measures, speaker identification and adds presence indicators in form of personalized background noises to the conference call in a desktop setting. Kilgore reports a lower level of perceived difficulty when using spatial audio in a desktop audio conference system [13]. Their system increases the quality of the conference call but was not applied in the more safety-critical automotive context. No significant benefit of spatial audio was found by Inkpen et al. [10]. However, they found that displaying a visual indicator in form of spatial video improved people's ability to follow the conversation. It is important to note that the application of visual indicators during manual driving has to be handled carefully because the main focus of the driver should be on the traffic and surroundings. Wickens' theory of shared resources argues against providing visual cues while driving because driving is a highly visual task and the resources for processing visual information should be allocated for the driving and not for a secondary task [32]. However, recent advances in ambient lighting (e.g. Loecken et al. [20]) provide an interesting possibility to implement a non-intrusive vision-based support system for multiparty conference systems. Also, previous work on stereoscopic 3D dashboard visualizations showed that they do not necessarily decrease driving performance if designed carefully - especially for change detection tasks [30, 31]. Furthermore, S3D can improve user experience while driving when designed carefully [3, 4].

The high mental workload induced by phone calls motivated us to explore spatial audio and stereoscopic 3D visualizations as support systems during such calls more closely.

Tab. 1: Experimental design with four groups.

	No visualization	Stereoscopic 3D
Channel-based audio	CBA-NoV	CBA-S3D
Object-based audio	OBA-NoV	OBA-S3D

3. Study

Our study focuses on multiparty conference calls. The situation resembles a remote job interview where several employees of a company talk with a potential new employee who is currently driving on a highway.

3.1. Experimental Design

The experiment had a between-subject design. Dependent variable is the type of user interface. Participants experienced one of the four user interfaces as indicated in Table 1. Independent variables are described in Section 3.4.

In *CBA-NoV*, the three speakers' voices come from the same direction in front of the user and no supporting visualization was offered. This condition acts as the baseline. In *OBA-S3D*, the three speakers' voices can be perceived from distinct directions (left, center, and to the right of the driver) while supporting S3D visualization was offered.

We chose a between-subject design for two reasons: first, to shorten experiment duration. Second, a mixed design would have required a second simulated conversation. It turned out to be a challenging task to design two similar conversations that induce the same workload and therefore allow a comparison of all conditions, but do not repeat themselves. Hence, we applied the between-subject design that requires more participants but allows easy comparisons across groups.

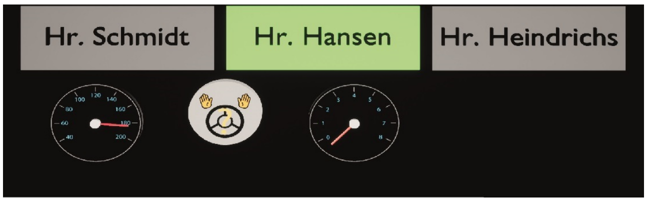

Fig. 1: The user interface adapts to the number of participants. The speaking person is highlighted using a green color. Location of tags matches position of speakers in auditory scene for OBA-conditions.

Fig. 2: Default view of the driving simulation.

3.2. Sample

The final sample consisted of $N = 56$ participants (male = 38 or 67.9%, female = 18 or 32.2%) with a mean age of 30.5 years ($SD = 8.22$). All possessed a valid driving license, had no hearing impairments, had normal or corrected-to-normal vision, and passed a stereopsis test if they were assigned to a S3D group (Random Dot Stereogram, [28]). 45 out of 56 (80.36%) had experience with stereoscopic displays and 26 (46.43%) had experience with 3D Audio. Overall, 31 participants (55.36%) said that they regularly make phone calls while driving. Mean MSSQ score is 9.53 ($SD = 8.62$).

3.3. Apparatus

Figures 2 and 3 illustrate the simulation environment. A screen (2D, 3.6×2.26 m, 2560×1600 pixel at 120 Hz) displays the driving environment. The car mock-up has an integrated spatially augmented reality dashboard (L-shaped, 90×60 cm, 2560×1600 pixel at 120 Hz, in S3D mode: 60 Hz per eye). It is further equipped with a Thrustmaster TX Racing Wheel Leather Edition and pedals. Volfoni Active 3D glasses[1] enable the stereoscopic visualization. An Optitrack tracking system [2] realizes head tracking and the head-coupled perspective. The rear-view was visible via a real car mirror and a 1920×1080, 30 Hz projection. All visualizations were realized using Unreal Engine 4.18.3 [3]. Field of view was set to 70°.

Figure 1 illustrates the user interface supporting the multiparty conversation. The name tags appear one by one according to the callers in the multiparty conversation. The name tags and colors are carefully calibrated to be non-intrusive and non-glaring while driving. They are located at the top of the dashboard so that they can act as an ambient lighting cue. Disparity of the name tags is $D = 0,218°$. That means that they appear 4 cm behind the projection plane with a viewing distance of 80 cm and and interpupillary distance of $IPD = 63$ mm. Disparity of the speedometer, tachometer, and cruise control indicator is $D = 0°$. Stereoscopic 3D was chosen as an additional cue for the ambient visualization.

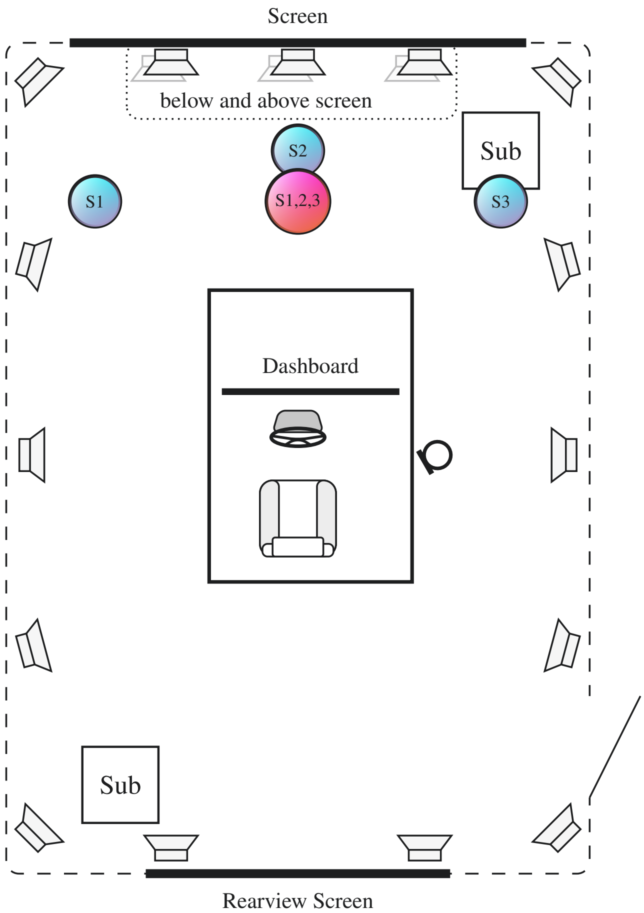

Fig. 3: Overview of simulation environment. Positions of dialogue partners marked for OBA in blue and CBA in red.

Locations of the name tags correspond to the intended seat distribution of the dialogue partners. In the OBA-condition, speakers' voices were located at the intended position relative to the driver. In the CBA-conditions, all three voices come from the same direction directly in front of the user resembling mono audio playback.

To reproduce audio, the wave field synthesis based 3D audio system SpatialSound Wave[4] (SSW) using 18 Seeburg TS-nano speakers and two Seeburg TSM subwoofers was installed. The speakers are positioned in a horizontal plane around the listening position at a height of 1.4 m. See Figure 3 for speaker positions. The front speakers are positioned above at 2.5 m and under the screen at ground level not blocking view to the screen. Using delay and gain adjustments, the sound sources can still be perceived in the horizontal plane.

[1] http://volfoni.com/, 2019-01-03
[2] http://optitrack.com/, 2019-01-03
[3] http://unrealengine.com/, 2019-01-02

[4] https://www.idmt.fraunhofer.de/en/institute/projects-products/spatialsound-wave.html, 2019-01-30

The audio system was calibrated and equalized to compensate room acoustics and speakers' characteristics.

Synchronization between visualization using Unreal Engine and auralization using SSW was implemented by exchanging scene data graphs as XML via UDP. By that, all sound sources (environment, other cars, engine and tires of ego car, wind noise and voices of the speakers) are managed in real-time and positioned in 2D space around the driver. Audio reproduction also included doppler effect of passing objects like cars.

3.4. Measures

To describe the sample, we apply the motion sickness susceptibility questionnaire (MSSQ, Golding et al. [7]). We further ask for simulator sickness using the Simulator Sickness questionnaire (SSQ, Kennedy et al. [12]) and for user experience using the user experience questionnaire (UEQ, Laugwitz et al. [17]). In order to assess workload, we apply the Driver Activity Load Index (DALI, Pauzie et al. [24]). The DALI is a modified version of the NASA TLX [**?**], especially designed to assess workload of drivers. Driving performance was operationalized using number of steering reversals and number of lane departures (measured using widest part of the vehicle) [27]. We count a steering wheel reversal when the driver turns the steering wheel 6 degrees in one direction and within 2 seconds 6 degrees in the opposite direction. We count a lane departure when the car crosses the center of a lane marking without making a lane change.

3.5. Procedure

When participants arrived, they filled out a consent form, the MSSQ, and a general questionnaire. They were randomly assigned to one of the four groups. They took a seat in the driving simulator and received general information about the phone interview, the car, and its capabilities. Depending on their assigned group, participants experienced different audiovisual user interfaces (c.f. Section 3.1). However, all wore stereoscopic 3D glasses. The virtual car was equipped with cruise control and participants were instructed how to use it. They were further told to respect traffic rules. In our study, participants had to drive along a 12 km curved three-lane highway with low traffic (approximately 5 cars per km; primary task) and engage in a multiparty-conversation (secondary task). After about 3 km, which were considered training, the phone rang. When participants had accepted the call with a button on the steering wheel, a simulated job interview started. In our scenario, the participant had previously applied for an internship at the fictional "Institute for Thermodynamics". Three virtual members of the institute took part in the interview. For playing questions, answers, and other sounds (e.g. agreeing sounds) of the callers, we used a wizard-of-oz-based audio player for multiparty conversations[5] which was controlled by an operator from another room. Approximately 10 seconds after they had hang up, the drive ended. After that, they filled out the UEQ, SSQ, and DALI. That concluded the experiment. Total experiment duration was approximately 25 minutes. Participants were not paid, but had the chance to win 50 Euro.

[5]https://github.com/JoReIMT/Dialogue-Sample-Player, Dialogue-Sample-Player, GPL-3.0, 2019-03-01

4. Results

All data follows normal distribution (tested with Shapiro-Wilk test, histograms, and QQ-plots, [33]) and shows homogeneity (tested with Levene's test) if not stated otherwise. An α-value of 0.05 was used as significance criterion when necessary. Data was analyzed using R 3.5.2 (afex 0.22.1, fBasics 3042.89, bestNormalize v1.3.0, and MASS 7.3-51.1).

The overall drive lasted about 6 minutes and 4.97 seconds ($SD = 75.07$ seconds;), depending on the answers given by the participants. Mean driving speed was $M = 116.47$ km/h ($SD = 8.00$ km/h). Participants drove on average 11.81 km.

4.1. Simulator Sickness

Data of the SSQ is not normal distributed. Hence, we applied a Yeo-Johnson transformation to correct for normality [34]. A two-way ANOVA found no evidence for significant differences in simulator sickness on Yeo-Johnson transformed data of the SSQ (untransformed data: $M_{Nausea} = 18.91; SD = 17.59; M_{Oculomotor} = 18.27; SD = 15.55; M_{Disorientation} = 14.42; SD = 21.21; M_{Total} = 20.30; SD = 17.51; p > .340$).

4.2. User Experience

Results of the UEQ are presented in Figure 4. Results of a two-way ANOVA suggest that there is a significant interaction effect on the *Attractiveness* scale ($F(1, 52) = 6.82$, $p = .012$, $\eta_p^2 = .12$). Closer inspection using Tukey's HSD confirms that CBA-S3D ($M = 0.88$ $SE = 0.09$) was perceived significantly more attractive than CBA-NoV ($M = 0.46$ $SE = 0.08$; $t(52) = 2.858$, $p = .0302$, Cohen's $d = 1.26$) as Figure 5 indicates. There is also a significant main effect of the video condition on *Dependability* stating that the S3D condition was perceived more dependable than the NoV condition (S3D: $M = 0.60$, $SE = 0.12$; NoV: $M = 0.24$, $SE = 0.12$; $F(1, 52) = 4.31$, $p = .043$, $\eta_p^2 = .08$). No other main or interaction effects were found.

4.3. Driver Activity Load Index

Results of the DALI questionnaire are shown in Figure 6. Data is not normal distributed and data transformations do not lead to a reasonable normal distribution. Hence, we decided to analyze data using multiple Wilcoxon Rank Sum tests. Because of the large standard deviations and rather small sample size, we deliberately did so without correction for multiple comparisons. This allows for data exploration, detection of possible effects, and making suggestions for future research. However, interpretation of any significant differences requires taking into account this very liberal procedure.

On the *Visual Demand* scale, there is a significant difference between OBA-NoV and OBA-S3D ($W = 52, p = .035, r = .281$. indicating that OBA-NoV ($Mdn = 48.5$) was less visually demanding that OBA-S3D ($Mdn = 15.0$). On the *Situational Stress* scale, we found evidence for a significant main effect of visualization between S3D and NoV with $W = 266.5, p = .394, r = .276$ suggesting that the S3D condition ($Mdn = 19.0$) leads to less stress than the NoV condition ($Mdn = 39.5$). For all other comparisons, test results are non-significant with $p > .079$. This tells us that the other sub-scales measured by the DALI do not significantly differ between groups and conditions.

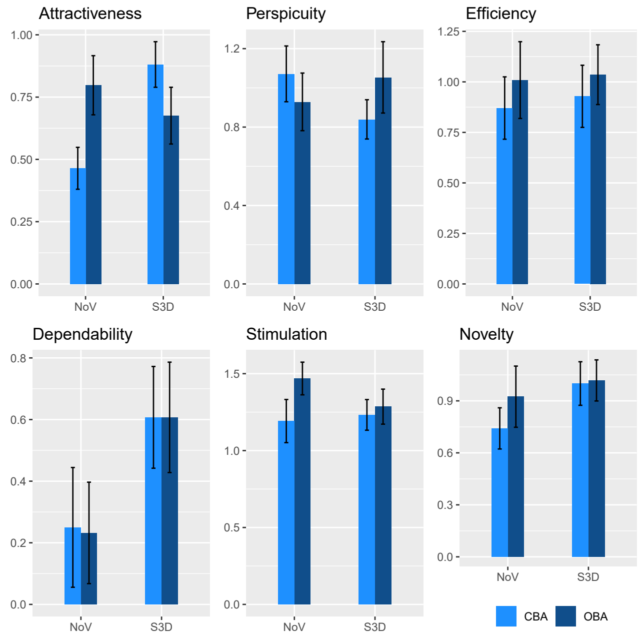

Fig. 4: Means and standard errors for the UEQ ([-3;3], higher is better).

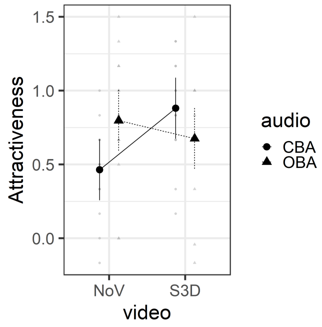

Fig. 5: Interaction plot for "Attractiveness" of the UEQ. CBA-S3D was perceived more attractive than CBA-NoV

Because some authors suggest that the ANOVA is very robust to violations of the normality assumptions (e.g. Field et al. [5]), we ran a two-way ANOVA for the two found effects. The more conservative ANOVA confirms the main effect of the visualization on *Stress* ($F(1, 52) = 4.75$, $p = .034$, $\eta_p^2 = .08$) but not the interaction effect on *Visual Demand* ($F(1, 52) = 3.30$, $p = .075$, $\eta_p^2 = .06$).

4.4. Driving Performance

4.4.1. Number of Steering Wheel Reversals

We applied a Tukey transformation to establish normal distribution. A two-way ANOVA found no significant effects in steering wheel reversals ($p > .406$). By that, we can assume that the different audiovisual support systems did not significantly influence this measure of driving performance.

4.4.2. Number of Lane Departures

Data on number of lane departures does not follow a normal distribution so we applied a Tukey transformation. Again, a two-way ANOVA found no significant effects in the Tukey-transformed number of steering wheel reversals ($p > .582$). This indicates that there is insufficient evidence for an influence of the support systems on the number of steering wheel reversals.

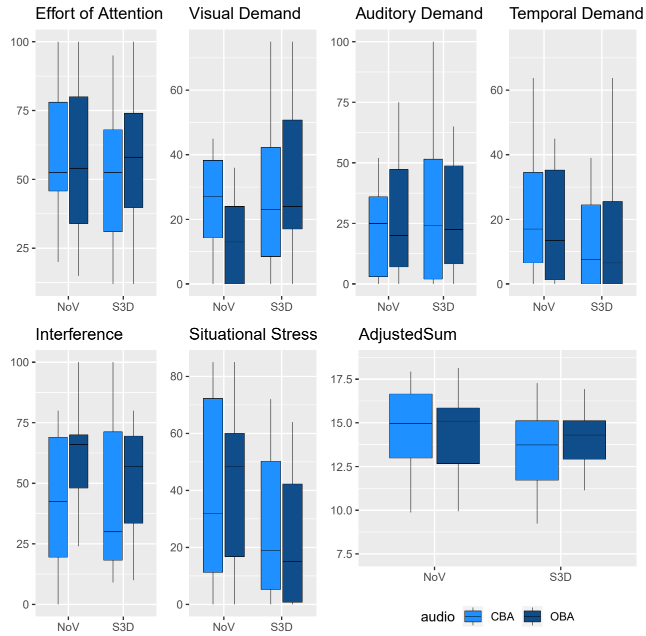

Fig. 6: Medians for the scales of the DALI ([0;100], lower is better).

5. Discussion

We explored two support systems: an object-based audio which emits the speakers' voices from distinct directions and a stereoscopic 3D dashboard visualization providing indications who is speaking at the moment via non-intrusive ambient lighting and name tags. We expected that each support system on its own leads to less workload and higher user experience (UX). We further expected that the combination of both support systems outperforms each support system regarding workload and UX.

Results indicate that participants' responses to the different support systems are very conservative. We did not find a clear favorite with respect to user experience or workload.

Prior research indicates that spatial sound makes it easier for people to follow a conversation and reduces perceived difficulty in multiparty conversations [1]. In our study, spatialized audio reproduction did not significantly improve driving performance or reduce the drivers' workload during phone conference. Whereas prior research on multiparty conversation was done without any ambient noise [1, 29], the complete auditory scene including environmental sounds was reproduced with spatial audio in our study. Maybe the differences between spatialized speech and spatialized noise were not prominent enough to affect the measures positively.

Nevertheless, results indicate that perceived *Situational Stress* is lowered by the S3D support system. As mentioned in Section 4, this is based on a liberal test procedure. However, it acts as a strong motivation for further research in this area.

SAE International mentions values for comparison of lane departures as guidance ranging from 7.1 to 16.4 lane departures per 100 miles and reversal rates of 1 to 2 per minute - without any distracting tasks [27]. The measured number of steering wheel reversals falls within this range. However, measured number of lane departures in our study was approximately twice as high - regardless of condition and with a mean distance of only 11.81 km, meaning that participants drove serpentine-like but without abruptly moving the steering wheel. On the one hand, it is important to note that the experiment was performed in a medium fidelity simulator with naturally unrealistic vehicle physics. On the other hand, this suggests that hands-free multiparty phone conversations are very demanding and impair driving performance. Another point that could have potentially impaired driving performance is that all participants wore 3D glasses all the time. Overall, we did not uncover any main or interaction effects on driving performance suggesting that the support systems did not negatively affect driving performance.

However, the S3D visualization was perceived more *dependable* (UEQ, Figure 4) than the NoV-condition, regardless of the audio condition. Further, we did not measure any driving performance impairments of the visualization. That means that participants felt more in control and perceived the UI as more secure and predictable (c.f. Laugwitz et al. [17]) when the system presented visual indicators of who is speaking compared to when they had to rely on audio information only. This is supported by the mentioned effect on the *Situational Stress* scale of the DALI. A more dependable system might induce less stress. It is likely that the more obvious visual support system made people feel more certain in who is speaking and by that, made it easier to follow the conversation. Also, acting as an ambient lighting cue, participants can perceive the information in their peripheral vision without taking the eyes of the road. The presented names further enhance the experience by adding meta information. This is confirmed by responses which indicate the necessity of research on UIs containing additional information like the role of the speaker, faces or avatars, and locations during international phone conferences.

5.1. Limitations

Data suggests that the sample size was too small to generalize results. Follow-up experiments with more participants are necessary to confirm our conclusions. We missed to ask participants how often they engage in multiparty conversations. Considering that these support systems enhance conference calls while driving, another study with a sample that is very likely to often engage in such calls (unlike students and university staff in our study) is necessary to confirm results of our exploratory study. The experimental design did not investigate the full spectrum of visualizations like HUD or windshield displays as well as traditional 2D displays. While S3D is not crucial for this study and acts only as a design element, we chose S3D to explore the application domain of this visualization technique, in particular for structuring information. Further research is necessary to put the results in perspective to traditional 2D visualizations.

It is important to note that mental workload can vary substantially due to the type of conversation. Involvement and engagement is another factor that can influence workload. In our study, participants might not have cared much about the conversation which influences final workload results.

Participants reported simulator sickness scores ranging from significant to problematic symptoms according to Kennedy et al. [11]. Taking a closer look at the scores reveals that the symptoms "difficulty concentrating" ($\sum_{n=0}^{N} = 49$) and "sweating" ($\sum_{n=0}^{N} = 29$) were reported especially high among participants compared to the average scores of the other symptoms ($\sum_{n=0}^{N} = 8.9$). The first one is likely due to exhaustion. Participants needed a high effort of attention and reported high situational stress especially because of the simulated job interview. The latter is most likely due to the warm laboratory environment without appropriate air conditioning but many projectors, workstations and loudspeakers.

We need to mention that the used spatial sound setup is different from a typical setup encountered in a car. Especially the acoustics in the laboratory and a common car lead to differences in perceived proximity of audio sources and by that, the virtual speakers. Hence, replicating our experiment in a real car might lead to different results.

Also, being a simulator study, vehicle physics as well as overall replication of real-world conditions is limited by technology. Hence, results can only be interpreted within the context of our simulation. The absence of a baseline drive without any conversational task makes it hard to specify the reason for the impaired driving performance. However, comparing our results with previous work suggests that doing a multiparty call - e.g. a job interview - can lead to impaired driving performance. Hence, our data indicate that it is not recommended to engage in a multiparty conversations for manual driving with SAE Level 1 and 2 automation - even with support systems like ours. We suggest exploring the proposed support systems for higher levels of automation, e.g. Level 3 where participants have to monitor the environment but do not have to engage while automation is enabled.

6. Conclusion

In this work, we investigated two audiovisual support systems and their potential to make multiparty conferencing while driving more pleasant: a stereoscopic visualization and an object-based spatial auralization. We found no positive - but also no negative - impact of object-based audio in this context. This might be due to the spatialized noise reproduction. We propose to conduct further experiments of multiparty conferencing in realistic acoustic environments to verify if it generally reduces the benefits of using spatialized speech reproduction. However, results of our study indicate that a visualization presenting information about the callers potentially reduces perceived stress and is likely to increase attractiveness. Participants further perceived the conditions with a supporting stereoscopic 3D visualization as more dependable compared to user interfaces without a vision-based support system.

Considering the positive aspects of our prototype paired with the absence of any impairments, an intelligent visualization has a lot of potential to support drivers during conference calls. Our results can be the basis for a user interface that offers more security, predictability, and makes users feel more in control during such phone calls. In our fast-paced society, where workplace availability and flexibility but also constantly keeping in touch with friends and family is of utmost importance, such a system could be highly beneficial for user experience and safety.

Since the impact on driving performance of such systems, the applied display, and the application in (semi-)autonomous vehicles are important aspects in this research domain, follow-up experiments with slightly modified designs that integrate these factors are planned to investigate this areas further. Especially a test with people who regularly participate in conference calls, the comparison with other types of displays (HUD, windshield, and perspective 3D), and the application of in vehicles with Level 3 automation seem promising.

7. Acknowledgements

This work has partially been funded by the Free State of Thuringia, Germany (FKZ: TUI-I-01-14 and 2015-FE-9109).

8. References

[1] AHRENS, J., GEIER, M., RAAKE, A., AND SCHLEGEL, C. Listening and conversational quality of spatial audio conferencing. In *Audio Engineerig Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space* (Oct 2010).

[2] BERKHOUT, A. J. A holographic approach to acoustic control. *J. Audio Eng. Soc 36*, 12 (1988), 977–995.

[3] BROY, N., ANDRÉ, E., AND SCHMIDT, A. 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 - AutomotiveUI '12* (2012).

[4] BROY, N., GUO, M., SCHNEEGASS, S., PFLEGING, B., AND ALT, F. Introducing novel technologies in the car - Conducting a Real-World Study to Test 3D Dashboards. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15* (New York, New York, USA, 2015), ACM Press, pp. 179–186.

[5] FIELD, A. P., MILES, J., AND FIELD, Z. *Discovering statistics using R.* Sage, London, 2014.

[6] FRÖHLICH, P., SACKL, A., TRÖSTERER, S., MESCHTSCHERJAKOV, A., DIAMOND, L., AND TSCHELIGI, M. Acceptance Factors for Future Workplaces in Highly Automated Trucks. In *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '18* (New York, New York, USA, 2018), ACM Press, pp. 129–136.

[7] GOLDING, J. F. Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. *Brain research bulletin 47*, 5 (nov 1998), 507–16.

[8] HOLMAN, T. Surround Sound (Second Edition). 2nd ed. Focal Press, Oxford, 2008.

[9] HOWARD, I. P., AND ROGERS, B. J. *Binocular Vision and Stereopsis.* Oxford University Press, New York, New York, USA, 2008.

[10] INKPEN, K., HEGDE, R., CZERWINSKI, M., AND ZHANG, Z. Exploring spatialized audio & video for distributed conversations. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10* (New York, New York, USA, 2010), ACM Press, p. 95.

[11] KENNEDY, R. S., DREXLER, J. M., COMPTON, D. E., STANNEY, K. M., LANHAM, S., AND HARM, D. L. Configural scoring of simulator sickness, cybersickness and space adaptation syndrome: Similarities and differences?

[12] KENNEDY, R. S., LANE, N. E., BERBAUM, K. S., AND LILIENTHAL, M. G. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology 3*, 3 (jul 1993), 203–220.

[13] KILGORE, R., CHIGNELL, M., AND SMITH, P. Spatialized Audioconferencing: What Are the Benefits? In *Proceedings of the 2003 Conference of the Centre for Advanced Studies on Collaborative Research* (2003), CASCON '03, IBM Press, pp. 135–144.

[14] KUN, A. L., AND MEDENICA, Z. Video Call, or Not, That is the Question. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 2012), CHI EA '12, ACM, pp. 1631–1636.

[15] KUN, A. L., VAN DER MEULEN, H., AND JANSSEN, C. P. Calling While Driving: an Initial Experiment With Hololens. *PROCEEDINGS of the Ninth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (2017), 200–206.

[16] KUN, A. L., VAN DER MEULEN, H., AND JANSSEN, C. P. Calling while Driving using Augmented Reality: Blessing or Curse? (accepted for publication. *Presence: Teleoperators and Virtual Environments* (2018), 1–28.

[17] LAUGWITZ, B., HELD, T., AND SCHREPP, M. Construction and Evaluation of a User Experience Questionnaire. 2008, pp. 63–76.

[18] LAURIER, E. Doing Office Work on the Motorway. *Theory, Culture & Society 21*, 5 (2004), 261–277.

[19] LEUNG, S., CROFT, R. J., JACKSON, M. L., HOWARD, M. E., AND MCKENZIE, R. J. A Comparison of the Effect of Mobile Phone Use and Alcohol Consumption on Driving Simulation

Performance. *Traffic Injury Prevention 13*, 6 (2012), 566–574.

[20] LÖCKEN, A., MÜLLER, H., HEUTEN, W., AND BOLL, S. AmbiCar: Towards an in-vehicle ambient light display. *AutomotiveUI 2013*, October (2013), 107–108.

[21] MAURER, M., CHRISTIAN, G., LENZ, B., AND WINNER, H. *Autonomous Driving - Technical, Legal and Social Aspects*. 2016.

[22] MELCHIOR, F. *Investigations on spatial sound design based on measured room impulse responses*. PhD thesis, Delft University of Technology, Delft, Niederlande, 24.06.2011.

[23] PAL, M., ROY, R., BASU, J., AND BEPARI, M. S. Blind source separation: A review and analysis. In *2013 International Conference Oriental COCOSDA held jointly with 2013 Conference on Asian Spoken Language Research and Evaluation (O-COCOSDA/CASLRE)* (nov 2013), pp. 1–5.

[24] PAUZIE, A. A method to assess the driver mental workload: The driving activity load index (DALI). *IET Intelligent Transport Systems 2*, 4 (2008), 315.

[25] PERTERER, N., MOSER, C., MESCHTSCHERJAKOV, A., KRISCHKOWSKY, A., AND TSCHELIGI, M. Activities and Technology Usage while Driving: A Field Study with Private Short-Distance Car Commuters. *Proceedings of the Nordichi '16: the 9Th Nordic Conference on Human-Computer Interaction - Game Changing Design* (2016).

[26] RAJAN, R., CHEN, C., AND SELKER, T. Considerate Audio MEdiating Oracle (CAMEO). In *Proceedings of the Designing Interactive Systems Conference on - DIS '12* (New York, New York, USA, 2012), ACM Press, p. 86.

[27] SAE INTERNATIONAL. Operational Definitions of Driving Performance Measures and Statistics J2944_201506, jun 2015.

[28] SCHWARTZ, B. L., AND KRANTZ, J. H. Random Dot Stereograms, 2018.

[29] SKOWRONEK, J., AND RAAKE, A. Investigating the effect of number of interlocutors on the quality of experience for multi-party audio conferencing. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, August (2011), 829–832.

[30] SZCZERBA, J., AND HERSBERGER, R. The Use of Stereoscopic Depth in an Automotive Instrument Display: Evaluating User-Performance in Visual Search and Change Detection. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 58*, 1 (sep 2014), 1184–1188.

[31] WEIDNER, F., AND BROLL, W. Exploring Large Stereoscopic 3D Dashboards for Future Automotive User Interfaces. In *Proceedings of the 9th International Conference on Applied Human Factors and Ergonomics AHFE*. Springer, Cham, Orlando, FL, 2019, pp. 502–513.

[32] WICKENS, C. D. Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science 3*, 2 (jan 2002), 159–177.

[33] YAP, B. W., AND SIM, C. H. Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation 81*, 12 (2011), 2141–2155.

[34] YEO, I.-K., AND JOHNSON, R. A. A New Family of Power Transformations to Improve Normality or Symmetry. *biometrika Biometrika 87*, 4 (2000), 954–959.