Travelers’ information need in automated vehicles – a psychophysiological analysis

Charlotte Brandebusemeyer
Institute of Cognitive Science
University of Osnabrück
Osnabrück, Germany
cbrandebusem@uos.de

Klas Ihme
Institute of Transportation Systems
German Aerospace Center (DLR e.V.)
Braunschweig, Germany
klas.ihme@dlr.de

Esther Bosch
Institute of Transportation Systems
German Aerospace Center (DLR e.V.)
Braunschweig, Germany
Esther.Bosch@dlr.de

***Abstract*—With increasing technological progress in autonomous driving, careful design of passenger-autonomous vehicle interaction is necessary. The current study examined this during an on-demand drive with a real-world Wizard-of-Oz study. The vehicle collected each participant from a virtual stop, drove to a destination and returned to the starting location. On the way back, two unexpected events occurred – an abrupt braking event and an unexpected detour. During the events, 19 out of the 37 participants received an error notice followed by content information, while the remaining 18 only received an error notice. Psychophysiological data was gathered during and after the drive in the Wizard-of-Oz autonomous vehicle and served to determine a passenger’s information need during the drive. Analyses of the physiological data showed a cardiological reaction to the events, especially during the abrupt braking event. The subjective evaluations served as the primary source of insight into a passenger’s wish for information during events and the whole drive. The results indicate that information presentation positively impacted the passengers’ feeling of pleasure, safety and understandability of the vehicle.**

I. Introduction

Transportation research currently promotes a transition from traditional human-driven to autonomous driving vehicles. Reasons for this transition in transportation are that transferring driving responsibilities from a human driver to an automated vehicle may increase driving safety [1] and reduce the use of fossil fuels and emissions [2]. Passenger-autonomous vehicle interaction impacts a passenger’s expectations and attitude toward automated vehicles. A recent study [3] found that people have a low-level negative attitude towards technology but a high level of trust in automated vehicles. On the one hand, technology is seen as dehumanizing [3] and passengers fear a loss of control and of fun in driving [1]. On the other hand, people have expectations of reliability, functionality, and helpfulness of an autonomous vehicle [3]. However, individual and socio-demographic differences in people’s expectations in automated vehicles exist [4].

In general, passengers traveling in autonomous vehicles wish to interact with the vehicle, mainly in the form of receiving information [5], [6]. Therefore, a system that is as much self-explanatory as possible is seen as most satisfactory, trustworthy and achieves the highest situational awareness [5]. According to the concept of “locus of control,” the passengers feel that either they or the automated system are mainly responsible for the behavior of the vehicle [7].

Insufficient information may lead to an experienced passive role, and a feeling of failing to maintain a sense of control [8]. Information need is the desire of a person to receive information to satisfy the subjectively experienced lack of information. Ongoing research investigates what type of information to display and the timing of giving specific information to satisfy this need. Explaining actions can promote trust and acceptance in the automated vehicle [8]. Investigating psychophysiological reactions may help to understand passengers’ information needs. For instance, cardiological activity can be measured via electrocardiography (ECG). Cardiovascular parameters extracted from ECG, such as heart rate (HR) and heart rate variability (HRV), are indicators of psychological stress and have previously been used to assess a passenger’s state in automated vehicles [9]. Emotions and arousal are related to presented information. A study by Koo and colleagues [10], for example, showed with the help of subjective evaluations in questionnaires that information about the causes of events, given during a driving simulation with events, decreased anxiety. Only objective measurable physiological data and subjective evaluations of the participants together can give a holistic picture of people’s emotions and their need for information. Hence, the focus of the current study is to investigate a person’s need for information when traveling in an autonomously driving vehicle to improve future mobility concepts. Especially the interaction of the actions and explanations given by the vehicle and the passenger are examined. Physiological data (ECG) as well as subjective evaluations were analyzed. The current study was conducted as a real-world Wizard-of-Oz study in which an on-demand drive in an automated vehicle of automation level four was simulated.

II. Methods

A. Data Collection

1) Study Design: The real-world study was a Wizard-of-Oz study in which the participants were made to believe that they are traveling in an autonomously driving vehicle. During the travel, two events occurred: an abrupt braking event and an unexpected detour. During the events, the participants received different information depending on their assigned groups:

- withInfo group: During the events, the participants received an error message followed by content information whenever an event occurred.

- noInfo group: During the events, the participants received an error notice without content information whenever an event occurred.

This design enabled a between-group comparison of the participants' reactions to different information displays.

2) Participants: Forty-two subjects signed up for the study. Four participants were excluded from the analyses due to malfunctioning technical equipment. We excluded one further participant because he was under time pressure for private reasons during the whole experiment. Therefore, the physiological data was not comparable to that of the other participants. Hence, 37 participants remained for the analyses (11 female, 26 male). The mean (M) age over all participants was 33 years with a standard deviation (SD) of 11.0 and an age range of 36 years (min=21 years, max=57 years). The participant groups consisted of 19 participants for the withInfo group (ten female, M = 33 years, SD = 11.5, range = 35 years) and 18 participants for the noInfo group (one female, M = 34 years, SD = 10.7, range = 33 years). Participants were recruited via a participant database of the German Aerospace Center and via social media postings. For each commenced half hour, the participants received 5€ as financial reimbursement. All participants were informed about the process of the study and gave their informed consent to participate.

3) Technical Set-Up: The research vehicle we used for the study was a Volkswagen Passat. A GPS sensor recorded the location coordinates and the vehicle speed during the drive. In the vehicle, a tablet served as a display of information. We collected physiological data with the ECG sensor EcgMove 3 from movisens GmbH (Karlsruhe, Germany). We gathered subjective ratings during the drive via an application on a Google Pixel 4a smartphone. Participants could additionally use the application to book drives with the vehicle, fill in questionnaires, and navigate.

4) Experimental set-up:

a) Wizard-of-Oz study: The participants were told that they were traveling in an automated vehicle while the vehicle was actually driven manually. They were told that the person behind the steering wheel was only present for safety reasons. A partition between the front and back seats prevented the participants from realizing who was in charge of the vehicle. This way it was possible to examine the driving experience of participants in an automated vehicle. At the end of the study, the participants were asked whether they believed that the vehicle drove autonomously and all of them confirmed this. The participants were then informed that the vehicle did not drive autonomously.

b) Scenario: The experiment took place in a scenario that embedded the trip in an autonomously driving vehicle in a possible everyday situation. Each participant was instructed to imagine that they could not work for the upcoming three hours due to technical difficulties at their workplace. Therefore, they decide to return an old book to the nearest sidewalk library and borrow a new one. To get from the workplace to the sidewalk library and back, they had to book an autonomous shuttle car with the help of the smartphone application and were picked up from a virtual stop.

Fig. 1. Driving route of the vehicle. Light blue and orange lines indicate the forward run. The orange sections mark the baseline routes for the analyses of the physiological data. The dark blue line indicates the return path. Two events occurred, which are marked in red.

c) Route: The automated vehicle picked up the participant at a virtual stop, to which the participants navigated with a smartphone application (see Figure 1). Next, the vehicle drove the participant to a stop near the sidewalk library. The participant then navigated to the sidewalk library and booked a vehicle for the way back to the campus. The participant navigated to the next virtual stop and was collected by the vehicle, which drove back in the direction of the destination. On the way back, two events occurred: The first was that the vehicle halted abruptly in a parking lot for one minute. The second event occurred when the vehicle deviated from the route indicated on the smartphone application. The detour was about one kilometer long and took about two minutes. Finally, the vehicle arrived at the campus, the participants left the vehicle and navigated back to the study's starting point. The whole trip in the vehicle took about 30 min, and the drive was about 15 km long.

d) Information display: During the whole drive, a display showed the starting and end address of the trip and informed the participant that the automated driving mode was turned on. Depending on which participant group the passenger in the car was assigned to, different information was displayed on the screen when the events occurred. Participants assigned to the withInfo group received an error notice followed by content information on what happened and why the vehicle reacted to the event in a specific way. The information for the withInfo group for the first event stated that "[t]his is a stop to synchronize the GPS data. The drive will continue shortly." and for the second event that "[a] problem with a public-transit bus driving in front was communicated and the route was adapted.". Participants assigned to the noInfo group only received an error notice without content information on the event.

5) Measures:

a) Physiological recordings: To measure cardiological activity during the drive, the participants wore an ECG sensor on their skin as a chest belt.

b) Subjective rating: During the whole experiment, the participants were told to rate their well-being on a 10-point Likert scale from unpleasant (one) to pleasant (ten) with a slider at the bottom edge of the smartphone screen.

c) Post-questionnaire: The post-questionnaire contained Likert-scaled and open-ended questions concerning the participants' information needs. Quantitative statements were posed to the participants regarding the event of the abrupt braking, the event of the unexpected detour and the whole drive in the vehicle. For each of these three categories, three statements regarding the information that was given to the participants during the drive ("The display in the autonomous vehicle helped me to understand the present actions of the vehicle", "In that situation / during the drive I felt insecure due to missing information", "In that situation / during the drive I wished for more information") had to be rated on a 5-point Likert scale. A rating of one meant total disagreement and five meant total agreement.

"The following information would have helped me to understand the situation better" was the open-ended question that was posed to the participants for the two events. To evaluate the whole drive in the vehicle, the participants were additionally asked to describe up to three situations in which they would have wished for more information during the drive in the vehicle.

d) Study procedure: For the study, the participants came to the German Aerospace Center. First, documents regarding data privacy, information about the study, and a consent form were signed by the participant. Then the participant was equipped with a mobile ECG and a smartphone. A demographic questionnaire was given to the participant and filled in. Then the scenario of having to return a book to the sidewalk library by using an autonomously driving vehicle to get there was introduced. Due to the complexity of the study, a summary with the most important information was given to the participants to take with them on the trip. After the experiment, the participants were debriefed and informed that the vehicle did not drive autonomously. The whole experiment took about two hours. The procedure of this study follows the Helsinki Declaration and was reviewed and approved by the ethics committee of the German Aerospace Center (reference no. 5/21). The participants gave their informed consent to partake in the study, and a hygiene concept according to the present Covid-19 rules was designed and implemented.

III. Data Processing

1) Data preparation:

a) Cardiological data: The HR and the HRV are the cardiological parameters of the ECG recordings, which are considered in the following. The ECG data of only 23 out of 37 participants could be examined because, in 13 cases, the data was not recorded due to a nonfunctional device. Of the ECG data, we calculated heart rate (HR) and heart rate variability (HRV) measures. Per participant, two baseline route segments were selected to perform a z-transformation of the HR values. The median (MD) and the median absolute deviation (MAD) were chosen as robust measures for the z-transformation. Baek and colleagues [11] showed in their study that the Root Mean Square Successive Difference (RMSSD) metric for HRV is reliable for 30s time windows. The normalized RMSSD (nRMSSD) values were calculated for each participant by dividing the RMSSD values by the mean RR-intervals to take account of individual differences in HRV [12].

b) Subjective rating: The 10-point Likert-scaled questions did not need any preparation before the quantitative analyses. Seven participants had to be completely excluded from the analyses of the subjective ratings. In six cases, the recording of the subjective ratings stopped at the time points of the events, and, in one case, the rating remained at one level during the whole recording. Three further participants were only considered during the first event and excluded for the analysis of the second event due to missing recordings. Therefore, 30 participants were examined for the first event, and 27 participants were examined for the second experiment.

c) Post-questionnaire: All 37 participants completed the post-questionnaire. The statements to the open-ended questions were categorized into a two-level category system. The statements were categorized on a first level into the categories "Display of route," "Display of time," and "Display of further information", and on a second level into the categories "on the tablet," "Duration of event", "Duration of the rest of the trip until arrival," "More context" and "Need for action or not." Residual categories contain statements that could not be fully assigned to the level-1 or level-2 categories.

d) Time intervals for analysis: For each participant and each event, time intervals are defined to perform analyses of the physiological data and subjective ratings during the events and during the information display. The mean values of the physiological parameters and the subjective ratings of every time interval were then calculated to perform comparative analyses in the next step. The time intervals before the event, during the event and after the event were examined for each of the two events on trip to identify the impact of the events on the participants. GPS coordinates and the timing of the information display were used to define the time intervals. Each time interval had the same time length.

2) Data Analyses and Statistics: Linear mixed effects analyses were performed to systematically analyze the physiological reactions and subjective ratings of the participants during the two events. R and the package "nlme" [13] with the function lme were used for the analysis. The results are going to be reported separately for the event of abrupt braking and the event of the unexpected detour. The time interval (levels: before event, during event, after event) and the group (levels: withInfo: passenger received error notice and content information, noInfo: passenger received error notice but no content information) were entered as fixed effects. The time interval was seen as a categorical fixed effect to compare within-subject data. The participants (P) were examined as random effects with random slopes for the effect time interval. The reason to take account of the time interval as random slopes is the assumption that the responses differ between participants for the time intervals. Some might be more affected by the events than others. Since an effect of the subdivision of the participants into groups is only expected in the time interval during the event and not in the other time intervals, the assumption that the responses of the participants differ between groups is neglected at this point. For this reason and because of too little data, the effect group will not be modelled as random slopes. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity. Since linear models are robust to the violation of the normality assumption when it comes to hypothesis testing even for relatively small sample sizes [14], [15], the normality assumption will not be further considered in the following analyses. P-values were obtained by likelihood ratio tests with an analysis of variance (ANOVA) of the full model with the effect in question 1 against the model without the effect in question.

$$value \sim timeInterval + group + timeInterval|P \quad (1)$$

Frequency analysis: A frequency analysis was performed on the categorized statements of the post-questionnaire to examine what information was retrospectively most desired by the participants during the events and the whole drive. Instead of counting the number of specific statements, the number of participants that gave specific statements was analyzed. This way, a frequency bias due to some participants answering the questions more extensively than others is prevented.

A. Results

The linear mixed effect analysis revealed a significant difference in the participants' z-transformed heart rates (zHRs) between the time intervals during the abrupt braking event ($\chi^2(7) = 16.3117, p < .05$). A Tukey post-hoc test showed that the zHR was significantly lower during the event ($p < .001$) and after the event ($p < .05$) than before the event (Figure 2). No significant difference in zHR could be detected between the two groups. The nRMSSD values neither differed significantly between the time intervals nor between the groups. The results of Spearman's rank correlation show a neglectable correlation between the driving speed and the heart rate during a normal driving situation as in the baseline interval ($\rho = 0.069$). A weak positive correlation between speed and heart rate can be observed during the first event ($\rho = 0.26$). Hence, slightly lower heart rates were found at lower speeds. The unexpected detour showed a minimal correlation between speed and heart rate ($\rho = 0.099$).

A linear mixed effect analysis was performed for the subjective ratings of the participants during the drive. The time intervals differed significantly ($\chi^2(7) = 54.9866, p < .0001$) during the abrupt braking event. The post-hoc Tukey test revealed significantly lower ratings of the participants during the first event compared to before ($p < .0001$) and after ($p < .01$). After the event, the ratings were significantly lower compared to before the event ($p < .01$). No significant difference could be observed between the withInfo and noInfo groups. During the unexpected detour, the time intervals also differed significantly ($\chi^2(7) = 50.7342, p < .0001$), but the more conservative post-hoc Tukey test did not show a significant difference between specific time intervals. A significant difference for the fixed effect group was found during the mixed linear effects analysis ($\chi^2(1) = 0.5432, p < .05$), but no significant difference was revealed by the Tukey test. Figure 3 shows that the withInfo group has overall higher ratings than the noInfo group in the case of the unexpected detour. A stronger reaction in terms of ratings can be observed during the abrupt braking event compared to the unexpected detour (Figure 3). The mean subjective ratings of the withInfo group are highest before the unexpected detour (mean=9.23) compared to before and after the abrupt braking event (mean before=8.90, mean after=8.67). The highest mean subjective ratings of the noInfo group were before the abrupt braking event (mean=9.21), while the mean ratings after the abrupt braking event (mean=8.58) and before the unexpected detour (mean=8.81) remained comparable (Figure 3).

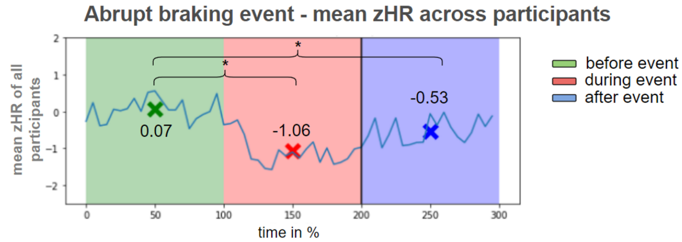

Fig. 2. The mean zHR across participants (n=23) for the first event is depicted (blue line). Average zHRs across participants were calculated for time in percent. The mean z-scores over all participants in the respective time interval is marked with a cross. The asterisk denotes a significant difference between the time intervals (p¡.05).

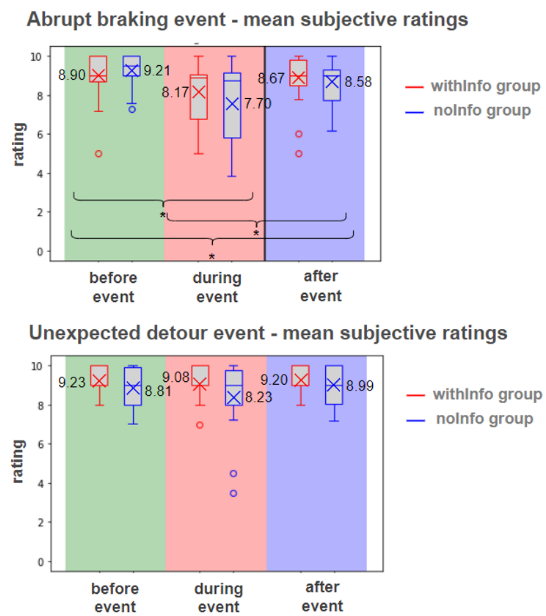

Fig. 3. Mean ratings during the abrupt braking event and the unexpected detour. The red boxplots contain the mean ratings of the withInfo group (nwithInfo=16) and the blue boxplots contain the mean ratings of the noInfo group (nnoInfo=14). The asterisks indicate significant differences between the time intervals.

Significant differences between the answers of the participant groups could be observed in the post-questionnaire. The question whether the display helped the participant to understand the actions of the vehicle was affirmed significantly more strongly by the withInfo group than by the noInfo group in all three categories as tested by Mann-Whitney U test (MW)

$(U^{Event1}_{Mann-Witney} = 279.0, p < .001; U^{Event2}_{MW} = 304.5, p < .0001; U^{WholeDrive}_{MW} = 277.5, p < .001)$ (see Figure 4). Results further showed that the withInfo group felt significantly less insecure than the noInfo group during the first event($U_{MW} = 103.0, p < .05$), the second event ($U_{MW} = 80.0, p < .05$) and the whole drive ($U_{MW} = 97.0, p < .05$) (see Figure 3). The withInfo group also significantly less affirmed the question whether more information was wished for during the first ($U_{MW} = 113.0, p < .05$), the second event ($U_{MW} = 92.0, p < .05$) and the whole drive ($U_{MW} = 107.0, p < .05$) than the noInfo group (see Figure 4). Content information was therefore in general preferred over merely an error message. This is the case for both event types equally (see Figure 4).

The frequency analysis of the open-ended questions of the post-questionnaire showed that during the first event, the second event and the whole drive, participants in the noInfo group wished for more context information than participants in the withInfo group. Although participants in the withInfo group received content information, participants declared that more context information would have helped them understand the event situations better and/or wished for more context information during the drive. More information on the exact reason for e.g. the stop during the first event that exceeded the presented content information was asked for. A participant proposed to add a button "with 'further information' for those people, who would like to have more background information" (P6). During the first event, both groups would have liked to receive information on the need of them to take action or not.The results indicate that the error messages did not suffice for the participants to understand the event situations and that more information was whished for by participants of the noInfo group and by some participants of the withInfo group. Those participants which received content information during the event tend to wish more for information concerning the whole drive, such as the current route and time displays than participants that did not receive content information. The focus of interest of the noInfo group was to gain more information on the cause and actions of the vehicle during the events.

IV. Discussion

This paper presents a Wizard-of-Oz study in which participants experienced the feeling of driving in an automated car. We assessed physiological as well as subjective data continuously during the study, with additional subjective evaluations after the study. During the drive, two events were induced, and differences in measurements were compared between a group that received information and a group that received no information.

We found that the events had an impact on the passenger's physiology and subjective ratings. The HR changed during both events, and more negative subjective ratings could be observed, especially during the abrupt braking event. The HRV, however, was not influenced by the events. Contrary to expectations, the zHRs of the participants decreased significantly from before to during the abrupt braking event. A correlation between driving speed and heart rate, which would explain the reduced HRs when the vehicle stopped for the abrupt braking, could be dismissed as an explanation. Therefore, we propose that the abrupt braking event as such had an influence on the participants' HRs. Beggiato et al. [16] also

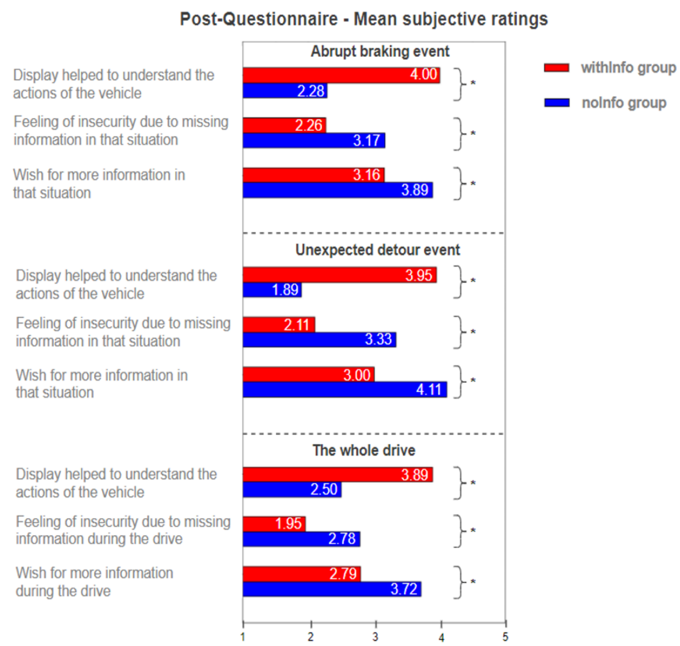

Fig. 4. Mean participant ratings of the post-questionnaire. The asterisks indicate significant differences between the groups as a result of a one-sided calculated Mann-Whitney U-test.

witnessed a decrease in heart rates when participants were in discomfort periods during a simulation of automated driving. The authors explained their results with the "preparation for action" hypothesis [16]. The hypothesis refers to a decrease in HR before a planned action [17]. A significant decrease in zHR from before to during the event but, however, no significant increase in zHR between the time intervals during and after the event [16] can be observed in our study. An "action"-reaction after the event was therefore not present. An emotional fear reaction to the events, which is expressed physiologically by a decrease in HR and no change in HRV, may be a better explanation [18]. A literature review [18] reported that fear paradigms that elicit a high degree of self-involvement lead to the feeling of higher imminence of threat. The participants in the present study retrospectively expressed a wish for information on whether they had to act in the case of the abrupt braking event. This wish indicates that the participants felt somewhat involved in that situation. According to [18], this involvement could have led to an imminent fear response which can be characterized by immobilization rather than by an active coping response. The physiological responses to the abrupt braking event could therefore be an indication of the participants' experience of fear.

The participants' subjective ratings during the drive indicate a negative reaction to the abrupt braking event. This seems to have had a prolonged negative effect on the experienced pleasantness of the participants in the noInfo group during the remaining drive but not for the withInfo group. While the withInfo group regained their pleasure before the unexpected detour began, the noInfo participants did not fully regain their pleasure. This difference cannot be attributed to a bias in the participant constellation in the groups since the subjective ratings were comparable between the two groups during the baseline conditions. We propose that this difference can be attributed to the type of information that the participants received.

The retrospective evaluations of the participants revealed a preference for more information, especially context information. The quantitative analyses of the post-questionnaire showed that information about the event and its causes was preferred over only an error message to help understand the actions of the automated vehicle and to feel less insecure during an unexpected event. In sum, the subjective ratings during the drive combined with the retrospective evaluations of the participants indicate a preference and wish for information during a drive in an automated vehicle. The drive becomes more pleasurable, the feeling of insecurity due to missing information is lower and occurring events are perceived as more understandable when more context information is given.

The current study was conducted as a real-world Wizard-of-Oz study in which an on-demand drive in an automated vehicle of automation level four was simulated. The experimental set-up enabled us to examine the impact of autonomous driving on passengers in a real-world setting. In the past, most studies on automated driving were conducted in driving simulators, but people may behave differently in driving simulators than in real-world driving conditions [8], [19]. The real-world setting of this study decreases the bias of driving simulation studies in the research area and enables a comparison of the results of driving simulation and real-world studies. Future research will need to investigate whether optimizing the content and timing of the information display will lead to a decreased physiological reaction to the event situation. The present results, together with the results of prior research, indicate a preference for information during events and during the whole drive. The exact timing of the information display before an event and the interaction between the time and the content of the display still needs to be investigated in future research [8]. Using live explanations when events occur should also be considered in the future [20].

V. Conclusion

A person's need for information while traveling in an automated vehicle is best satisfied by receiving content information over merely an error notice. Subjective evaluations of the participants emphasized the benefit of receiving information on the events during the drive. As a result, the drive becomes more pleasurable, the feeling of insecurity due to missing information is lower and occurring events are perceived as more understandable when context information is given. The passenger-automated vehicle interaction is therefore improved.

Acknowledgment

This work was developed within the framework of the project ViVre, which was funded by the German Federal Ministry for Digital and Transport (funding number: 01MM19014A).

References

[1] D. Howard and D. Dai, "Public perceptions of self-driving cars: The case of berkeley, california," in *Transportation research board 93rd annual meeting*, vol. 14, no. 4502, 2014, pp. 1–16.

[2] H. Hwang and C.-K. Song, "Changes in air pollutant emissions from road vehicles due to autonomous driving technology: A conceptual modeling approach," 2020.

[3] I. P. Tussyadiah, F. J. Zach, and J. Wang, "Attitudes toward autonomous on demand mobility system: The case of self-driving taxi," in *Information and communication technologies in tourism 2017*. Springer, 2017, pp. 755–766.

[4] Q. Zhang, X. J. Yang, and L. P. Robert Jr, "Individual differences and expectations of automated vehicles," *International Journal of Human–Computer Interaction*, vol. 38, no. 9, pp. 825–836, 2022.

[5] J. Maarten Schraagen, S. Kerwien Lopez, C. Schneider, V. Schneider, S. Tönjes, and E. Wiechmann, "The role of transparency and explainability in automated systems," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 65, no. 1. SAGE Publications Sage CA: Los Angeles, CA, 2021, pp. 27–31.

[6] J. Meurer, C. Pakusch, G. Stevens, D. Randall, and V. Wulf, "A wizard of oz study on passengers' experiences of a robo-taxi service in real-life settings," in *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 2020, pp. 1365–1377.

[7] N. A. Stanton and M. S. Young, "Vehicle automation and driving performance," *Ergonomics*, vol. 41, no. 7, pp. 1014–1028, 1998.

[8] Q. Zhang, X. J. Yang, and L. P. Robert, "What and when to explain? a survey of the impact of explanation on attitudes toward adopting automated vehicles," *IEEE Access*, vol. 9, pp. 159 533–159 540, 2021.

[9] P. Zontone, A. Affanni, R. Bernardini, D. Brisinda, L. Del Linz, F. Formaggia, D. Minen, M. Minen, C. Savorgnan, A. Piras, R. Rinaldo, and R. Fenici, "Comparative assessment of drivers' stress induced by autonomous and manual driving with heart rate variability parameters and machine learning analysis of electrodermal activity," *European Heart Journal*, vol. 41, 2020.

[10] J. Koo, J. Kwac, W. Ju, M. Steinert, L. Leifer, and C. Nass, "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)*, vol. 9, no. 4, pp. 269–275, 2015.

[11] H. J. Baek, C.-H. Cho, J. Cho, and J.-M. Woo, "Reliability of ultra-short-term analysis as a surrogate of standard 5-min analysis of heart rate variability," *Telemedicine and e-Health*, vol. 21, no. 5, pp. 404–414, 2015.

[12] C. C. Grant, C. Murray, D. C. Janse van Rensburg, and L. Fletcher, "A comparison between heart rate and heart rate variability as indicators of cardiac health and fitness," *Frontiers in physiology*, vol. 4, p. 337, 2013.

[13] J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, S. Heisterkamp, B. Van Willigen, and R. Maintainer, "Package 'nlme'," *Linear and nonlinear mixed effects models, version*, vol. 3, no. 1, 2017.

[14] A. Gelman and J. Hill, *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press, 2006.

[15] U. Knief and W. Forstmeier, "Violating the normality assumption may be the lesser of two evils," *Behavior Research Methods*, vol. 53, no. 6, pp. 2576–2590, 2021.

[16] M. Beggiato, F. Hartwich, and J. Krems, "Using smartbands, pupillometry and body motion to detect discomfort in automated driving," *Frontiers in human neuroscience*, vol. 12, p. 338, 2018.

[17] A. Cooke, M. Kavussanu, G. Gallicchio, A. Willoughby, D. McIntyre, and C. Ring, "Preparation for action: Psychophysiological activity preceding a motor skill as a function of expertise, performance outcome, and psychological pressure," *Psychophysiology*, vol. 51, no. 4, pp. 374–384, 2014.

[18] S. D. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biological psychology*, vol. 84, no. 3, pp. 394–421, 2010.

[19] A.-M. Brouwer, E. Van Dam, J. B. Van Erp, D. P. Spangler, and J. R. Brooks, "Improving real-life estimates of emotion based on heart rate: a perspective on taking metabolic heart rate into account," *Frontiers in human neuroscience*, vol. 12, p. 284, 2018.

[20] T. Schneider, J. Hois, A. Rosenstein, S. Ghellal, D. Theofanou-Fülbier, and A. R. Gerlicher, "Explain yourself! transparency for positive ux in autonomous driving," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–12.