# Driver Emotional States & Trust: Interactions with Partial Automation On-Road

Liza Dixon
Hochschule Rhein-Waal
University of Applied Sciences
Kamp-Lintfort, Germany
email: lizadixon@gmail.com

William M. Megill
Hochschule Rhein-Waal
University of Applied Sciences
Kleve, Germany
email: william.megill@hochschule-rhein-waal.de

Karsten Nebe
Hochschule Rhein-Waal
University of Applied Sciences
Kamp-Lintfort, Germany
email: karsten.nebe@hochschule-rhein-waal.de

***Abstract*—Many vehicles on-road today are equipped with Advanced Driver Assistance Systems (ADAS) which enable a driver to handover primary driving tasks to the vehicle under specific conditions, provided the driver continues to supervise the system (Level 2 automation). Various tools and methods are used in the study of human-machine interaction with vehicle automation in order to assess a driver's experience and interactions with a system. When used in-vehicle, Facial Emotion Recognition (FER) offers researchers the possibility of a quantitative reading of the driver's changing emotional state in response to interactions with the system. This paper presents a method of correlating FER data post-drive with participants' reported feelings of trust in the system. FER visualizations of the duration of the test drive sessions as well as visualizations of specific driving events are presented. Challenges in the use of FER in-vehicle, "in the wild" (on-road) are also discussed. Participants with a gain in trust post-drive and those with a loss in trust post-drive more frequently displayed the emotions *happy* and *angry*, respectively. Results indicate that trust increases after a user's first experience with an ADAS and further that FER may be predictive of user trust in automation.**

***Keywords—Facial Emotion Recognition; Driver Emotions; Advanced Driver Assistance Systems (ADAS); Human-Machine Interaction; Automation.***

## I. INTRODUCTION

Exponential improvements in computing speeds, computer vision, and machine learning over the past decade are fundamentally changing what it means to "drive a car". Inside the vehicle, a revolution is taking place—one in which the primary role of the driver is shifting to that of a passenger [1]. This shift in the role of the user presents substantial challenges in the acceptance of this technology and raises important social, ethical, and legal concerns about the future of road transportation.

From a road safety perspective, the incentive for user acceptance of vehicle automation is clear—a majority of auto accidents are due to human error, killing 1.35 million people each year and leaving up to 50 million injured or disabled, internationally [2]. In addition to a humanitarian concern, acceptance is also an economic concern for companies heavily invested in the research and development of vehicle autonomy [3]. Failure to support users in their exchange with Advanced Driver Assistance Systems (ADAS) on-road today, "will become increasingly costly and catastrophic [4]," as vehicle automation grows in its capabilities and prevalence.

The term Advanced Driver Assistance System refers specifically to the current state of the art in production vehicles, also known as a Level 2 (assisted or partially automated) system, whereas Automated Driving System (ADS) refers to vehicles that are conditionally, highly or fully automated (Level 3, 4 or 5, respectively). The pervasiveness of partial automation (Level 2) in the form of ADAS and the dawn of conditional automation (Level 3) in production vehicles necessitates a robust understanding of Human-Machine Interaction (HMI) challenges in vehicle automation [5].

Each level of vehicle autonomy requires differing levels of supervision from the user and presents various challenges from a HMI perspective [6]. While an ADAS supports the driver in longitudinal and lateral control of the system under specific conditions, the driver is still required to remain focused on the road and prepared to take control of the vehicle at all times. Because these systems are characterized by their limitations as *assistive* systems [5], ADAS involves frequent handovers of control between the system and the user, leaving substantial room for error [6], [7].

Studies have been conducted to gather insight regarding user sentiment towards vehicle autonomy, the results of which point to trust as a major factor in the acceptance of the technology [8]–[10]. Trust is defined by Mayer et al. [11] as an attitude; it is not risk-taking, "but rather it is a willingness to take risk." Adjusted for the context of automated systems by Körber, trust is "the attitude of a user to be willing to be vulnerable to the actions of an automated system based on the expectation that it will perform a particular action important to the user, irrespective of the ability to monitor or to intervene [12]." Trust affects reliance on automation, and reliance aids the user in navigating the complexities of automated systems, especially when the context of use demands adaptive behavior, as is the case with ADAS [4]. Furthermore, a study of trust in automation by Miramontes et al. [13] concluded that people with "high emotional stability…reported higher levels of trust in automation."

Whether it is getting cut off in rush-hour traffic or a series of traffic lights ahead signaling green—the act of driving can be an emotional roller-coaster. In turn, a driver's emotional state influences their driving performance. A literature review of emotion on the road by Eyben et al. [14] reveals that "happy drivers are better drivers" and that "aggressiveness and anger are emotional states that extremely influence driving behavior and increase the risk of causing an accident." As a tool, Facial Emotion Recognition (FER) can facilitate a deeper understanding of the user experience in-vehicle [15], [16] and aid in the identification of drivers with a tendency towards *happiness* and those with a tendency towards *anger* on the road. In theory, it is possible that FER may be indicative of a users' trust in automation.

While closed courses and simulators provide a stable research environment for human-machine interaction with vehicle automation, they are not fully aligned with the context of use, nor the state of the art. The release of one's personal safety to the system occurs exclusively while engaging with partial automation in an on-road setting. Hence, existing research is unable to specify how exposure to and experience with a Level 2 system on-road might impact user trust. While Facial Emotion Recognition is typically used to assess the user experience, it has yet to be applied in an on-road study in correlation with driver trust in vehicle automation. In order to address this, the following research questions were explored in an experiment:

Q1: How does a driver's first experience with an Advanced Driver Assistance System on-road affect their level of trust in the system?

Q2: Which emotions (*neutral, happy, surprise, angry*) do first time drivers of an ADAS display and is there a relationship between the emotions displayed and their reported levels of trust in the system?

Q3: Which emotions (*neutral, happy, surprise, angry*) accompany specific assisted driving events?

Incorporating a mixed-method approach utilizing verbal trust scores, the Trust in Automation questionnaire [12], and Facial Emotion Recognition, and qualitative/observational data, it was expected that the results of the experiment would indicate the following:

H1: Participants will report higher levels of trust in ADAS after their first experiential drive with an ADAS.

H2: FER analysis will reveal a relationship between a participant's Trust in Automation score and their emotions displayed during the drive.

H3: Participants will display varying emotions according to the driving scenario and behavior of the ADAS, including an initial emotional response to a driving event, followed by emotional resolution i.e., a return to their respective normal emotional state as defined by FER.

Section II of this paper provides related works and background information, and Section III discusses participant demographics, the technical capabilities of the vehicle utilized for test drive sessions, and the experiment procedure. Results are reported in Section IV and Section V summarizing qualitative and quantitative findings, respectively. Section VI is a discussion of the results, followed by Section VII, which outlines the limitations of this study. The paper concludes with Section VIII, which offers an outlook and future work.

This paper is an extension of work originally presented in VEHICULAR 2019: The Eighth International Conference on Advances in Vehicular Systems, Technologies, and Applications [1]. This work varies from the original in the following ways, adding: 1) more background on driver emotions 2) an additional research question (Q3) 3) more detailed information about the participants in the study 4) the results of an exploratory analysis of each participant's emotional journey throughout the experiment (Section V.C) and during a selection of specific driving events (Section V.D.) 5) discussion about the limitations and difficulties of working with FER in-vehicle (Section VI and Section VII).

## II. Related Work

Experiments to measure trust in vehicle automation have been carried out in both closed courses [11][12] and simulators [19], [20]. For example, in an experiment with 72 participants, Gold et al. [19] utilized a driving simulator modeling a Level 3 system to "investigate how the experience of automated driving will change trust in automation and the attitude of the driver towards automation." A questionnaire was administered before and after a 15-20 minute driving experience. Gaze behavior was also recorded in an effort "to measure a change of trust by a change in [eye] scanning behavior." The results of this study revealed that participants reported a higher level of trust in automation after the driving experience, however gaze behavior could not be established as a valid measurement.

In an experiment using a Wizard of Oz setup (simulating an automated vehicle), Ekman et al. [18] explored a mixed-methods approach for the assessment of trust during a 15 minute drive on a closed course with 18 participants. The results of the study indicated that "data should not only be collected at the very end of a trial only but be complemented with data collection also during a trial, in particular in relation to events that may influence and contribute to a user's overall experience."

Researchers have developed frameworks, models and scales for the assessment of user trust in vehicle autonomy. Ekman et al. [17] constructed the Lifecycle of Trust (LCoT) framework, to serve as a tool for HMI design. The LCoT identifies 11 trust-affecting factors throughout the *Pre-Use Phase* (*Implicit/Explicit Information*), *Learning Phase* (all activities from *Entering the Vehicle* to transitions from *Manual to Automated Control*, to *Exiting the Vehicle*) and *Performance Phase* (covering *Continuous Usage, Change of Context* & *Incidents*). Validation of LCoT factors, specifically through the *Pre-Use* and *Learning Phases* are an area of interest for this study, as it is the most current, comprehensive framework for understanding the development of trust in automation.

Based on empirical research, Jian et al. [21] developed the "Checklist for Trust between People and Automation" a 7-

point Likert scale comprised of 12 questions designed for use as a general scale in any area where human-automation interaction occurs. Based on this, the work of Mayer et al. [11], Lee & See [4] and others, Körber [12] developed a refined model of Trust in Automation (TiA) with an accompanying 19-item, 5-point Likert scale questionnaire covering the following factors: *Reliability/Competence, Understanding/Predictability, Intention of the Developers, Familiarity, Propensity to Trust* and *Trust in Automation*. The questionnaire features questions, such as "The system is capable of interpreting situations clearly," (*Reliability/Competence*) and inverse items such as "The system reacts unpredictably," (*Understanding/Predictability*) which correspond to the underlying factors. To the knowledge of the authors, this questionnaire has yet to be applied in a study of trust in partial automation on-road.

### A. Driver Emotions

Facial Emotion Recognition via the analysis of facial expressions extracted from images/video frames is of growing interest to HMI researchers in the area of automated driving. FER consists of three main events: 1) face and facial component detection, 2) feature extraction, and 3) expression classification [22]. The facial expressions which are associated with the emotions *happy* (*joy*), *anger*, and *surprise* are thought to be the most relevant in the context of automated driving and are used by commercial software companies in their analysis [16]. Studies have confirmed that the emotions *happy* and *angry* are the most influential on how a car is driven [23].

When used in-vehicle in real time, systems can perhaps use FER data to align its behavior with the changing emotional state of the user. This concept of system adaptation, was originally presented by Picard [24] and is known as *affective computing*. Picard suggests that computers which "...sense and respond to users' emotional states may greatly improve human-computer interaction" [4]. Affective computing by means of FER is of particular interest in vehicle automation, as improving the naturalness of interaction with the system facilitating trust and acceptance [25]. Further, according to Lee & See [4], understanding "Emotional response to technology is not only important for acceptance, it can also make a fundamental contribution to safety and performance."

A well-designed, supportive system assists the driver in achieving their goals by taking over the monotonous tasks of driving and correcting errors such as inattentiveness, smoothing out the driving experience. However, a system which is not well-aligned with the user or behaves unexpectedly may cause friction, resulting in displeasing emotions and disuse of the system. Acceptance of automated driving systems and appropriate reliance from the user is crucial to ensure its efficient and safe use [4]. Understanding the emotional states of a driver before and after driving events, while operating an ADAS on-road may support affective computing techniques in next generation vehicles [14] and in turn, trust in automation.

When used as part of a mixed-method approach in post-production, FER may therefore enable the observation of correlations between driver emotional states, vehicle behaviors and reported trust in automation.

## III. Methods

A total of n=10 participants were introduced to the same Level 2 vehicle and completed one individual test drive session. All participants completed their test session within the same two-week period. The driving route included driving time on the autobahn (including a construction zone), country roads and in urban settings. Each experiment session lasted 1 hour and 30 minutes, approximately an hour of which was driving time. The same moderator accompanied all of the participants; participants were not explicitly told to activate the ADAS. A pilot test was conducted to refine the experiment structure and equipment, after which it was determined the route did not include enough autobahn time and was therefore revised.

### A. Participants

Of the ten participants selected for this study, there were six females and four males. All participants were members of the university community and were recruited via email and flyer. Participants were screened prior to the experiment session to ensure they met specific requirements for the study: holding a valid driver's license, experience with automatic transmission, have no prior experience with the vehicle class (Mercedes-Benz GLC), not own or regularly operate a Mercedes-Benz, have no prior first-hand experience with ADAS, any semi-autonomous or autonomous vehicle systems (including for example: autopilot systems, adaptive cruise control or lane keeping assistants. Excluding: standard cruise control/speed limiters, back up cameras or blind spot assistants).

The driving route was designed as a loop, beginning and ending at each participant's respective campus. All participants drove the same section of the autobahn, with additional driving time in country road and urban settings which varied slightly based on the participant's starting location. Participant 1 (going forward, participants are referred to as "PX" where X is their individual coded reference number), P2 and P3 began at campus A, while P4-P10 began at campus B.

TABLE I. Participant Profile

| Participant | Age | Gender | Education | *Technical |
|---|---|---|---|---|
| P1 | >30 | F | Vocational | No |
| P2 | <30 | F | Vocational | No |
| P3 | <30 | F | Vocational | No |
| P4 | >30 | M | Masters | No |
| P5 | >30 | F | PhD | Yes |
| P6 | >30 | F | Masters | Yes |
| P7 | <30 | M | Vocational | No |
| P8 | <30 | M | Vocational | Yes |
| P9 | >30 | M | Masters | No |
| P10 | >30 | F | Masters | Yes |

*Denotes whether or not the participant had a computer science or other engineering background.

The mean age of the participants was $M$ = 31.66 years ($SD$ = 9.17, ranging from 20 to 48 years old). Six of the participants had been driving for over ten years while four had been driving for ten years or less. Educational level was split 50/50 between the participants, half holding a master's degree or above and the other half having received vocational training. Table I profiles each participant.

*1) Disclosure:* Participants were informed only that they would be taking part in a study of Human-Machine Interaction in ADAS which involved an on-road test drive. The focus of the study being specifically about their trust in the ADAS was intentionally withheld from the participants. It was not disclosed explicitly nor by accident (e.g., titles removed from trust questionnaire, discussion about experiment sessions prohibited) in order to mitigate the *Hawthorne Effect*. This effect refers to the inclination of research participants to adjust their behavior and act in a way that they believe is aligned with the expectation of the moderator [26]. This decision was also made in part due to the high cognitive demands of the experiment [14] (driving an unknown vehicle with unfamiliar technology on public roadways, while under observation) and to obtain unbiased and natural reactions in any participant commentary related to the discussion of trust in the system.

### B. Vehicle

The same Mercedes-Benz GLC-250 4Matic was driven by all participants. This vehicle was equipped with the Driving Assistance Package Plus option which includes an Advanced Driver Assistance System (sub-systems relevant to this study are listed in Table II). These features qualify the vehicle as a partially automated, Level 2 system [5].

TABLE II. SELECT DRIVER ASSISTANCE PACKAGE PLUS FEATURES

| Feature | Function | Active |
|---|---|---|
| Distance-Pilot DISTRONIC with Steering Assist and Stop&Go Pilot | "Autonomous intelligent cruise control system" able to accelerate and decelerate according to traffic conditions. Steering interventions help the driver stay in lane. The system can follow the vehicle ahead even where there are no or unclear lane markings (<130 km/h). | 0-200 km/h, driver activated |
| Hands-Off Warning | A haptic (steering wheel vibration) and graphic warning (in the multi-function display, next to the speedometer), alerts the driver to return their hands to the wheel. If this is not heeded, it is enhanced via an auditory warning tone. | Active with DISTRONIC |
| Active Lane Keeping-Assist | Detects unintentional lane drift by monitoring road markers. Can tell if the vehicle veers out of lane without signaling, and will vibrate the steering wheel. Brakes individual wheels for correction, keeping the vehicle within the road markers. | 60-200 km/h, (conditional) |
| PRE-SAFE® Brake with Pedestrian Detection | Able to detect pedestrians ahead and will apply the brakes automatically. | Up to 50 km/h |
| Traffic Sign Assist | Identifies traffic signs and speed limits on the instrument display via camera and GPS data. | Always active. |

Source:[27].

The purpose of this study is not to cross-compare various technologies, but rather, to analyze the inherent trust in a particular vehicle's systems, holding this as a constant.

### C. Procedure

In order to ensure consistency and objectivity between the experiment sessions, the moderator adhered to a set procedure (see Table III) and script. At the start of the session, the moderator greeted the participant outside of the vehicle in the parking lot. This is when the participant was first exposed to the make and model of the vehicle. The participant was invited to enter the vehicle, where they were then interviewed regarding their initial impressions of the vehicle, Mercedes-Benz, thoughts about ADAS, vehicle autonomy, and their expectations of the system, including their initial feelings of trust in the system. They were then asked to give a verbal rating of trust in the ADAS on a scale from 1 to 5 (1=low, 5=high). Next, they watched an introductory video featuring original content from Mercedes-Benz, which was produced for this experiment by the researchers to reflect the capabilities of the specific vehicle used for testing. The participant was informed that they were in full control of the vehicle at all times and responsible for obeying all traffic laws and posted signs. Next, the Trust in Automation questionnaire (modified from [12]) was administered to the participant in their native language (German or English, translation from [12]). After, they were encouraged to ask questions, to ensure their understanding of the system's functions and capabilities. They were asked to rehearse how to activate/deactivate the system while the car was parked. Following the introduction, each participant was asked for a second time to give a verbal rating of trust in the ADAS. Trust in Automation, Pre-drive vs Post-drive.

TABLE III. EXPERIMENT PROCEDURE

| Pre-Drive | | Drive | Post-Drive |
|---|---|---|---|
| ***Introduction I*** | ***Introduction II*** | ***Test Drive*** | ***Closing*** |
| 1) Interview<br>2) VTS #1 | 1) Intro video<br>2) TiA #1<br>3) Interview<br>4) VTS #2 | 1) Planned route<br>2) Think aloud<br>3) FER | 1) TiA #2<br>2) Interview<br>3) VTS #3 |

VTS = Verbal Trust Score, TiA = Körber's Questionnaire for Trust in Automation, FER = Facial Emotion Recognition

As the participants began the test drive with the route pre-programmed into the vehicle's GPS, the GoPro cameras were activated at 60fps. The driver-facing camera was mounted to the windshield to the right of the steering wheel for later FER analysis. The driving scene (roadway ahead), multi-function display, and participant's interaction with the system's interface was were captured by a second camera mounted behind/next to the driver's right shoulder.

During the test drive, the moderator did not give any tasks to the participants other than to follow the route on the GPS. The moderator played an observatory role, giving instruction only when prompted (e.g., clarifying a system limitation). Participants activated the system only as they felt comfortable, in the appropriate conditions and were encouraged to think aloud [4] while doing so. Participants were asked to state aloud whether they or the car was performing certain actions (steering, braking, acceleration/deceleration) throughout the drive and to share their thoughts on the vehicle's behavior as it occurred. Top speed with ADAS active was recorded for each participant as well as adjustments in posture (positioning of hands, arms and feet on/off pedals). Immediately following the drive, the TiA questionnaire was administered a second time. Participants then completed a post-drive interview and gave a final verbal rating of trust in the ADAS on a scale from 1 to 5 (1=low, 5=high), based on their experience. All interview audio was recorded for later reference.

*1) Data Analysis:* The responses to the TiA questionnaire were scored following the procedure used by the System Usability Scale [28]. Adjusted for the number of questions, responses were reverse coded, added together and then multiplied by a factor to convert the original scores of 0-68 to a 0-100 value, in order to better identify discrepancies in participant's pre-use and post-use scores (the factor *Familiarity* was removed from analysis, as all participants were selected purposefully to have no prior experience with the technology). A Wilcoxon Signed-Rank Test was used to examine differences in pre-drive and post-drive, reverse coded TiA questionnaire medians. This method was chosen as it is appropriate for the comparison of medians in ordinal data from related groups with a symmetrical distribution [29]. Wilcoxon was performed for all TiA factors together (*Reliability/Competence*, *Understanding/ Predictability*, *Intention of the Developers*, *Propensity to Trust* and *Trust in Automation*) and for each factor's respective set of questions. Friedman's Test (adjusted for ties) was used at to analyze shifts verbal trust scores (pre-introduction, post-introduction and post-drive). Friedman's was selected as the data is ordinal, came from a single group measured at three intervals, and there are no interacting effects between the groups [30], [31]. Statistical analysis and plotting of TiA and verbal trust scores was completed in RStudio [32] using the *stats* [33], *agricolae* [34], and *ggplot2* [35] packages.

*2) Facial Emotion Recognition:* Driver facing video footage captured at 60fps was processed by a convolutional neural network (CNN) with 3 convolutional layers and two fully connected layers (including the output layer). The CNN was trained for the facial emotion recognition of seven emotions [36], however a reduced set of emotions was selected for analysis: *neutral*, *happy*, *surprise*, and *angry* [14], [16]. Classification performance using this set of emotions was reported 81% accurate by Mathworks MATLAB 2018b [37], which was used to run the network and output the data in text files. The text files were then compiled, cleaned and analyzed in RStudio.

The analysis returned a value from 0-1.0 for each emotion, (where *neutral*, *happy*, *surprise* and *angry* share a portion of the 1.0 value) for each frame and output the data in text files. The text files were then compiled and plotted in MATLAB using movmean [38]. FER scores were reviewed for each test drive as a whole, as well as for specific driving events. These events included: removing hands from the wheel (Hands Off), strong deceleration and steering through curves (DISTRONIC PLUS with steer assist), full stop and restart (Stop&Go Pilot), driver intervention, and a vehicle malfunction.

## IV. Qualitative Results

Participant commentary from the interviews (pre-intro, post-intro, post-drive) and during the test drives was recorded. This Section includes excerpts from the commentary and participant behaviors recorded. The commentary was transcribed and categorized according to the TiA factors examined by the questionnaire [12], and additionally participant *Driving Style* and *Weather*.

### *A. Reliability/Competence*

Toward the end of the drive, P9 said, "I think the benefits [of ADAS] are clear and undeniable. Every system here is intended to improve safety. I don't think there's any danger posed by the system," and "I was skeptical. Having seen it in action, having felt it under my hands…it is a good thing and I could recommend this kind of system to other people as well. I could talk positively about my experiences on the road. I wouldn't be averse to having this kind of system in my own vehicle."

P4 took back control while in an autobahn construction zone due to discomfort with Steer Assist, stating, "It's keeping us in the lanes but before it was a little bit problematic. It went too far to the right and then it went to the left and then I intervened because I was not sure if it would do it itself."

After the drive, P4 said, "There were a couple of mistakes [with Steer Assist] and it was not too clear to me if it was on or off. I guess that's not the point of the system, that I have to focus more on the [system] than the road. It's not as useful because I have to keep my hands at the wheel anyway." During the course of the test drives the system experienced one malfunction, which occurred during P4's test drive. While driving on a country road, the system drifted the vehicle out of its lane. P4 allowed the vehicle to continue drifting out of lane until half of the vehicle was in the lane of the oncoming traffic before intervening.

### *B. Understanding/Predictability*

After the introduction to the ADAS, P7 stated, "I was curious and skeptical at the beginning but now that I know more about how [the systems] work and what they can do for me, it makes me more confident. I think I may struggle with using them due to a lack of experience. I can trust [assistance] more than a full take-over of my driving."

P7 expressed how unpredictable vehicle behaviors affected feelings of trust, "I felt that I was mostly in control of the system, but not when resuming my settings. I knew I was faster when I last used the feature, so I wanted to use it to accelerate. But sometimes it was much faster than I expected. It was alarmingly fast. I did not trust in the braking after such a strong acceleration."

*1) Mental Model:* Several participants made comments that revealed changes in their mental model of the system as the drive progressed.

During the test drive, when P5 activated the system, they did not release their foot from the pedal until ~40 minutes in to the drive, which automatically caused the distance control system (Distance Pilot) to become passive, leading P5 to become confused about the system's functionality.

P2 said, "I see the lane is not there, so I will not trust [the system.]" and "It is steering but I want to make sure I keep my hands on the wheel because there is no lane marking right there." P6 expressed, "Maybe the [steer assist] is off because it has to "take some information in to analyze the situation," indicating their perception of the system's functionality. P9 stated, "I thought maybe if I moved to the right a little bit, [the system] would start to see a pattern in the lanes. I was trying to show the car what the lane looks like. I thought maybe if you just adjust the position of the car within these lanes that it could find a pattern within it, that it could orient itself."

P8 mentioned a time they adjusted their mental model and hence behavior, "Most times I felt in control. Except for the two times when the car ahead [moved into the adjacent lane and] turned off the road. I thought the system misinterpreted it a bit. I anticipated it the second time, based on the first time, that it could happen, and my anticipation was right."

### *C. Intention of the Developers*

Participants (P1, P3, P5, P7) said that they felt "safe and comfortable" in the vehicle due to the brand. Several mentioned that they would prefer to drive a vehicle from another brand (P1, P8, P9). After an introduction to the system, P9 said, "I wouldn't say I necessarily feel better about [the system]…I think the developers did their best to create a system that doesn't put people in harm. I have faith in them but there are so many variables on the road that I don't feel comfortable putting my full trust in the system."

*a) Implicit Information:* P9 referenced stories they have heard in the media: "On the news you'll see, in the United States in particular, where people are very excited about automated driving systems, that someone runs into something because they are not paying attention and then ends up in a fatal accident. That puts me on edge about the whole thing. I am not exploding with excitement [to use the system]."

### *D. Propensity to Trust*

In the initial interview, P1 said, "I am like a dinosaur. I don't have a big trust in the system. I feel safe with things I can see." After an introduction to the system, P1 expressed their feelings towards driving: "If I drive, I am concentrated on it and I like it. I would not like a car to do things for me that I could do myself." P7 stated, "I am curious and also suspicious if it really works. I think that I can do better than the automatic calculations of the system. It is making me curious but also cautious."

### *E. Trust in Automation*

P5 expressed a desire to test this system but at each opportunity they intervened and took back control; in the end, stating, "I can't trust the system so fast. You would have to put a wall out of cushions in front of me before I try that" (referring to the Stop&Go Pilot). After the test drive, P5 stated "It's up to me how much I trust the system. If I would drive with such a system for a long time, I would put more trust in it. But now it's a very big contrast in driving for me."

P6 said, "I am not totally trusting but for me a [verbal trust score of] 4 is very high because I normally do not like these systems. But it is very comfortable to me now."

Throughout the test drive P10 was animated and expressive stating at the end, "It is a bit creepy for me…to trust a car. Normally, you trust a driver. I hate to go by airplane. Because you have to trust someone else. A stranger! But here...you have to trust a car...something with no inside, no feelings! It is only a system, a machine, and you have to trust it. The longer you drive it and the more you get familiar with it…you get a feeling for it and you start to trust the system."

### *F. Driving Style*

During the test drive, P9 reflected on the effect the ADAS might have on driving styles, "I could imagine the system really reducing reckless driving. I don't feel the need to even worry about passing this person. I kind of just feel comfortable letting the car takeover. It kind of takes the pressure off me to take some sort of an action. If driving manual transmission, I would probably be more aggressive right now."

P9 stated they would "feel better" if other cars around them were using ADAS. "There are a lot of really bad drivers. I would know the car is going to adjust to keep them within tolerance limits automatically. I would probably feel better about being in traffic with the person." P10 said, "Maybe I would pay more attention to a car [with ADAS]. Because I know it is new a technology. But you can expect more of what a system would do than a human. A system would work or not work, not be in-between like humans. Maybe these are the cars on the road now that you think, 'Oh that driver drives very correctly.'"

*1) Risk-taking:* Shortly after expressing their apprehension at the start ("I am very nervous about what we're going to do today"), P3 engaged in repeated attempts to test the system at high speeds while on the autobahn. P3 intentionally drifted over lane markings several times to see if the Active Lane Keeping system would reorient the vehicle properly in the lane.

20 minutes into drive, P8 "provoked" the system, stating, "Now [the system] steered, because I provoked it. I tried to go straighter than I should have into the turn. If you get the angle of the curve wrong, it's nice that someone assists you with it."

When testing the Stop&Go Pilot, P7 said, "The car will stop? Cautiously, I am trying that. I've got my foot over the pedal. The car…the car completely stopped. It is new and weird. Okay, wow. Now it's going again on its own. If I know this is doing the job for me, I feel comfortable in releasing my foot and not keeping it directly above the pedals." P9 stated, "So it's going to stop completely? That's giving me a little bit of apprehension right now. That feeling, do I let it? Cause that's like twenty years of driving experience inching up towards that bumper."

Several of the participants drove through or attempted to drive through roundabouts in urban settings while the system was active. P4 followed a vehicle in front through the roundabout and out of the second exit. P8 also followed a vehicle into the roundabout but intervened as the system accelerated once the vehicle in front exited. P6, P7, and P10 approached the roundabout with ADAS active but intervened.

P4 was the only participant who activated the ADAS near its top speed at 190 km/h.

*a) Hands Off:* All participants with the exception of P5 removed their hands from the wheel long enough to trigger the hands-off warning graphic and/or auditory warning tone.

Halfway through the drive, while traveling on the autobahn (160 km/h), P1 crossed their arms. P1 also adjusted the position of their headrest at high speed, stating “See, this is something I would do now, because the system is on” as they put both arms behind their head and adjusted the headrest.

P4 stated, “You get a warning to put your hands on, but you don't have to do anything, it’s just going on its own anyway." P4 discovered a work-around for disabling the hands-off warning; by briefly nudging the steering wheel slightly from side to side they were able to cease the warnings temporarily. P4 continued to workaround these warnings, through a narrow construction zone.

P3 told the hands-off auditory warning to “shut up” and P4 referred to it as “annoying.” P4 received the most hands-off warning notifications of all participants (over 45 notifications).

*2) Risk-aversion:* Participants P1, P2, P5 and P6 did not allow the Stop&Go Pilot to come to a full stop while all other participants did. P2 had the lowest top speed at 130 km/h and possessed the most cautious driving style.

### G. Weather

One instance of rain occurred which lasted approximately 15 minutes toward the end of P3’s test drive. P3 said, “I like the system. I trust the system. But because of the rain I have not such a safe feeling because I don’t know this car and I am driving it for the first time. It’s not like you just sit here and feel safe, because it’s up to all of the things that can happen around you.” At the end of the drive, P3 stated, “If there was no rain, I would give the system a [verbal trust score of] 5 because it worked, and it did what it was supposed to do. Because of the rain, I didn’t feel so safe, so I will say 4.”

## V. QUANTITATIVE RESULTS

### A. Verbal Trust Scores

Participants were asked to give a verbal rating of trust in the vehicle’s ADAS on a scale from 1 to 5 (1=low trust, 5=high trust) three times throughout the experiment session: pre-interview, post-introduction, and post-drive.

Verbal trust scores indicated that six of the participants had an increase in trust after receiving an introduction to system (ranging from +0.25 to +1.50 compared to pre-interview scores). Three showed no change and one reported a decrease in trust (-0.50). Post-Drive, seven of the participants reported an increase in trust (ranging from +0.50 to +2.50), while one reported no change and two reported a decrease in trust (-1.0). A non-parametric Friedman test of differences among repeated measures, adjusted for ties was conducted and rendered $\chi2(18)=0.07$, $p>0.05$, which was nearly significant (see Figure 1).

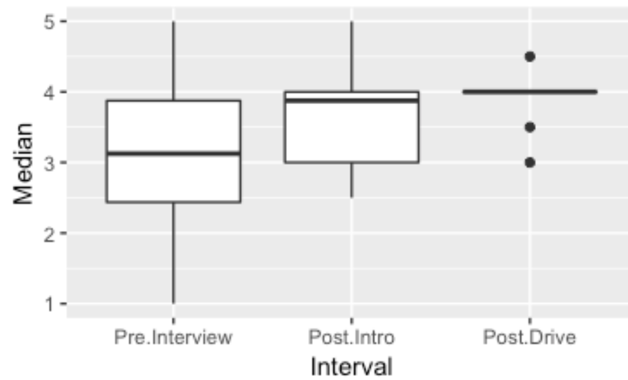


Figure 1. Comparison of median verbal trust scores for all participants at key intervals during the experiment, which were nearly significant at p>0.05.

### B. TiA Questionnaire

The Trust in Automation questionnaire was administered twice during the experiment session: after participants were introduced to the system (pre-drive) and again after the test drive (post-drive). Scored results (see Section III.C.1. *Data Analysis*) from the questionnaire indicated that three participants had a decrease in TiA after the test drive (P1, P5, P4) while all other participants reported an increase of TiA after the test drive (see Table IV). All of the participants who reported a decrease in trust after the drive were >30 years of age with ten or more years of driving experience.

TABLE IV. TRUST IN AUTOMATION SCORES: PRE-DRIVE VS POST-DRIVE

| Participant | Interval | | |
|---|---|---|---|
| | ***Pre-Drive*** | ***Post-Drive*** | ***Change in TiA*** |
| P9 | 57.408 | 82.432 | + 25.024 |
| P8 | 54.464 | 72.128 | + 17.664 |
| P10 | 45.632 | 61.824 | + 16.192 |
| P7 | 47.104 | 63.296 | + 16.192 |
| P2 | 64.8 | 75.1 | + 10.3 |
| P3 | 63.3 | 72.1 | + 8.8 |
| P6 | 57.408 | 61.824 | + 4.416 |
| P1 | 50 | 47.8 | - 2.2 |
| P5 | 75.1 | 63.3 | - 11.8 |
| P4 | 54.5 | 35.3 | - 19.2 |

By participant, pre-drive and post-drive. An increase in TiA occurred in all participants while a decrease in TiA was observed in P1, P4 and P5.

In one instance, a participant’s (P1) verbal trust scores did not align with their self-reported TiA responses. P1 verbally reported a gain in trust post-drive, but in the post-drive questionnaire reported a loss of trust.

Each participant’s TiA response (based on the 5-point Likert scale) was recoded, and a pre-drive median and post-drive median value was given for each participant. A Wilcoxon Signed-Rank Test performed on all participant’s pre-drive and post-drive medians indicated that the post-drive TiA median scores were not significantly higher than pre-drive median TiA scores ($Z$=-0.86, p>0.05) (Figure 2, factor: all).

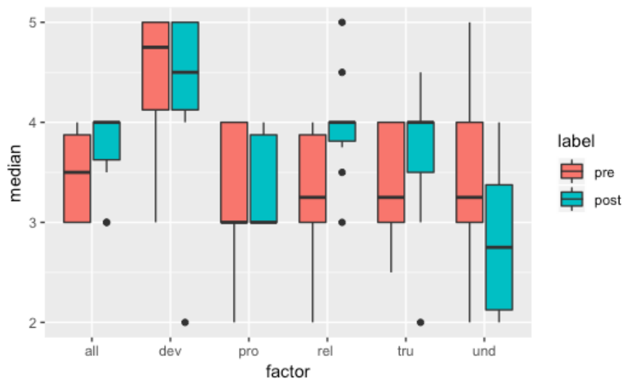


Figure 2. Trust in Automation, pre-drive (pre) & post-drive (post) medians. Shown overall (all) and for each factor (dev=*Intention of the Developers*, pro=*Propensity to Trust*, rel=*Reliability/Competence*, tru=*Trust in Automation*, und=*Understanding/Predictability*).

The factors *Reliability/Competence*, *Understanding/ Predictability*, *Intention of the Developers*, *Propensity to Trust*, *Trust in Automation* were also considered individually for analysis. The post-drive median score for *Reliability/Competence* was found to be significantly higher than the pre-drive median score ($Z$=-2.17, p<0.05) (see Figure 2, factor: rel). The pre-drive vs. post-drive median scores for the other factors were not found to be statistically significant at p>0.05.

### C. Facial Emotion Recognition

The driver facing camera footage for each participant's test drive was processed by the convolutional neural network for FER. A value ranging from 0 to 1.0 for each emotional state (where each of the four emotions, *happy*, *angry*, *surprise* and *neutral* share a portion of a 1.0 value) were returned every one tenth of a second for the entirety of the drive. FER scores were calculated for the entire duration of each participant's test drive, which lasted approximately one hour. The visualization of each journey over time was plotted in MATLAB. Frames in which no face was detected were included in the analysis with *movmean* and the *dim* property set to 5,000 (smoothing) (Figure 5, see next page). Each graph is labeled with the respective participant number, their FER score for four emotions out of 1.0 is shown in the y-axis. As shown in the legend below the plots, *neutral* is plotted in blue, *happy* is plotted in red, *surprise* is plotted in yellow, and *angry* is plotted in purple. The x-axis plots location in time, in which *Start* marks the beginning of the test drive and *End* marks the end of the test drive. Entry (EA) and exit of the autobahn (XA) are also annotated (see the caption of Figure 5).

### D. Facial Emotion Recognition + Sample Events

A selection of driving events was timestamped and extracted from FER analysis and the visualization of each event was plotted with movmean in MATLAB with the *dim* property set to 100 (smoothing). The selection here represents only a portion of the driving events which occurred. As shown in the legend for each graph, *neutral* is plotted in blue, *happy* is plotted in red, *surprise* is plotted in yellow, and *angry* is plotted in purple.

#### 1) Hands Off

While the Mercedes-Benz ADAS used in this study is not a hands-free system, a majority of participants elected independently to remove their hands from the wheel at some point during the test drive. The steering wheel was equipped with sensors (Mercedes-Benz, standard production) as a means of measuring driver attentiveness. When the system detects a lack of contact with the steering wheel, the system will display a graphic warning indicating the driver should return their hands to the wheel. If this warning is not heeded, it is then accompanied by an auditory tone. Figure 3 highlights the emotional state of the participant as they experiment with removing their hands from the wheel for the first time.

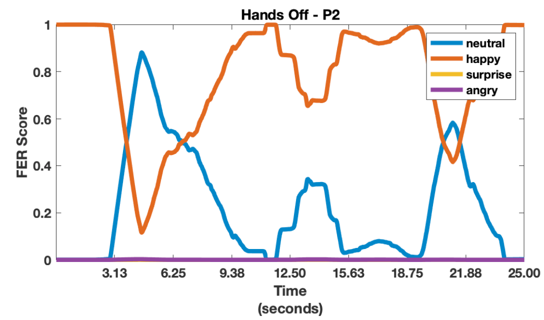


Figure 3. Hands Off, Participant 2. At 3.33 seconds the driver removed their hands from the wheel. At 10-15 seconds, they stated, "This is interesting. It keeps the [lane] all by itself." At 21 seconds, the driver said, "Hmm. Crazy."

#### 1) Deceleration

When active in an autobahn setting, DISTRONIC PLUS (Mercedes-Benz adaptive cruise control) is capable of autonomously decelerating the vehicle at high speeds. Figure 4 and Figure 6 are samples of participants emotional state as the vehicle decelerated autonomously.

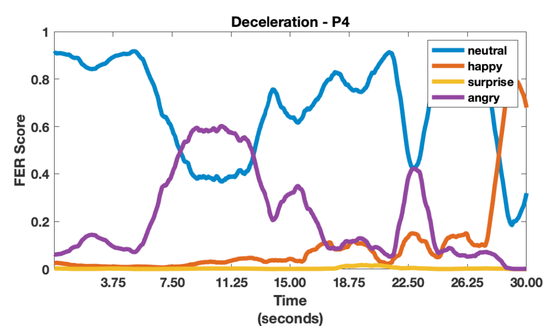


Figure 4. Deceleration, Participant 4. While on the autobahn with ADAS active near top speed (190km/h), the vehicle slowed itself autonomously, approximately 50% to 90km/h. The vehicle begins decelerating at 4 seconds, and the driver says, "Ok!" At 13 seconds the driver says, "That was not too comfortable for me, but I guess that's because my car doesn't brake so hard." At 23 seconds they state, "I would have braked way before that."

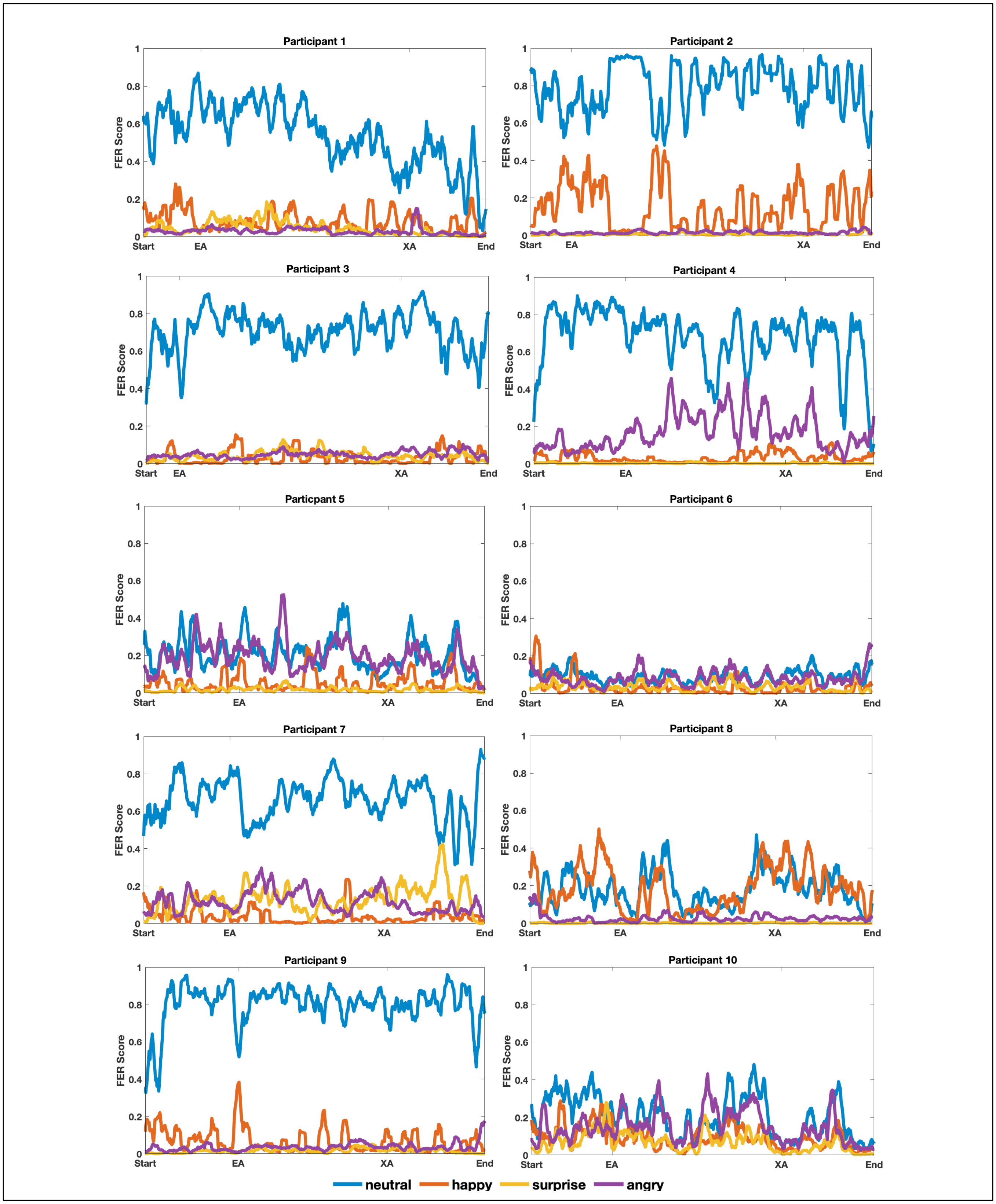


Figure 5. Facial Emotion Recognition Scores for all participants' entire test drive session. Y-axis expresses participants' FER scores from 0-1.0 where all emotions share a portion of 1.0. X-axis shows location in time, in which *Start* marks the beginning of the test drive session, *EA* marks "Enter Autobahn", *XA* marks "Exit Autobahn", and *End* marks the end of the test drive session. Between *Start* and *EA*, *XA* and *End*, participants drove in country road and urban settings. Between *EA* and *XA* participants drove on the autobahn. Test drive sessions averaged 70 minutes in total. P5, P6, P8, and P10 wore corrective eyeglasses which impeded facial detection. This is visible in the relatively low FER scores of these participants (see Section VI for more).

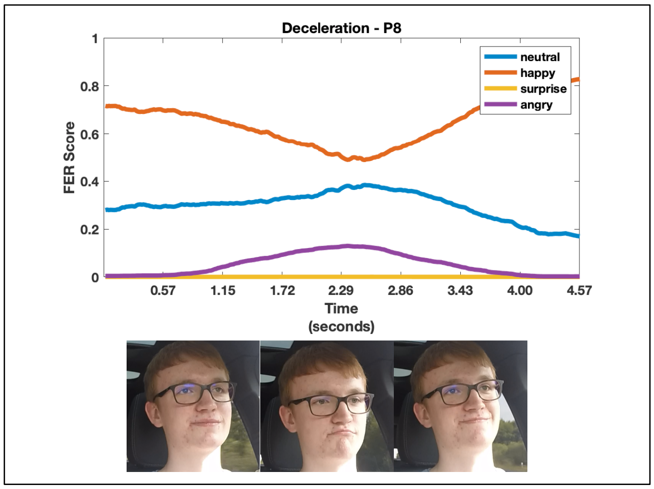


Figure 6. Deceleration with photos of corresponding frames, Participant 8. While traveling on the autobahn with ADAS active, the vehicle braked autonomously as it approached the vehicle in front. The photo on the left was just before 2.5 seconds, the middle at 2.5 seconds and the third after.

### 2) *Steering*

DISTRONIC PLUS with Steering Assist is capable of autonomously handling light curves with the supervisor of the driver. Figure 7 highlights the emotional state of the driver as they experience the vehicle steering through a curve autonomously for the first time.

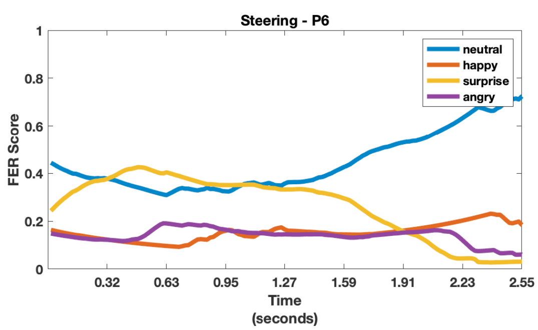


Figure 7. Steering, Participant 6. The curve of the FER score for neutral mimics the physical curve the driver was passing through on-road (as *neutral* and *happy* rises, the vehicle is coming out of the curve).

### 3) *Intervention*

Participants elected to turn off the ADAS and take back full control of the vehicle at different times or different reasons. This includes events in which the driver felt that the vehicle was not acting quickly enough, for example: DISTRONIC braking occurring to slowly, DISTRONIC acceleration occurring too quickly/slowly, uncertainty and/or distrust that the vehicle would come to a complete stop on its own (Stop & Go). Figure 8 shows the emotional state of Participant 7 as they elected to intervene and take back control.

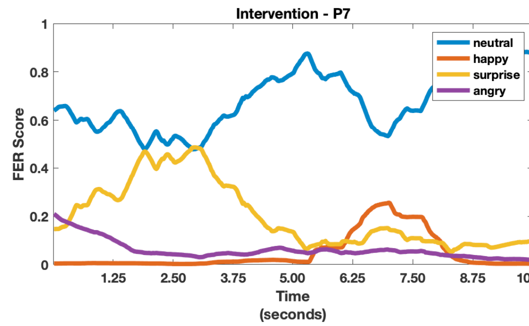


Figure 8. Intervention, Participant 7. In a country road setting, the vehicle began to slow slightly as it was approaching a vehicle in front turning off the road, however the driver intervened between 1.66-3.33 seconds. At 6.66 seconds the participant said, “I took over, it was me.” They explained that they felt vehicle was not stopping quickly enough.

### 4) *Malfunction*

During the test drive sessions, while the ADAS was active, one obvious malfunction occurred. While active on a country road, the vehicle drifted out of lane and halfway into the lane of on-coming traffic. Figure 9 shows the emotional state of Participant 4 as they silently corrected this error.

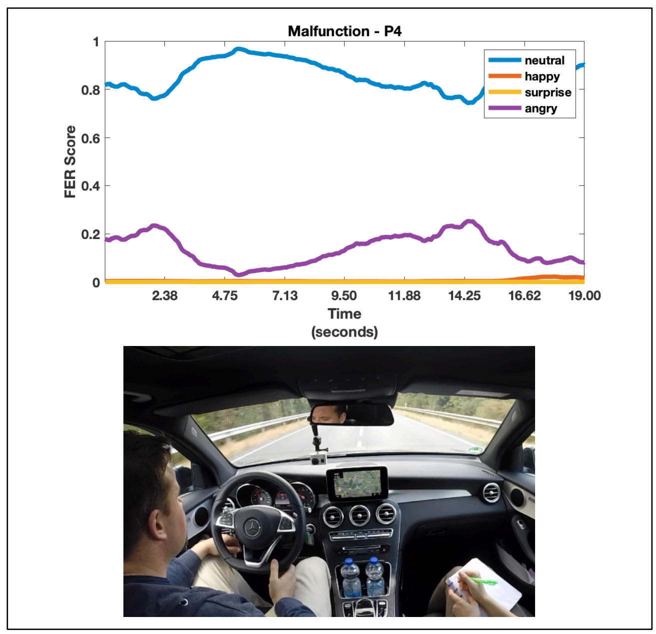


Figure 9. Malfunction with corresponding photo depicting vehicle having drifted into the lane of on-coming traffic (at 6 seconds), Participant 4. Interestingly, the driver remains neutral at the time of the event, with slight rise in anger following, continued by a very slight detection of *happy*.

### 5) *Stop&Go*

A portion of the participants elected to test the Stop & Go feature of the system, which allows the vehicle to come to a complete stop. This feature was used while in traffic in autobahn and urban settings. Figure 10 highlights the first time Participant 7 used this feature in an urban setting.

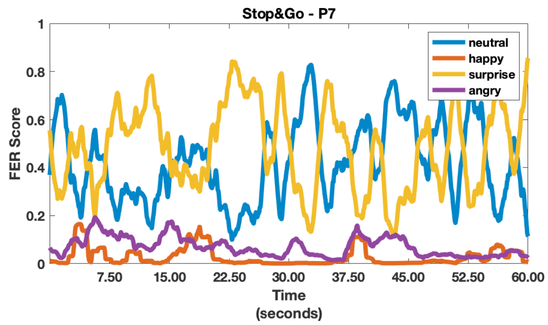


Figure 10. Stop & Go, Participant 7. While in an urban setting P7 allows the vehicle to come to a full stop. At 10 seconds, they say in anticipation, "The car is braking" and then, "I [will] try to trust the car and use the feature as it's offered. I am not doing anything." The vehicle stops at 30.83 seconds. At 35 seconds, because the car was stopped for less than three seconds (the threshold for the driver having to intervene from a full stop), the vehicle accelerated on its own to follow the car in front.

### E. *Facial Emotion Recognition + TiA Score*

The overall mean values for each of the four emotions were noted separately for each participant. Values were then converted into a percentage, indicating which emotions were most dominant throughout each drive for each participant, respectively. Figure 11 displays the relationship between participant's reported TiA scores and FER scores. The y-axis reflects the change in participants pre-drive vs. post-drive TiA scores, in order from the greatest gain to the greatest loss in TiA. The x-axis presents the FER score as a percentage, indicating the dominant emotion for each participant's drive. *Neutral* was the dominant emotion among all participants, however participants displayed differing frequencies of the emotions *happy*, *angry* and *surprise*. Participants with a gain in TiA post-drive tended to display *happy* whereas participants with a loss in TiA post-drive tended to display *angry* (see Figure 11). Note the distribution of angry among participants and its prevalence in those with a loss in TiA.

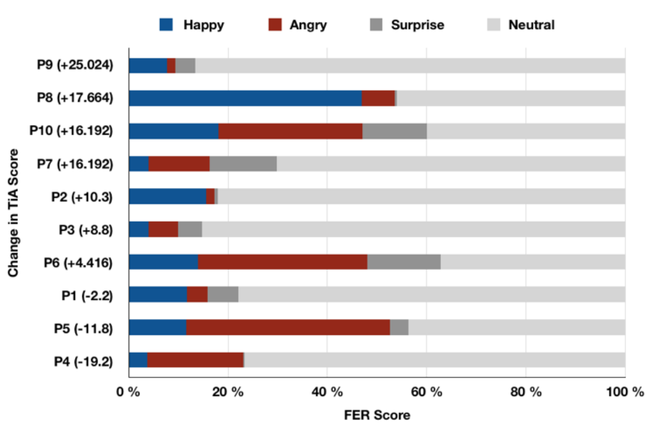


Figure 11. Participant change in TiA Score (y-axis) compared with FER scores (x-axis) as a percentage of their total test drive session.

According to their TiA scores, participants P9 and P8 reported the greatest gain in trust and (with the exception of *neutral*) displayed *happy* as a dominant emotion. Participants P4 and P5 reported the greatest loss of trust and (with the exception of *neutral*) displayed *angry* as a dominant emotion. A loss of trust was observed in both P4 and P5's verbal trust scores, which decreased at the same interval (both -1.0 post-drive), while P8 and P9 reported an increase in trust at the same intervals (both +1.25 post-introduction and post-drive).

P7 and P10 reported the same gain in TiA post-drive yet displayed different dominant emotions according to FER. A comparison of the verbal trust scores of P7 and P10 do not reveal the same changes in trust (P7 reported verbally no change in trust post-drive, while P10 reported +1.0 post-drive). Additionally, P10 displayed the dominant emotions as those drivers which had the greatest loss in trust (P4 & P5).

Relative to the other drivers in this study, P4 possessed the most aggressive driving style (e.g., speed, triggering of hands off warning graphic/tone). A contrast to P5, who was the most reluctant to give over control to the system and did not take their hands-off the wheel. However, both P4 and P5's TiA scores revealed the greatest loss of in trust in automation post-drive. This loss of trust is also reflected in their verbal trust scores, as they were the only participants to verbally report a decrease in trust post-drive. Further, P4 and P5's both displayed *angry* as a dominant emotion.

## VI. Discussion

Nesting automation in safety-critical systems requires careful consideration from a human-machine interaction perspective. In order to determine what effect a user's first contact with an ADAS has on their level of trust in the system, an on-road experiment with ten participants with no prior experience with ADAS was conducted. The use of the Trust in Automation questionnaire [20], verbal trust scores, Facial Emotion Recognition, and interviews/observational data, enabled a mixed method analysis of each participant's experience. It was hypothesized that participants would report higher levels of trust in ADAS after their first experiential drive, and that FER results would reveal a relationship between a participant's TiA score and the emotions displayed (*happy*, *angry*, *surprised*) during the drive. Further, an exploratory analysis of participant's FER scores, over the course of their test drive session and during a selection of driving events was conducted.

The scored results of the TiA questionnaire revealed that trust in automation increased after the test drive in a majority of the participants. A comparison of pre- vs. post-drive median TiA scores however, did not reveal a statistically significant difference in trust in automation ($p>0.05$).

The significant rise in factor *Reliability/Competence* ($p<0.05$, Figure 2) after the drive indicates that based on their experience driving with the ADAS, participants believe that the system performed in a way that reliably assisted them in achieving their goals [4]. Perhaps inherent trust in the established Mercedes-Benz brand had an effect on the initial scores, but it is not clear what effect it might have had on the scores post-drive. Additionally, participants made comments corresponding to the underlying factors of trust in automation, for example, P1 referred to themself as a "dinosaur" regarding their approach to technology, which may be interpreted as an indicator of their *Propensity to Trust*. Reviewing P1's median TiA score for the factor *Propensity to Trust* reveals a low score (pre-drive: 3, post-drive: 2). P9 mentioned their, "faith in the developers" during the experiment session. P9's median TiA score for the factor *Intention of the Developers* was high (pre-drive: 5, post-drive: 5).

As noted, a system malfunction occurred during the later part of P4's test drive. One can assume that this was weighed as a factor in their reported feelings of trust in the system. Additionally, it is plausible that the *Hawthorne Effect* [26] may have occurred in the instance where a participant (P1) reported an increase in trust verbally post-drive, but a decrease in trust on the questionnaire post-drive.

Regarding FER scores, participants with a gain in Trust in Automation post-drive tended to display *happy* more frequently in their FER score while those with a loss in TiA post-drive tended to display *angry* more frequently (Figure 11). This finding is of interest, as self-reported feelings of trust in automation and emotional states appear to follow a similar pattern. This gives validity to the combination of TiA, verbal trust score and FER data, suggesting that this approach may be able to identify a specific persona, who may be less trusting and therefore less accepting of ADAS. However, due to some discrepancies (see Section V.E), additional research is needed to determine if a relationship between trust in automation and emotions captured via FER can be replicated.

The driving behaviors of the participants demonstrated a willingness to take risks with the system, for example: using Stop&Go Pilot, hands-off events, and attempting to use the system in complex scenarios such as construction zones, roundabouts, and urban settings. This is aligned with the definition of trust by Mayer et al. [11]. Based on the results of this study however, displaying a willingness to take risks with the system alone is not a reliable indicator of trust as the participant who took the most risks with the system (P4) reported the greatest loss of trust post-drive.

Referring back to the LCoT framework by Ekman et al. [17], the *Learning Phase* events, *Control Transition 1, Automated Mode* and *Control Transition 2* (handover scenarios) do not list *Mental Model* as a trust-affecting factor. This is contradicted by the observations in this study. For example, during the test drive, while thinking aloud, participants stated their beliefs about how the system would behave prior to engaging in *Control Transition 1*. While in *Automated Mode* participants stated their expectations of the system's behavior. When the system behaved in a way that was not aligned with their expectations, participants engaged in *Control Transition 2*. Participants then stated why they took back control, and based on their learning from the scenario, adjusted their mental model to adapt their future interaction with the system (see Section IV.B.1). Participants who more easily developed an accurate mental model aligned with the functionality of the system handed control over to the system more easily whereas those whose mental model was not well aligned with the system had a difficult time handing over control to the vehicle. For example, P5 was reluctant to release their foot from the accelerator for an extended period of time, indicating they possessed a poor mental model of the system's functionality. After several *Transitions*, P5 made adjustments to their mental model and their interaction became more fluid. In contrast to P9, who gained an understanding of the system functionality quickly, expressed a desire to work with the system by showing "the car what the lane(s) looks like." P9's accurate model of the system (that the vehicle is tracking the lane markings) allowed for them to place more trust in the system. This suggests that during the *Learning Phase,* both *Automated Mode* and *Control Transition* events are impacted by the trust-affecting factor *Mental Model*.

This study confirms Ekman et al.'s [18] conclusion that a mixed methods approach is required to understand trust in automation. Results also suggest that the finding by Gold et al.'s [19] simulator study ("driving experience increased self-reported trust in automation") does in fact carry over to the on-road context of use.

An exploratory analysis of facial emotion recognition data captured via the driver facing camera from the on-road test drives was conducted to investigate which emotional states accompanied the participants total test drive session and assisted driving events (e.g., assisted steering). The selection of driving events and their corresponding FER plots illustrate that driver emotional states during specific driving events can be quantified by means of FER.

Analysis of FER for the whole of the participant's test drive revealed that participants who wore corrective eyeglasses (P5, P6, P8, P10) had lower FER scores. Further analysis of the raw FER data indicated that this was due to the FER algorithm having increased difficulty in accurately detecting the faces of participants who wore corrective eyeglasses. Frames with an undetected face ranged from 20% missing (P5) up to 78% missing (P6). A post-hoc review of literature on eyeglass detection for FER revealed that this issue is not uncommon [39] and CNN techniques for improving detection [40], [41]. Perhaps further training of the CNN on a broader dataset which includes more domain-specific footage (i.e. images of drivers, and drivers wearing eyeglasses) may be the best approach to improving detection for this domain.

Still, FER plots in Figure 5 are challenging to compare between participants. A variety of factors unrelated to the act of driving may have affected a participant's FER score. For example, driving events occurred in differing durations and frequencies, and participants may have been discussing a prior driving event while engaging in another driving event (e.g., discussing a time they had to intervene while their hands were off of the wheel, as participants were encouraged to think aloud). Other factors impacting FER score may have been things like sunny weather, which may have caused a driver to squint, perhaps causing them to appear *angry*, when in fact they were not. A participant's individual culture or personality may also play a role in the way they display emotions. Participants' accompaniment by a moderator in the passenger seat may have also affected the emotions displayed.

This raises an important consideration regarding affective computing in-vehicle. Observations of participants' emotional state fluctuating in response to vehicle behaviors may in the future enable the calibration of system behaviors to meet user expectations of the system's response to the on-road environment in a development setting. However, often while the ADAS was active during extended periods of autobahn driving, participants would begin a conversation on topics unrelated to the driving environment, becoming expressive while talking about past or future events. One can assume that in conditional or highly automated driving (in which drivers are able to be fully out of the loop) that users will display emotions correlated with non-driving related tasks versus the actual on-road environment or system interactions. While FER may be a useful tool for researchers & system developers, FER may not be useful for affective computing in-vehicle at the consumer level.

## VII. Limitations

Studies in simulators and on closed courses allow for significant control over research conditions, whereas studies on public roads leave much open to chance. Due to the spontaneity of the on-road environment, each participant was exposed to a variety of different scenarios at varying frequencies and intervals throughout the drive, with unplanned events such as rain or a system malfunction occurring simply by chance. Furthermore, because there is no baseline or reference point for each participant's emotional state in this study (e.g., FER scores for a drive along the same route in their daily driving vehicle), one should be cautious in their interpretation of this data. Participants also may have had a different FER scores had they been driving alone and not accompanied by a moderator.

While this study provides insight into the development of trust in ADAS on-road and using FER "in the wild", one should be cautious in generalizing the results of this study. The sample size was small (n=10), as it was limited by the loan period of the vehicle supplier. Further, the results of this study are specific to the design of Mercedes-Benz ADAS, and experienced or daily drivers of this system may display different emotions under similar conditions.

## VIII. Conclusion and Future Work

An enhanced understanding of the exchange between the user and the system on-road and the resulting effects on trust, will aid in the design of safer, more efficient automated systems. A method for correlating FER data and TiA scores is presented here which may be explored in future studies. FER accuracy in terms of both detection and classification could be improved by means of data augmentation or by training the network on a more robust dataset, should it be used again in the future. Running further test drives without a moderator in vehicle may result in different FER scores. Analysis with an emotion recognition method specifically for driver speech may also be insightful, especially with the movement towards voice user interfaces in-vehicle. The finding that all participants who reported a decrease in trust after the drive came from a similar demographic (>30 years of age, ten or more years of driving experience) warrants further investigation.

The results presented here support future research of trust in vehicle automation and other applications of FER in-vehicle. More research is needed to improve the understanding of the development of trust in automation, in order to aid the user in their acceptance of this safety-critical technology. Tackling issues of trust in ADAS today lays the groundwork for the acceptance of higher levels of autonomy in the future, eventually leading to fewer deaths, less injury, disability and a safer more enjoyable on-road experience for all.

## Acknowledgments

Thank you to Herbrand Mercedes-Benz and Mr. Sven Ingenpaß for the generous loan of the vehicle for this project. To the administration and staff of Hochschule Rhein-Waal University of Applied Sciences for their support and to the university staff who participated in this study, thank you. Special thanks to Dr. Claudio Abels for his assistance with logistics, Dr. André Frank Krause for his assistance on the CNN for FER, and Ms. Sabine Lauderbach for her knowledge of statistics.

## References

[1] L. Dixon, W. M. Megill, and K. Nebe, "Trust in Automation : An On Road Study of Trust in Advanced Driver Assistance Systems," in *VEHICULAR 2019 : The Eighth International Conference on Advances in Vehicular Systems, Technologies and Applications*, 2019, no. c, pp. 85–93.

[2] World Health Organization, "Global status report on road safety 2018," World Health Organization, 2019.

[3] C. Kerry and J. Karsten, "Gauging investment in self-driving cars," *The Brookings Institution*, 2017. [Online]. Available: https://www.brookings.edu/research/gauging-investment-in-self-driving-cars/. [Accessed: 09-Oct-2019].

[4] J. Lee and K. See, "Trust in Automation: Designing for Appropriate Reliance," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 46, no. 1, pp. 50–80, 2004.

[5] Society of Automotive Engineers (SAE) International, "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles: J3016_201806," 2018.

[6] H. Abraham, B. Seppelt, B. Mehler, and B. Reimer, "What's in a Name : Vehicle Technology Branding & Consumer Expectations for Automation," *Proc. 9th ACM Int. Conf. Automot. User Interfaces Interact. Veh. Appl. (AutomotiveUI '17)*, pp. 226–234, 2017.

[7] S. Trösterer *et al.*, "What We Can Learn from Pilots for Handovers and (De)Skilling in Semi-Autonomous Driving: An Interview Study," *Proc. 9th Int. Conf. Automot. User Interfaces Interact. Veh. Appl. (AutomotiveUI '17)*, pp. 173–182, 2017.

[8] N. Hutchins and L. Hook, "Technology Acceptance Model for Safety Critical Autonomous Transportation Systems," *2017 IEEE/AIAA 36th Digit. Avion. Syst. Conf.*, pp. 1–5, 2017.

[9] C. Rödel, S. Stadler, A. Meschtscherjakov, and M. Tscheligi, "Towards Autonomous Cars: The Effect of Autonomy Levels on Acceptance and User Experience," *Proc. 6th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.*, pp. 1–8, 2014.

[10] M. Nees, "Acceptance of Self-driving Cars: An Examination of Idealized versus Realistic Portrayals with a Self- driving Car Acceptance Scale," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 60, no. 1, pp. 1449–1453, 2016.

[11] R. Mayer, J. Davis, and D. Schoorman, "An Integrative Model of Organizational Trust," 1995.

[12] M. Körber, "Theoretical considerations and development of a questionnaire to measure trust in automation," in *20th Triennial Congress of the IEA*, 2018.

[13] A. Miramontes *et al.*, "Training Student Air Traffic Controllers to Trust Automation," *Procedia Manuf.*, vol. 3, pp. 3005–3010, 2015.

[14] F. Eyben *et al.*, "Emotion on the Road—Necessity, Acceptance, and Feasibility of Affective Computing in the Car," *Adv. Human-Computer Interact.*, pp. 1–17, Jul. 2010.

[15] J. Izquierdo-Reyes, R. A. Ramirez-Mendoza, M. R. Bustamante-Bello, J. L. Pons-Rovira, · Jose, and E. Gonzalez-Vargas, "Emotion

recognition for semi-autonomous vehicles framework," *Int. J. Interact. Des. Manuf.*, 2008.

[16] Affectiva Inc., "Affectiva Automotive AI: Metrics in Affectiva Automotive AI," 2018. [Online]. Available: https://www.affectiva.com/product/affectiva-automotive-ai/. [Accessed: 17-Sep-2018].

[17] F. Ekman, M. Johansson, and J. Sochor, "Creating Appropriate Trust in Automated Vehicle Systems: A Framwork for HMI Design," *IEEE Trans. Human-Machine Syst.*, 2017.

[18] F. Ekman and M. Johansson, "Understanding Trust in an AV-context : A Mixed Method Approach," *Proc. 6th Humanist Conf.*, no. June, pp. 13–14, 2018.

[19] C. Gold, M. Körber, C. Hohenberger, D. Lechner, and K. Bengler, "Trust in Automation – Before and After the Experience of Take-over Scenarios in a Highly Automated Vehicle," *Procedia Manuf.*, vol. 3, no. November, pp. 3025–3032, 2015.

[20] M. Körber, E. Baseler, and K. Bengler, "Introduction matters: Manipulating trust in automation and reliance in automated driving," *Appl. Ergon.*, vol. 66, no. January, pp. 18–31, 2018.

[21] J. Jian, A. Bisantz, and C. Drury, "Foundations for an Empirically Determined Scale of Trust in Automated Systems," *Int. J. Cogn. Found. an Empirically Determ. Scale Trust Autom. Syst.*, no. January 2015, pp. 37–41, 2000.

[22] B. Ko, "A Brief Review of Facial Emotion Recognition Based on Visual Information," *Sensors*, vol. 18, no. 2, p. 401, 2018.

[23] T.-K. Tews, M. Oehl, F. W. Siebert, R. Höger, and H. Faasch, "Emotional Human-Machine Interaction: Cues from Facial Expressions," Springer, Berlin, Heidelberg, 2011, pp. 641–650.

[24] R. W. Picard, "Affective Computing," Cambridge, MA, 1995.

[25] H. Hastie, X. Liu, and P. Patron, "Trust Triggers for Multimodal Command and Control Interfaces," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI'17)*, 2017, pp. 261–268.

[26] L. Bortolotti and M. Mameli, "Decpetion in Psychology: Moral Costs and Benefits of Unsought Self-Knowledge," *Account. Res.*, vol. 13, no. 3, pp. 1–20, 2006.

[27] Daimler AG, "Active safety: Intelligent Drive: Assistance in all driving situations," *Daimler Global Media Site*, 2018. [Online]. Available: https://media.daimler.com/marsMediaSite/en/instance/ko/Active-safety-Intelligent-Drive-Assistance-in-all-driving-situations.xhtml?oid=10001778. [Accessed: 14-Sep-2018].

[28] J. Brooke, "SUS - A quick and dirty usability scale," *Usability Eval. Ind.*, vol. 189, no. 194, pp. 4–7, 1996.

[29] Lund Research Ltd, "Wilcoxon Signed Rank Test in SPSS Statistics," *Laerd Statistics*, 2018. [Online]. Available: https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spss-statistics.php. [Accessed: 12-Oct-2018].

[30] M. Berenson, D. Levine, and T. Krehbiel, *Basic Business Statistics: Concepts and Applications*, New Jersey. Upper Saddle River: Prentice Hall, 2012.

[31] Statistics How To, "Friedman's Test / Two Way Analysis of Variance by Ranks," *Statistics How To*, 2014. [Online]. Available: http://www.statisticshowto.com/friedmans-test/. [Accessed: 20-Sep-2018].

[32] RStudio, "RStudio: Integrated development environment for R." Boston, MA, 2018.

[33] R Core Team, "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria, 2018.

[34] Felipe de Mendiburu, "agricolae: Statistical Procedures for Agricultural Research." R package version 1.2-8 , 2017.

[35] H. Wickham, "ggplot2: Elegant Graphics for Data Analysis." Springer-Verlag, New York, 2016.

[36] I. J. Goodfellow *et al.*, "Challenges in Representation Learning: A report on three machine learning contests."

[37] MathWorks Inc., "MATLAB." 2018.

[38] The Mathworks Inc., "Moving mean - MATLAB movmean R2018b," 2018. [Online]. Available: https://www.mathworks.com/help/matlab/ref/movmean.html. [Accessed: 06-Nov-2018].

[39] A. Bernin *et al.*, "Towards more robust automatic facial expression recognition in smart environments," *ACM Int. Conf. Proceeding Ser.*, vol. Part F128530, pp. 37–44, 2017.

[40] A. M. Basbrain, I. Al-Taie, N. Azeez, J. Q. Gan, and A. Clark, "Shallow convolutional neural network for eyeglasses detection in facial images," in *2017 9th Computer Science and Electronic Engineering Conference, CEEC 2017 - Proceedings*, 2017, pp. 157–161.

[41] A. M. Basbrain, J. Q. Gan, A. Sugimoto, and A. Clark, "A Neural Network Approach to Score Fusion for Emotion Recognition," in *2018 10th Computer Science and Electronic Engineering Conference, CEEC 2018 - Proceedings*, 2018, pp. 180–185.