Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '16), October 24–26, 2016, Ann Arbor, MI, USA

Automated Driving System, Male, or Female Driver: Who'd You Prefer? Comparative Analysis of Passengers' Mental Conditions, Emotional States & Qualitative Feedback

Philipp Wintersberger
CARISSMA
Ingolstadt, Germany
wintersberger@carissma.eu

Andreas Riener
University of Applied Sciences
Ingolstadt, Germany
andreas.riener@thi.de

Anna-Katharina Frison
University of Applied Sciences
Ingolstadt, Germany
anna-katharina.frison@thi.de

ABSTRACT
It is expected that automated vehicles (AVs) will only be used when customers believe them to be safe, trustworthy, and match their personal driving style. As AVs are not very common today, most previous studies on trust, user experience, or acceptance measures in automated driving are based on qualitative measures. The approach followed in this work is different, as we compared the direct effect of human drivers versus automated driving systems (ADSs) on the front seat passenger. In a driving simulator study (N=48), subjects had either to ride with an ADS, a male, or a female driver. Driving scenarios were the same for all subjects. Findings from quantitative measurements (HRV, face tracking) and qualitative pre-/post study surveys and interviews suggest that there are no significant differences between the passenger groups. Our conclusion is, that passengers are already inclined to accept ADS and that the market is ready for AVs.

Author Keywords
Automated Vehicles; User Acceptance; ISO 9241/29119; SAE J3016; PANAS; Affect Grid; Emotion; Facial Expression; ECG/HRV; Trust in Automation.

ACM Classification Keywords
Human-centered computing: Human computer interaction (HCI): HCI design and evaluation methods, User studies

INTRODUCTION
When talking about the ongoing transformation of cars from manually controlled to AVs (NHTSA levels 3 or higher), researchers and manufacturers often highlight only the benefits, such as increased road safety, more leisure time, improved flow of traffic, reduced fuel consumption or mobility for the impaired [3]. Currently it is not clear if people accept this technology *"just because it is there"* and how to deal with a potential behavior change in humans emerging from automation. This uncertainty is backed by questions such as *"perhaps the first question we should ask is, will drivers accept automated vehicles?"* [10]. According to a 2016 survey conducted in Austria, only 19% of respondents agreed to the statement that "*automated driving comes with higher safety*" and another 25% have the fear that automated driving would result in a higher number of accidents due to software failures [7]. Another survey [16] confirms the results. More than 80% of study participants "*wanted to gain control of the car in any situation*" and further would not allow an automated function to overrule their input. Only 22% of the interviewees responded that they "*can imagine to own a vehicle without lateral/longitudinal controls, i. e., wheel, pedals*" as introduced in the Google self driving vehicle (May 2014).

Figure 1. Between groups design: Test subjects driven by an ADS, male or female driver.

As *"new technologies usually require several years to build market acceptance"* [12], the benefit of automated driving systems for the individual must be communicated rather sooner than later. High user acceptance and adoption to personal (driving) preferences will be important features in this regard. Thill et al. [23] argue that nowadays *"people start to perceive vehicles as intelligent agents rather than mere tools"* – a shift in perception that shows many analogies to the "uncanny valley"-problem [14] discussed within human-robot interaction for quite a long time. Garcia and colleagues pointed out that researchers have not yet fully *"examined attitudes towards various autonomous features."* This is an important aspect as *"manufacturers and government agencies must understand if consumers are receptive of the ongoing advancements,"* because *"the individual trust will take a prominent role in the success of the future use of vehicles"* [6]. As most previous research addressing user acceptance of AVs is presented in the form of interviews or opinion surveys, we wanted to go a step further by looking not only on qualitative, but also on quantitative measures in a realistic setting. Therefore, we've set up a user study analyzing the effect of HAD

ACM 978-1-4503-4533-0/16/10…$15.00
DOI: http://dx.doi.org/10.1145/3003715.3005410

on the front-seat passenger in comparison to male/female human drivers (Figure 1). We argue that a difference in the user acceptance (sometimes also qualified as trust) between the three groups of drivers, would be – if verifiable – reflected in varying mental/emotional states, changed behavior or different qualitative feedback.

RELATED WORK

Various work addressing acceptance of automated vehicles was published recently. Bansal, Kockelman and Singh [1] performed an internet-based survey answered by 347 participants from the Austin area (Texas). They investigated how much people are willing to pay for specific autonomous features and concluded that people would explicitly pay more for higher levels of autonomy (NHTSA level 4 vs. level 3) and that *"high-income, tech-savy males living in urban areas"* have highest interest for those technologies, while older people seemed to be less interested, maybe due to the learning effort and a feared loss of driving pleasure. Kyriakidis et al. [11] analyzed the results of a similar survey containing over 5,000 responses from 109 countries and state that, even though many people liked the idea of automated driving, manual driving was the *"most enjoyable mode of driving"*. They also point out that the concept of automated driving polarized into two groups where one liked the idea while the other group seems to strongly refuse it [11].

Public opinion of 1,533 persons on AVs from the US, UK and Australia was examined by Schoettle and Sivak [22], showing that many respondents were concerned that *"automated vehicles may not perform as well as actual drivers"* and did not like the idea of vehicles without driver controls. Payre et al. [17] asked 421 French drivers about their attitudes to automated driving and gained a positive acceptance rate (68,1%), where the preferred situations were highway driving and automated parking.

Rödel et al. [20] tried to explicitly get insight into how people accept the individual automation levels as provided by NHTSA and concluded, that higher levels of automation seem to be less accepted, except by participants that already have experience with ADAS (advanced driver assistance system). They further point out, that *"it will be a goal for the automotive industry to implement highly autonomous cars adapted to the drivers' needs in terms of a pleasurable and authentic driving experience"*.

When talking about acceptance of AVs, trust is an additional important factor. People will accept such vehicles only if they also trust them to drive safely and responsibly. To gain a user's trust, Ekman and colleagues [5] developed a formal framework for HMI design in vehicles called the "Life Cycle of Trust" and state that a holistic approach is necessary as trust development "starts long before a user's first contact with the system, and continues long thereafter" [5]. Hergeth et al. [9] showed that there exists a connection between trust and monitoring frequency (by evaluating participants' gaze behavior). As trust in automated vehicles is often examined in the context of Take-Over-Requests (TOR), Hergeth and colleagues looked at the effects of TOR on trust with a user study containing participants from different nationalities (German and Chinese drivers). Their results suggest that initial differences between the nationalities regarding trust vanished after both groups used the system. Payre et al. [18] investigated the effect of trust on manual control recovery and showed that higher levels of trust (overtrust) can lead to higher reaction times. In a similar study, Helldin and colleagues [8] presented confidence of the vehicle's abilities to study participants and could demonstrate that this leads to a better Take-over performance, even the overall trust in the system was less than in the control group.

As in our user study, participants are not directly interacting with the vehicle and merely act as front-seat passengers, their attributed trust in AVs can only be influenced by its driving behavior. To provoke a loss of trust that might result in different passenger behavior based on the type of driver (human or ADS), we implemented some dangerous driving situations.

RESEARCH QUESTIONS

"User acceptance" is an abstract term that has been redefined multiple times based on need and purpose. Former approaches like the Technology Acceptance Model or the Unified Theory of Acceptance and Use of Technology (UTAUT) [24] mainly target classical systems and associate user acceptance with terms like "ease of use" or "performance expectancy". Some could argue that an AV would be more seen as a robot than a classical computer system. Problems similar to those addressed for the "uncanny valley"-problem [14] need to be discussed, which requires extended criteria of user acceptance [2]. With automated driving, users proverbially lay their own life into the "hands" of a complex computer system. Higher acceptance levels will thus have to consider "trust in technology" as well. For this work, a proper definition of user acceptance in the context of HAD is not the main target. Instead, our focus is to examine how people perceive AVs in contrast to manually controlled cars and if they are willing to accept an ADS as driver at all. Following this research questions, we identify two hypotheses to be deeply discussed in this paper.

RQ: Do people accept an ADS the same way as they accept human drivers?

Hi: There is no difference in the mental condition of front-seat passengers of an ADS, male or female driver. Mental conditions will be quantitatively analyzed from stress indicators extracted from ECG/HRV measurements.

Hii: There is no difference in the emotional state of front-seat passengers of an ADS, male or female driver. Emotional states will be analyzed quantitatively (classification of facial expressions from videos) as well as qualitatively (questionnaires and interviews).

STUDY SETUP

In the experiment, participants were instructed to take place in the front passenger seat and act as passive passenger in a driving simulator experiment. To run the experiment, the simulation environment IPG CarMaker was used in combination with a Hexapod simulator (moving platform). One run through the course took 9 minutes (with a preceding one minute phase for physiological baseline estimation) and the track was split into four parts: It started with an urban section (including dense traffic and multiple unregulated crossings) followed by a section in the mountains with narrow hairpin curves. Part three contained longer straight passages in a forest while the final section was again a mountain road (this time without other traffic). The whole driving process was automated by predefined maneuvers and contained multiple overtaking situations that would commonly be described as risky, e. g., overtaking in the city with oncoming truck (Figure 2) or overtaking in a curve. For the two groups with human drivers (front-seat passengers of male or female drivers), we implemented the experiment using a "Wizard of Oz" approach. The experimenter made the participants believe that the driver performs steering maneuvers on his/her own. The third group (front-seat passengers of the ADS) was told that they have the unique opportunity to test a novel automated driving system. All participants filled out questionnaires before and after driving. The questions included demographic information, their opinion on automated driving and psychological tests (PANAS, Affect Grid, Figure 3).

Figure 2. Overtaking maneuver with an approaching truck.

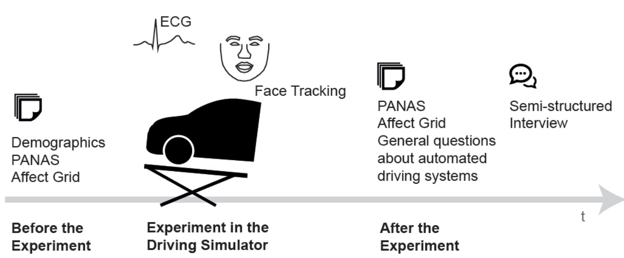

Figure 3. Timeline of our study setting containing all applied data collection methods.

Only subjects of the third group (driving with ADS) were individually interviewed about their opinions at the very end (just before debriefing). While in the car, subjects were continuously monitored with an ECG device (BioRadio Physiological Monitor) and video recordings (face tracking).

RESULTS

48 people participated voluntarily in the experiment (27 female, age $\overline{x_f} = 22.5 \pm 2.65$ and 21 male, age $\overline{x_m} = 24.04 \pm 3.38$). In this chapter, we discuss the results of the individual measurements (HRV measurements from ECG recordings, face analysis from videos, results of the qualitative questionnaires before and after driving) in detail.

ECG Stress Indicators

For evaluating emerging stress we measured the participants' heart rates using the BioRadio Physiological Monitoring Kit. From the raw data we extracted the temporal positions of the corresponding R-peaks using VivoSense software. Data sets were then used to calculate the heart rate variability (HRV), a measure that can estimate a person's level of stress [4]. A low HRV is an indication for stress or high mental load (but also illness), while a higher HRV can be associated with relaxation. As our baseline was measured only for one minute (participant already sitting in the simulator), followed by the 9 minutes of driving, we used the root mean square of successive differences (RMSSD) for our calculations. The RMSSD as a temporal parameter is suitable for short-time measurements [15], and higher values usually indicate relaxation. To quantify a potential increase/decrease of stress or relaxation over multiple measurements, we calculated RMSSD in consecutive time windows of one minute (resulting in 10 values for each participants, 1 minute baseline followed by 9 minutes driving). Afterwards we performed linear regression (RMSSD vs. time) and used the slope of the linear regression line as measure for increasing ($k<0$) or decreasing ($k>0$) stress tendency (Figures 4 and 5).

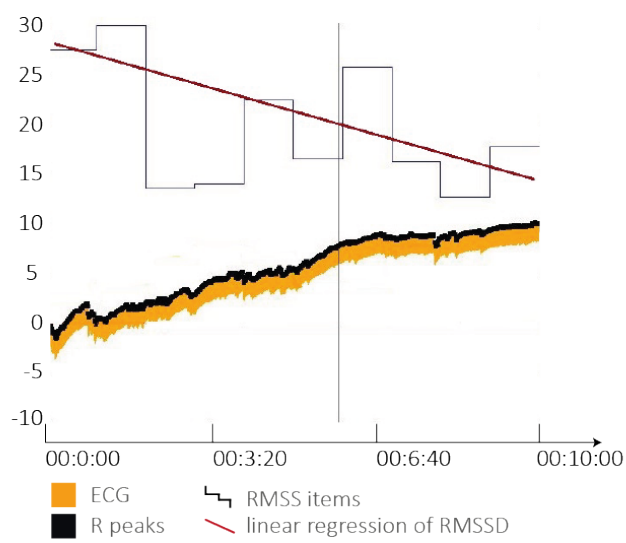

Figure 4. Sample image of a participant's RMSSD values and the corresponding regression line (red), indicating a decreased HRV during the stress situations.

The calculated slopes for each group were statistically analyzed using SPSS. Both ANOVA ($F(2,42) = 2.961$, $p = 0.063$) and Kruskal-Wallis (preceding test for normal distribution failed, $H(2) = 3.033$, $p = 0.219$) showed no statistically significant differences between the groups, neither between male/female nor between humans and ADSs.

Facial Expression Analysis from Video

The Microsoft Project Oxford Emotion API was used for face expression analysis of all subjects. The emotion API clas-

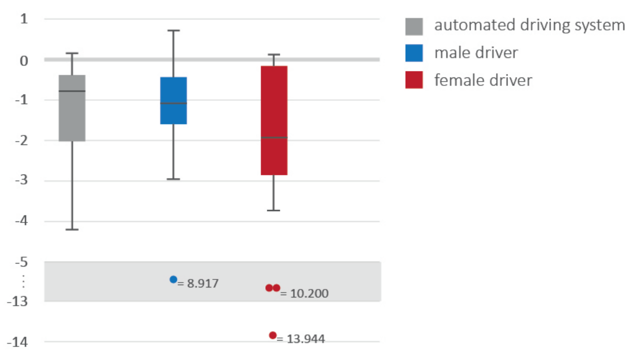

Figure 5. Boxplot of the RMSSD trend slopes for the individual groups.

sifies each frame as either *neutral* face or one of the 7 basic emotions *happiness, surprise, sadness, fear, disgust, contempt* and *anger*. In order to reduce data processing load, we extracted a single frame for every second of recorded video and used a C# program to classify the image according to the specification in the API documentation. From the total number of 31,859 images (all 3 groups), 31,180 could successfully be classified (97.8%). The results showed *neutral* face expressions for 94.3% and *happiness* for 4.4% of all images. The remaining 1.3% were attributed to *sadness* (0.9%), *surprise* (0.2%), *contempt* (<0.1%) and *anger* (<0.1%) – *disgust* and *fear* were classified not at all.
Even if only 5.7% of the classified face expressions are not neutral, looking at the individual groups reveals some interesting differences. Front-seat passengers of the ADS showed *happiness* in only 1.86% of the images, while this value reached 3.99% for passengers of a female driver and almost doubled for passengers of the male driver (7.60%), Figure 6. Performing ANOVA (normal distribution accepted by Kolmogorov-Smirnov for all groups, homogenity of variances accepted using Levene statistics) showed a significant difference between the male driver and the ADS situation ($F(2,45) = 6.035$, $p < 0.05$, significance of $p = 0.004$ using Tukey HSD Post-Hoc test).
Sadness was mostly present in the female-driver situation (1.8%), for male drivers and ADSs this value was almost similar and low (0.50%, 0.54%). Another emotion that reached a value worth to be mentioned, *surprise*, did not significantly differ between the three groups (ADS: 0.24%, male driver: 0.16%, female driver: 0.32%).
To gain additional insight about the distribution of the facial expressions, we further divided the dataset into 6 groups by splitting male and female front-seat passengers of male drivers, female drivers and ADSs. Although the number of participants is very little for some groups (only 4 male passengers of the female driver), this distinction indicates further interesting differences (Figure 6): Female faces showing *happiness* reached 9.27% as passengers of a male driver, while the direct opposite situation (male front-seat passengers of a female driver) reached only 1.56%. Contrarily, driver-passenger pairs with the same gender were (regarding *happiness*) nearly the same: 5.04% for female and 5.35% for male pairs.

Qualitative Survey Results

To analyze participants emotional states we conducted several qualitative examinations, PANAS (positive and negative affect schedule), Affect Grid, a qualitative survey for assessing the concept of automated driving, as well as a unstructured interview with the drivers of the ADS (group 3).

PANAS

PANAS is a standardized survey where participants have to rate their currently felt intensity concerning ten positive and ten negative feelings aka adjectives using a five-point Likert scale [25], where 1 means *"Very Slightly or Not at All"* and 5 means *"Extremely"*. To see the extent of changes within a group, subjects were asked to answer PANAS twice, before and after the experiment. We compared the answers after driving, as well as the level of change between the groups using SPSS. All effects are reported at $p < 0.05$. Regarding the answers of the PANAS which were conducted after driving, positive (PA) and negative affect (NA) have weak values (PA = 2.88, NA = 1.52), male (PA = 2.81, NA = 1.29) or ADS (PA = 2.95, NA = 1.34). Moreover these values do not differ between the groups for PA, $F(2, 27) = 0.411$, $p = 0.667$. Also the NA of the subjects is not significantly affected by the type of driver, $F(2, 27) = 1.666$, $p = 0.208$.
By visually inspecting Figure 7, showing the means of the ratings of each group and each adjective, we selected the adjectives *enthusiastic*, *irritable* and *jittery* for a closer look. All three analyzed adjectives showed no statistically significant difference in a Kruskal-Wallis test (*enthusiastic*: $H(2) = 1{,}727$, $p = 0.422$, *irritable*: $H(2) = 3.190$, $p = 0.203$, *jittery*: $H(2) = 3.679$, $p = 0.159$).

To examine the difference in the level of change between the groups, we subtracted the ratings of PANAS evaluated before and after the experiment. After a closer look (Figure 7), we selected the adjectives *excited*, *enthusiastic, inspired, attentive, irritable* and *jittery* as they differ more than others, Kruskal-Wallis resulted in no statistically significant differences for all adjectives (*excited*: $H(2) = 0.651$, $p = 0.722$, *enthusiastic*: $H(2) = 1.718$, $p = 0.424$, *inspired*: $H(2) = 1.295$, $p = 0.660$, *attentive*: $H(2) = 0.830$, $p = 0.660$, *irritable*: $H(2) = 1.486$, $p = 0.476$, *jittery*: $H(2) = 5.887$, $p = 0.053$).

Affect Grid

To get further insights, we applied a second standardized method to evaluate participants' emotional states before and after the experiment. In the Affect Grid, developed by Russel et al. [21], a subject must rate his/her current emotional state by marking a point in a two dimensional grid, where the x-axis represents *pleasure* from unpleasant to pleasant and the y-axis *arousal* from sleepiness to high arousal (Figure 8, left subimage).

To analyze the data, we again used a Kruskal-Wallis test. Regarding differences between the groups after driving we could not discover a significant affect on **pleasure** ($H(2) = 0.961$, $p = 0.619$). Also the dimension *arousal* is not significantly affected by a certain type of driver, $H(2) = 3.024$, $p = 0.220$ (Table 1).

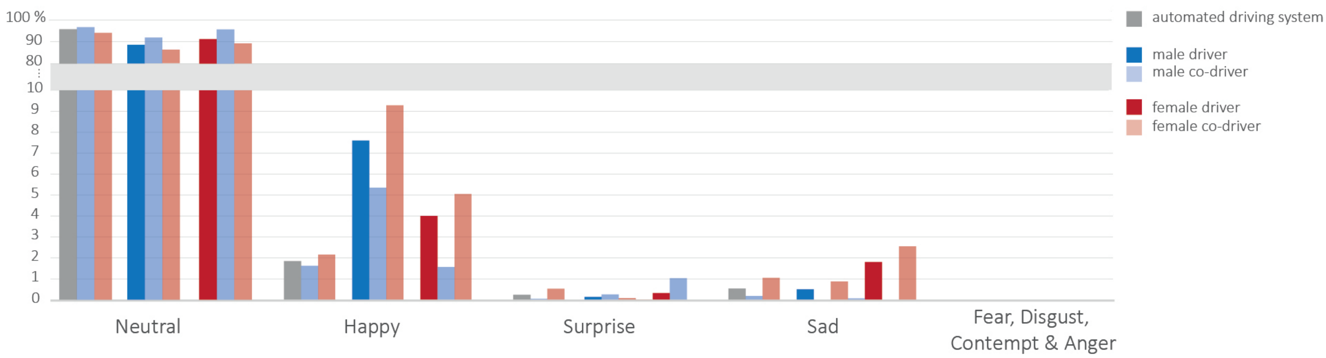

Figure 6. Classified emotional face expressions for the different groups: satisfaction (*happy*) for front-seat passengers of ADSs was generally low.

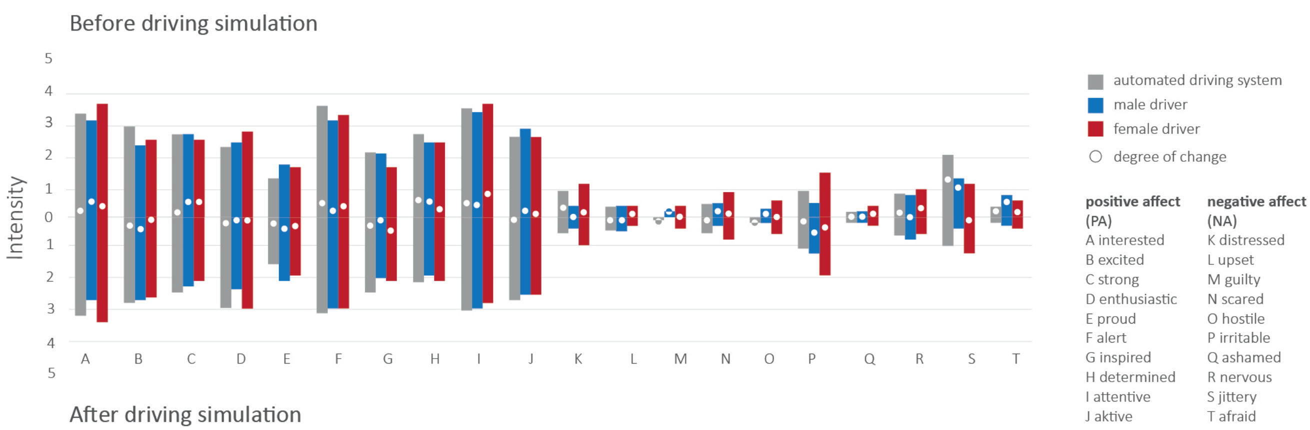

Figure 7. Results of PANAS before and after the driving experiment.

pleasure	ADS	male driver	female driver
Mean	6.20	5.67	6.20
SD	0.449	0.494	0.527
Variance	3.029	3.667	4.171

arousal	ADS	male driver	female driver
Mean	6.53	5.13	5.80
SD	0.291	0.576	0.500
Variance	1.267	4.981	3.743

Table 1. Results of Affect Grid analysis for pleasure and arousal.

We defined the mean values from the dependent variables pleasure and arousal of each group as center of blobs placed on the Affect Grid, where height and width represent the variances of the variables. One can see that all the blobs are placed in the same region (Figure 8, left subimage). Analyzing the differences between the groups concerning the level of change showed no statistically significant effects (*pleasure*: $H(2) = 2.853$, $p = 0.240$; *arousal* : $H(2) = 1.415$, $p = 0.493$), thus we conclude that the subjective self-evaluation of the subjects' current emotional state is not affected by a certain driver. All participants had a pleasant feeling while being a bit aroused or excited.

Reference to Circumplex Model

Moreover, we asked passengers of the ADS to describe driving with an AV by three adjectives. We tried to include the mentioned words in Russel' s circumplex model, as it was the basis to develop the Affect Grid. Both use the arousal-valence space [19, 13]. As shown in Figure 8 (right subimage), participants described driving with an automated vehicle mostly as interesting, exciting but also relaxing. Relaxing and exciting are contentious adjectives, but exciting and interesting are also located in the pleasure and aroused part of the grid.

Survey

As the subjective estimation of time duration can be an indicator of boredom or amusement [26], we asked all subjects to estimate the time of their trip in an unstructured survey. We used Kruskal-Walis again to compare the estimations of the three different groups. Front-seat passengers of the female driver estimated on average 8.8 minutes (SD = 0.852), passengers of the male driver estimated 8 minutes (SD = 0.552) and the group of the automated vehicle again 8.8 minutes (SD = 0.527). The subjective feeling on the trip duration thus is also not significantly affected by the type of driver ($H(2) = 1.045$, $p = 0.593$).
To gain more detailed information, we also asked them about their feelings concerning *safety, security, relaxation, control*

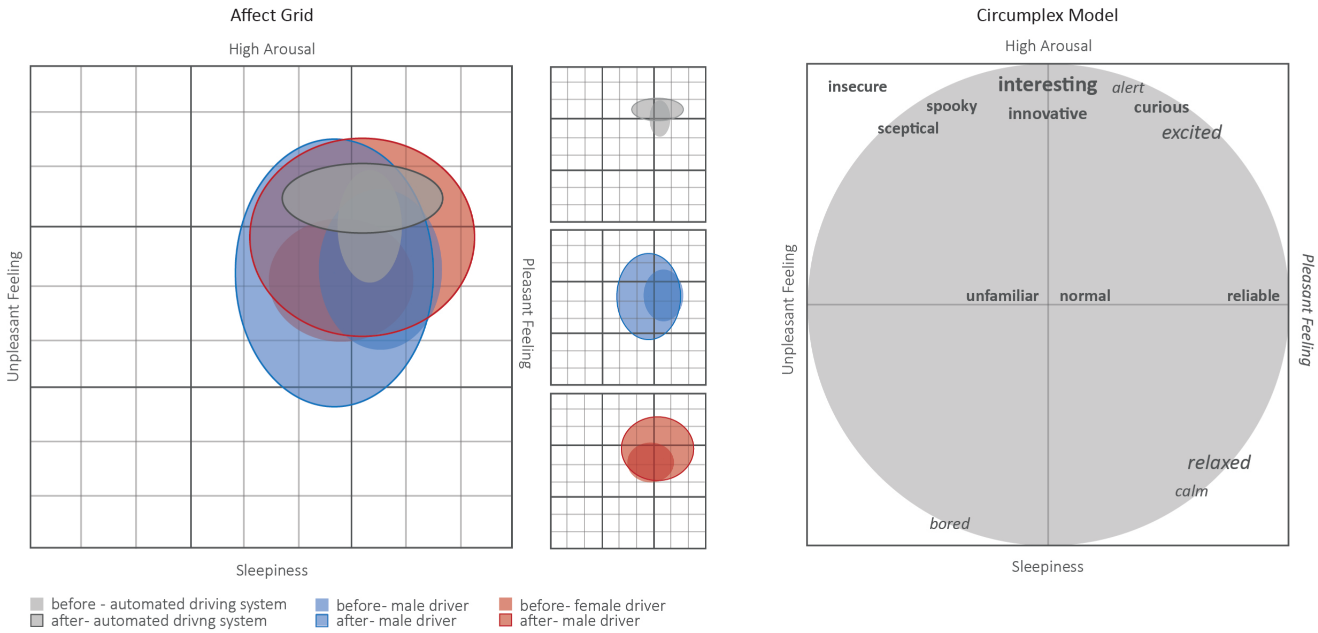

Figure 8. Results from the Affect Grid (left image) and Circumplex model (right image) with adjectives mentioned by the subjects. Adjectives already used in the model are typeset italic, new adjectives named by our subjects were arranged in the arousal-valence space according to their meaning and tyepset bold. The font size of the adjectives is proportional to the number of nomination.

and *trust* in a unstructured self developed questionnaire using a 7-Point Likert Scale. The groups do not feel different referring to the feeling safe and secure, (Kruskal-Wallis) H (2) = 1.282, p = 0.527. Regarding the medians, passengers of a female driver chose level 5 on the scale, while group male driver and group ADS chose level 6. So all agreed to feel save and secure. We asked them about their relaxation while car driving, also here there was no difference between the groups, H (2) = 2.831, p = 0.243 (female driver: median = 5; male driver: median = 6; ADS: median = 6). Moreover the results show that participants agreed that losing control, because they were not driving the car on their own was no problem for all subjects, H (2) = 0.998, p = 0.607 (female driver: median = 6; male driver: median = 5; ADS: median = 6). They all also submitted that they trusted the driver not to have an accident, H (2) = 1.346, p = 0.510 (female driver: median = 6; male driver: median = 6; ADS: median = 6). In another part of the survey, we asked about the subjects' general acceptance of AD. These questions are adopted from a 2016 survey [16] conducted in the four countries Austria (N=383), Germany (N=78), South Korea (N=81), and USA (N=324). We compared these results with ours. 64 % of our subjects responded being familiar with the concept of AD. 26.6 % replied that they have heard about it but do not know what it is and 8.9 % already have experience with AVs. This is almost similar to the VAN survey [16] where 29% have heard about it and 59% are familiar with the concept. As the best solution for AD, 66.7 % think that it is a 100% self-driving car, 24.5% wanted to keep control between phases of high automation. Only 8.9% would like to drive most of their time on their own while the automated mode is only available in certain driving situations. But the possibility of taking over whenever they want was requested by 84.5 % of our study participants. This correlates to the 2016 VAN survey, where 83 % responded likewise. Only 22.2 % of subjects in our study could imagine to own a fully automated car (18% VAN survey: 18%).

Interviews

All those results are confirmed by unstructured interviews, conducted with 14 passengers of the ADS. According to their qualitative statements (translated from German), 8 subjects confirmed that there is no difference in driving with an ADS compared to a human driver: *"It felt normal for me, there is no difference as if a person would sit next to me. I think each person I know drives much worse than an AV"*. When they were asked if they would also do a trip in reality, they agreed: *"Yes, definitely [] I think in some years it will be a standard and it will not be allowed anymore to drive on our own"*. Another subject said: *"Sure, I think errors can happen everywhere, and they happen less with computers than by human drivers who can be distracted"*. Concerning trust, one participant mentioned that for him there is no difference between public transport and AVs, where passengers also have to trust a driver who cannot be influenced. Nevertheless it was also mentioned that if there is no human being an active driver is unfamiliar and irritable. Skepticism of the concept of AD was expressed by 5 participants. They agreed they would try a trip in reality but *"[] would still like to keep control"*. Their reasons are, that they have fun driving a car but also that they do not yet trust the system totally, they trust themselves more. Only one (female) subject answered, that she does not an AV at all: *"I trust in people not in technology, if there is a human driver it feels more concrete"*.

DISCUSSION

The comparison of the heart rate variability by using the slope of the regression line of consecutive RMSSD measurements to determine an increase/decrease of stress (and correspondingly, relaxation) seems feasible for us, and the result showed an increased level or tendency of stress for all the groups, which could be interpreted as the result of multiple factors. One could be of course that our driving scenario with overtaking and fast driven hairpin curves in the mountains was perceived as dangerous, but also the participants' first experience with the high-fidelity hexapod driving simulator might have resulted in higher tension. Performing ANOVA and Kruskal-Wallis tests showed no statistically significant difference between all the three groups, thus we accept **Hi**. Nevertheless, as all of our participants were computer science students, it is hard to draw a general conclusion from these quantitative results - elderly people or humans with less technical background might have perceived an ADS differently. More interesting is the interpretation of the emotional states between the three groups. Front-seat passengers of ADSs seemed not to be very satisfied, and as male passengers of female drivers also had similar, low values, we would not connect this to the absence of other persons in the vehicle. Female front-seat passengers showed higher satisfaction (in terms of faces classified as *happy*) when performing the trip with male drivers, while on the other hand, male passengers of female drivers showed very low values of happiness. Regarding satisfaction, we found statistically significant differences between male drivers and ADSs, what is contrary to the self-evaluation of the participants. Taking a closer look on the PANAS adjectives, *enthusiastic*, *irritable* and *jittery* are the only ones showing minor differences (visual diagram inspection, no statistically significance). Overall, from all the analyzed factors (PANAS adjectives, Affect Grid, feelings reported in the unstructured survey, classified facial expressions), only facial expressions classified as *happy* between male drives and ADSs showed statistically significant differences. Therefore, **Hii**, stating that there is no difference in the emotional states between front-seat passengers of male drivers, female drivers or ADSs, has to be accepted.

Of course we did not allow our participants to perform side activities while driving, which might have had a positive influence on their satisfaction, still it could be worth trying to investigate how passengers could feel more joy and happiness when transported by an AV. More important for future HMI's, when designing automotive features that try to increase satisfaction, our study indicates that there might be different strategies needed to target male and female drivers/passengers. Regarding trust, answers from unstructured interviews also indicate that for most people it does not make much difference of whom they are passengers, some confirmed that AD will result in higher road safety, while only one out of 14 participants stated to trust humans more than machines.

CONCLUSION

This work is a contribution trying to answer the question how potential users perceive AVs and if they accept this technology. In contrast to similar studies that rely on qualitative surveys, we tried to quantify the impact of AD on the mental conditions and emotional states of front-seat passengers by performing a direct comparison with human drivers. We have shown that the choice of the driver (ADS, male or female driver) has only very little effects on their mental conditions and emotional states, while similar small differences can also be observed when comparing male and female drivers. Subjects in all groups showed slightly increased stress related parameters in provoked, potentially dangerous driving situations. Nevertheless, driving with an AV might not be per se a source of satisfaction, and similar to Roedel et al. [20] we finally conclude the need for strategies that try to maintain a pleasurable driving experience.

LIMITATIONS AND FUTURE WORK

This study leaves some room for future improvements: Compared to today's standards in computer graphics (e. g., computer games), the rendering of the environment shown in the driver simulator could have been more realistic. More convincing graphics might have led to higher immersion and, as consequence, could have resulted in a more believable perception of the predefined, dangerous driving situations. Also, using eye-tracking glasses like the Tobii Pro wearable eye tracker could be another important source for improving results, as gaze direction, eye blink frequency, and head movements are said to be trust related parameters. At the same time, however, this could have made it difficult to extract emotional facial expressions. In our view, the greatest influencing factor regarding the results is the choice of the sample: concerning user acceptance and trust, computer science students cannot truly represent the whole population and a sample size of about 50 subjects might be too small for representative results.

Our next steps include repeating the experiment, using an improved visual representation of the simulation environment and a proper eye tracker to get further insights into the monitoring process of front-seat passengers. Furthermore, we will try to allow more participants to take part in the experiment to include a greater variation of age, gender, level of education and general views on driving and vehicles.

REFERENCES

1. Prateek Bansal, Kara M Kockelman, and Amit Singh. 2016. Assessing public opinions of and interest in new vehicle technologies: An Austin perspective. *Transportation Research Part C: Emerging Technologies* 67 (2016), 1–14.
2. Jenay M Beer, Akanksha Prakash, Tracy L Mitzner, and Wendy A Rogers. 2011. *Understanding robot acceptance*. Technical Report.
3. Stephen M Casner, Edwin L Hutchins, and Don Norman. 2016. The challenges of partially automated driving. *Commun. ACM* 59, 5 (2016), 70–77.
4. R Castaldo, Paolo Melillo, Umberto Bracale, M Caserta, Maria Triassi, and Leandro Pecchia. 2015. Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis. *Biomedical Signal Processing and Control* 18 (2015), 370–377.

5. Fredrick Ekman, Mikael Johansson, and Jana L Sochor. 2016. Creating Appropriate Trust for Autonomous Vehicle Systems: A Framework for HMI Design. In *Proceedings of the 95th Annual Meeting of the Transportation Research Board, Washington, DC January 10-14, 2016*.
6. David Garcia, Christine Kreutzer, Karla Badillo-Urquiola, and Mustapha Mouloua. 2015. Measuring Trust of Autonomous Vehicles: A Development and Validation Study. In *HCI International 2015-Posters Extended Abstracts*. Springer, 610–615.
7. Anja Grießmeier. 2016. Selbstfahrende Autos in den Augen der Österreicher. *IMAS Report International* 7/2016 (2016), pp. 14.
8. Tove Helldin, Göran Falkman, Maria Riveiro, and Staffan Davidsson. 2013. Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, 210–217.
9. Sebastian Hergeth, Lutz Lorenz, Roman Vilimek, and Josef F Krems. 2016. Keep Your Scanners Peeled Gaze Behavior as a Measure of Automation Trust During Highly Automated Driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 58, 3 (2016), 509–519.
10. Andrew L Kun, Susanne Boll, and Albrecht Schmidt. 2016. Shifting Gears: User Interfaces in the Age of Autonomous Driving. *IEEE Pervasive Computing* 15, 1 (2016), 32–38.
11. Miltos Kyriakidis, Riender Happee, and JCF De Winter. 2015. Public opinion on automated driving: results of an international questionnaire among 5000 respondents. *Transportation research part F: traffic psychology and behaviour* 32 (2015), 127–140.
12. Todd Litman. 2014. Autonomous Vehicle Implementation Predictions. *Victoria Transport Policy Institute* 28 (2014).
13. Regan L Mandryk and M Stella Atkins. 2007. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International journal of human-computer studies* 65, 4 (2007), 329–347.
14. Masahiro Mori. 1970. The Uncanny Valley. *Energy* 7, 4 (1970), 33–35. in Japanese.
15. M Loretto Munoz, Arie van Roon, Harriëtte Riese, Chris Thio, Emma Oostenbroek, Iris Westrik, Eco JC de Geus, Ron Gansevoort, Joop Lefrandt, Ilja M Nolte, and others. 2015. Validity of (Ultra-) Short Recordings for Heart Rate Variability Measurements. *PloS one* 10, 9 (2015), e0138921.
16. Jeon Myounghoon, Andreas Riener, Jason Sterkenburg, Ju-Hwan Lee, Bruce Walker, and Ignacio Alvarez. 2017. An International Survey on Autonomous and Electric Vehicles: Austria, Germany, South Korea and USA. In *CHI '17: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Denver, AT, US, 2017 (under review)*.
17. William Payre, Julien Cestac, and Patricia Delhomme. 2014. Intention to use a fully automated car: Attitudes and a priori acceptability. *Transportation research part F: traffic psychology and behaviour* 27 (2014), 252–263.
18. William Payre, Julien Cestac, and Patricia Delhomme. 2015. Fully Automated Driving Impact of Trust and Practice on Manual Control Recovery. *Human Factors: The Journal of the Human Factors and Ergonomics Society* (2015), 0018720815612319.
19. Jonathan Posner, James A Russell, and Bradley S Peterson. 2005. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology* 17, 03 (2005), 715–734.
20. Christina Rödel, Susanne Stadler, Alexander Meschtscherjakov, and Manfred Tscheligi. 2014. Towards Autonomous Cars: The Effect of Autonomy Levels on Acceptance and User Experience. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, 1–8.
21. James A Russell, Anna Weiss, and Gerald A Mendelsohn. 1989. Affect grid: a single-item scale of pleasure and arousal. *Journal of personality and social psychology* 57, 3 (1989), 493.
22. Brandon Schoettle and Michael Sivak. 2014. *A survey of public opinion about autonomous and self-driving vehicles in the US, the UK, and Australia*. Technical Report.
23. Serge Thill, Maria Riveiro, and Maria Nilsson. 2015. Perceived intelligence as a factor in (semi-) autonomous vehicle UX. In *" Experiencing Autonomous Vehicles: Crossing the Boundaries between a Drive and a Ride" workshop in conjunction with CHI2015*.
24. Viswanath Venkatesh, Michael G Morris, Gordon B Davis, and Fred D Davis. 2003. User acceptance of information technology: Toward a unified view. *MIS quarterly* (2003), 425–478.
25. David Watson, Lee A Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* 54, 6 (1988), 1063.
26. John D Watt. 1991. Effect of boredom proneness on time perception. *Psychological Reports* 69, 1 (1991), 323–327.