Investigation of the dependency of the drivers' emotional experience on different road types and driving conditions

Marlene Weber, Joseph Giacomin, Alessio Malizia, Lee Skrypchuk, Voula Gkatzidou, Alex Mouzakitis

Abstract— The growing sophistication of technologies and sociological advances are major causes for the dramatic change the automotive sector is currently undergoing. To address changes from a human-centered design perspective an improved understanding of the occupants' emotional experience and behavior is required. Facial-Expression Analysis (FEA) is an emerging tool in support of such an approach, suitable for automotive research due to its non-contact application and low intrusiveness.

The research described here investigated the dependency of the occupants' emotional experience on road types and driving conditions by investigating emotional responses and their causes through FEA and observational analysis.

Twenty-one university students and staff were recruited for the real-time test on a planned road circuit covering different road types and conditions. Facial-expression data and video information from two in-car cameras were collected during an average driving time of 40 minutes per participant. A multi-method approach was applied for the data analysis, including both quantitative statistical analysis and qualitative observational analysis, as well as an inter-observer reliability test. Emotion frequencies were compared between the different road types, resulting in a percentage difference from the total average of emotion frequency of -6.09% below average for urban roads, +11.15% above average for major roads and +4.88% above average for rural roads.

The causes most frequently assigned to the emotional responses in this dataset were poor road conditions and causes related to the navigation device. The research supported the dependency of emotional experiences on the driving condition and type of road. The study presents the first step of a human-centered design approach towards modern automotive design. The results have wide application in automotive design, applicable to the development of, for instance, an affective human-machine interaction or a personalized autonomous driving experience.

Index Terms— Affective computing, Automotive case study, Emotion recognition, Human computer interaction

1.Introduction

Emotions play a significant role in the automotive environment. Emotional states can impact driving performance, behavior and safety. Anger can lead to aggressive driving behavior (Wells-Parker et al., 2002), stress can lead to a significant decrease in driving performance (Hoch et al., 2005; Uchiyama et al., 2002), and frustration and sadness can decrease levels of attention (Dula and Geller, 2003; Jeon, 2015; Lee, 2010). Emotional states can significantly influence goal generation, decision making, focus, attention and performance (Eyben et al., 2010).

Consequently, seeking to better understand human emotions has become a rapidly expanding research area (Noldus et al., 2017). Numerous studies have been conducted investigating emotional states, (Grimm et al., 2007; Healey, 2000; Healey and Picard, 2005;

Hoch et al., 2005; Jones and Jonsson, 2008; Lisetti and Nasoz, 2005), with a particular prevalence of aggression, workload and stress. Working to improve this understanding allows automotive design to directly respond to and address shortcomings and problem areas in current automobiles and road systems; through this, negative influencing factors can be mitigated, allowing use of the road to become a safer and more pleasant experience. Emotional factors and affective states are therefore crucial for acceptance, safety and comfort of future automotive design (Eyben et al., 2010).

As the automotive industry progresses, a host of new technologies, such as telematics, electrification, autonomous driving and other recent developments, offer many potential benefits for the future of the automotive industry (Bullis, 2011; Manyika et al., 2013). Autonomous automobiles are predicted to reduce CO_2 emission and fuel consumption (Bullis, 2001), increase safety and reduce fatalities (Manyika et al., 2013) and decrease congestion (Dumaine, 2012). Furthermore, developments like telematics and vehicle autonomy are anticipated to expand automotive revenues by 30% (Gao et al., 2016), with self-driving cars predicted to be a $87 billion opportunity by 2030 (Jacques, 2014). As these features are introduced, the emotional relationship between owner and automobile (Miller, 2001; Noldus et al., 2017), the role and significance of emotions in the wider automotive environment, and customer needs, desires and behaviors, will change (Gao et al., 2016). The automotive design process will need to adapt to the growing sophistication of in-car technologies and these changing requirements (Gao et al., 2016). To meet human requirements for coping with current and future automobile technology, it is important to understand the multi-layered emotional role of the automobile (Sheller, 2004).

One approach to responding to current and future developments is the application of affective computing, the study of systems or devices which can recognize, interpret or process human emotion (Picard, 2003) in automotive research. Numerous modern human-centered design approaches combining various methods have been applied to automotive research and design, to investigate the drivers' and passengers' behavior, emotion and needs and improve the driving experience (Giuliano, Germak and Giacomin, 2017; Gkatzidou, Giacomin and Skrypchuk, 2016).

An essential part of the study of the drivers' emotional behavior is the investigation of causes for emotions, which often include certain driving conditions or road types (Healey and Picard, 2005; Mesken, 2002). Certain emotional states have been directly linked to certain road types (e.g. rural, urban or major roads) in previous research, for instance aggressiveness (Carmona et al., 2016), frustration (Lupton, 2003) anger (Du et al., 2018) and stress (Mesken, 2002). While many automotive research studies investigated the influence of different road types on the automobile or traffic flow (DFT, 2017b; Rubino et al., 2007; Sheehan, 2017), research studies investigating road and driving conditions and their influences on the occupants are limited. Existing studies investigated accident rates on certain road types (RAC Foundation, 2009), driving behavior and speeding on different roads (Elliott, Armitage and Baughan, 2007) and risky and aggressive driving triggered by certain driving conditions (Dula and Geller, 2003). In-depth research approaches investigating the direct relationship between certain driving conditions and roads and emotional responses of occupants are scarce (Healey and Picard, 2005; Kuniecki et al., 2017; Mesken, 2002) and often restricted by their choice of measurement technique. Limitations caused by measurement techniques (e.g. sensors requiring direct contact with the participants' skin) include for instance high intrusiveness which often has an impact on the participants' behavior (Mesken, 2002). The choice of self-assessment has been criticized in previous research due to its subjectivity and influences of decaying memory strength, and fading affect bias due to the delay in the rating of emotions (Cerin, Szabo and Williams, 2001).

To avoid negative influences of the measurement tool on the participants' behavior a non-contact tool with low intrusiveness was chosen: Facial-Expression Analysis (FEA). FEA is a behavioral emotion measurement technique which requires a standard video camera. Conventional FEA approaches follow three steps for the recognition of facial expressions. The first step includes face and facial component detection. A facial image and its landmarks (e.g. corners of the eyebrows or tip of the nose) are detected and mapped from an input image through computer vision algorithms. The second step involves feature extraction, where spatial and temporal features are extracted from the facial components. In the third step expressions are classified. For this purpose machine learning algorithms, which are trained facial expression classifiers (e.g. support vector machines) are applied, producing a recognition result based on pixels analyzed in the extracted features (Ko, 2018; Lucey, et al., 2010). The classification algorithm is based on the Facial Action Coding System (FACS) (Ko, 2018). The FACS originates in Ekman's research in human facial expressions and is the most comprehensive and widely used taxonomy for the coding of facial behavior (McDuff et al., 2016).

To include a number of road types and driving conditions in the current study, a road circuit was planned based on the recommendation of existing studies (Miller, 2013; Schweitzer and Green, 2007) to include three different road types: rural, urban and major roads. An effort was made to include multiple driving conditions (e.g. high traffic density, roundabouts, poor road conditions) which may influence the emotional driving experience (Argandar, Gil and Berlanga, 2016; Cœugnet et al., 2013; Deffenbacher et al., 1994; Lee and Winston, 2016; Pau and Angius, 2001; Roidl et al., 2013).

This research combines the use of affective computing with a human-centered design approach, through investigating occupants' emotional responses during driving on different road types in different driving situations. To identify what aspects of the automotive environment are the most influential on the emotional experience, causes were assigned to the measured emotions. Facial-Expression Analysis, as a tool for the measurement of emotions was identified as suitable for the research purpose due to its low intrusiveness and non-contact application. Knowledge of the statistical frequencies and of the contextual causes would be expected to permit automotive designers to priorities a small number of road conditions and automotive systems, which may be having a disproportionate effect on the experiences and opinions of the vehicle users, for investigation.

The hypothesis of this research was therefore defined as the following:

Emotional responses during driving depend on driving conditions and road types. Differences in emotion frequencies between road types are statistically significant. An appropriate methodology for the real-time investigation of natures and frequencies of emotions during driving, and the assignment of their causes, combines both qualitative and quantitative research.

Results of this research reinforce the notion that emotions play a significant role during automobile driving and provide knowledge on causes of emotional responses on different roads in different conditions. The results of this research may be applied to the design of standardized road tests intended to investigate emotional responses during driving. Another possible application of the collected results could be an improved human-machine interaction through personification based on the individual's emotions and their causes, achieved through the avoidance of certain roads or driving situations for example.

1.1 Background research

A number of studies have investigated emotional states during driving in the past (Grimm et al., 2007; Healey, 2000; Healey and Picard, 2005; Hoch et al., 2005; Jones and Jonson, 2008; Lisetti and Nasoz, 2005;). While multiple emotion studies include different road types or driving conditions in the road circuit planning (Grimm et al., 2007; Klauer et al., 2005), results are often not analyzed from the perspective of comparing emotions between the different conditions. Approaches investigating differences in emotions on different roads are therefore limited.

One study including a comparison of emotions on different road types was conducted by Menken et al. (2007). In total 44 participants drove in an instrumented car while heart-rate measures were collected. During the test drive participants were asked to rate their emotional experiences thorough emotion scores every three minutes. When comparing heart-rate measurements on City, Ring road and Motorway roads, results showed that the three different driving conditions did not produce significantly differing results. Only small differences were noted between ring road and motorway. Self-assessed emotion scores showed that types and numbers of emotions did not differ for different driving conditions or road types. Nevertheless, the self-assessment method has been criticized in previous research due to limitations caused by the subjectivity of the measurement, difficulties in cross-cultural use and no distinct emotion measurement but measurement of general emotional states (Desmet, 2003).

Physiological data (electrocardiogram, electromyogram, skin conductance, and respiration) was recorded and combined with self-assessed data to investigate stress-levels in an on-road study with 24 participants (Healey and Picard, 2005). Highway, city-driving and rest-periods were compared. While difficulties of the application and use of the physiological sensors in the real-driving environment occurred, the self-assessed data showed that participants rated city driving as the most stressful, followed by highway driving as less stressful and the rest-period as the least stressful. Once again, the sole reliance of results on self-assessment can be criticized (Mesken, 2002).

Other research approaches investigated the relationship of workload, frustration or the driver's stress level and different road types (Miller, 2013; Schweitzer and Green, 2007; Sugiono, Widhayanuriyawan and Andriani, 2017). As workload, frustration and stress level are closely related to emotions and emotional states (Hou, Sourina and Mueller-Wittig, 2015) the research was considered relevant for the current study. Schweitzer and Green compared workload and task acceptability in urban situations, expressways, rural roads and residential roads based on ratings from video clips. Even though many exceptions were recorded, urban situations were associated with the highest workload, followed by expressways, rural roads and residential roads with the lowest workload (Schweitzer and Green, 2007). Sugiono, Widhayanuriyawan and Andriani investigated frustration and different demand and performance measures on city roads, motorways and rural roads based on subjective measurements using NASA TXL. Their results showed the highest level of frustration on city roads, followed by rural roads with the lowest frustration level on motorways (Sugiono, Widhayanuriyawan and Andriani, 2017). Miller investigated the effects of different roadways (expressways and rural roads) on driver stress using physiological measures (ECG data). The highest stress levels were measured on expressways, rural roads were notably less stressful (Miller, 2013).

In light of the scarcity and discrepancies of studies conducting in-depth investigations and comparisons of emotional states under different conditions, the research described here provides a methodology for the in-depth investigation of emotional responses during driving on different road types in different driving conditions, enabling the construction of methods and systems that will allow future research to address the highlighted issues.

2 Driving Study for observation of emotional responses on different roads

2.1 Measurement Equipment

FEA was chosen as a suitable tool for the measurement of emotions in the automotive environment due to its low intrusiveness and non-contact application (Kapoor, Qi and Picard, 2003). Furthermore FEA and has achieved up to 90% correlation with self-assessed emotions in previous research (Zeng et al., 2009).

Criteria including real-time measurement, low cost, user-friendliness easily adaptable to different participants, high portability, high robustness, customizable software and data synchronized with video feed, were defined for the choice of emotion recognition software.

Fulfilling all criteria, *Affdex Affectiva*, a real-time FEA tool, was chosen to be integrated into the data acquisition and integration platform *iMotions Attention Tool*. The *Affdex Affectiva* face detection is performed through the Viola-Jones face detection algorithm, calibrated using a large, independent set of facial images (iMotions, 2013). Taken in natural conditions with different posture and lighting, they were subsequently coded by experts (McDuff et al., 2016). The software is based on the Facial Action Coding System, which codes specific combinations of action units (contractions of facial muscles) into the six basic emotions (Ekman, Friesen and Ellsworth, 2013; McDuff et al., 2016) joy, anger surprise, fear, disgust and sadness.

Affdex Affectiva provides emotion evidence scores which correspond to the probability of the presence of each emotion in the facial image. The evidence score output from the software is between 0 (absent) and 100 (present). A threshold suggested through previous research for an emotion being present or absent of 50-70 (iMotions, 2013) is defined to determine the presence of absence of an emotion.

Limitations of the application of FEA in the automotive setting were identified in previous research (Gao, Yüce and Thiran, 2014; Tischler et al., 2007). Factors influencing the usability of the tool include lighting changes, head movement and high frequencies of expressions. In order to avoid noise and increase the usability of the chosen method in the study environment, adjustments were made. These included the creation of a threshold for the presence of an emotional response at a minimum expression duration of 1 second, adding an immediate median correction of the last 3 samples of the emotion evidence score and setting the evidence score threshold for an emotion being present at 70 (Weber, 2018).

2.2 Test Vehicle and Set-up

The research automobile was provided by Jaguar Land Rover for the duration of the study and insured by the university. The Land Rover Discovery Sport SE eD4 150PS, a four-wheel drive automobile had a 2.0L four-cylinder diesel engine and a manual transmission.

Two cameras (Logitech C920HD) were fitted in the automobile to capture the driving environment, the dashboard and the participants' face. The environment camera was fixed on the seat's headrest to capture both the dashboard and the environment of the automobile, while the face camera was fixed to the windshield (Figure 1). Both the FEA data and the recorded videos were collected on a laptop (Lenovo Thinkpad) by the researcher, seated on the backseat of the automobile.

Figure 1 Camera placement in the research automobile

Both cameras were placed such that they fulfilled the following requirements including minimal intrusiveness and impact on the participant's visual field, robust placement and avoiding camera movement through vibration or car movement. Specific requirements for the placement of the face camera included ideal location to avoid interruption of data transfer due to the participant's head movement and minimize impact on the visual field. The requirement for the scene camera was the placement to reach a wide angle covering parts of the dashboard and the driving environment to collect as much information about the driving environment and potential event triggers as possible (Figure 2)

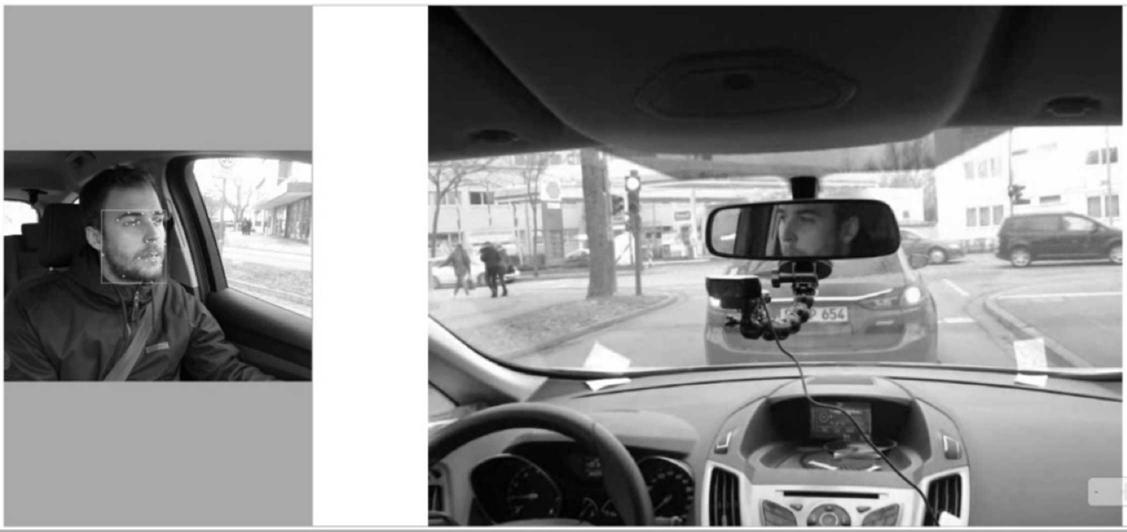


Figure 2 View of the face and scene camera during the study

2.3 Road Circuit Selection

To include a variety of road types and driving situations a road circuit was planned for the current study. Existing automotive studies (Miller, 2013; Schweitzer and Green, 2007; Schweitzer and Green, 2007; Sugiono, Widhayanuriyawan and Andriani, 2017) recommend the combination of three different road types for either the planning of road circuits or the comparison between them: rural, urban and major roads. A ratio of these three road types recommended in human factors and ergonomics research is 40% rural roads, 40% urban roads and 20% major roads (Giacomin and Bracco, 1995; Taylor, Lynam and Baruya, 2000). When planning the road circuit, the definition of road types (urban, major, rural) according to the UK Department for Transport (DFT, 2017, p.1-2) was followed (Table 1).

Table 1 Definition of road types according to the UK Department for Transport (DFT, 2017, p.1-2)

Road Type	**Definition**
Urban roads	These are major and minor roads within a settlement of population of 10,000 or more. The definition is based on the 2001 Communities and Local Government definition of Urban Settlements.
Major roads	Includes motorways and all ‘A’ roads. These roads usually have high traffic flows and are often the main arteries to major destinations.
Rural roads	These are major and minor roads outside urban areas (these urban areas have a population of more than 10,000 people).

An attempt was made to not only cover the suggested three road types but also to respect the suggested ratio in the restricted study time. Compliance with the university’s legal and ethical protocols (i.e. study length restricted to a maximum of one hour, any route point was required to be within 30 minutes of the university campus in case of emergency) suggested the choice of routes within a 30-minute radius of the university, which permitted a final configuration of (Figure 3) 4.5 miles of urban roads covering 30% of the total mileage and 17 minutes of driving on average, 6.7 miles of major roads covering 44% of the total mileage and 14 minutes of driving on average and 4.0 miles of rural roads covering 26% of the total mileage and 9 minutes of driving on average.

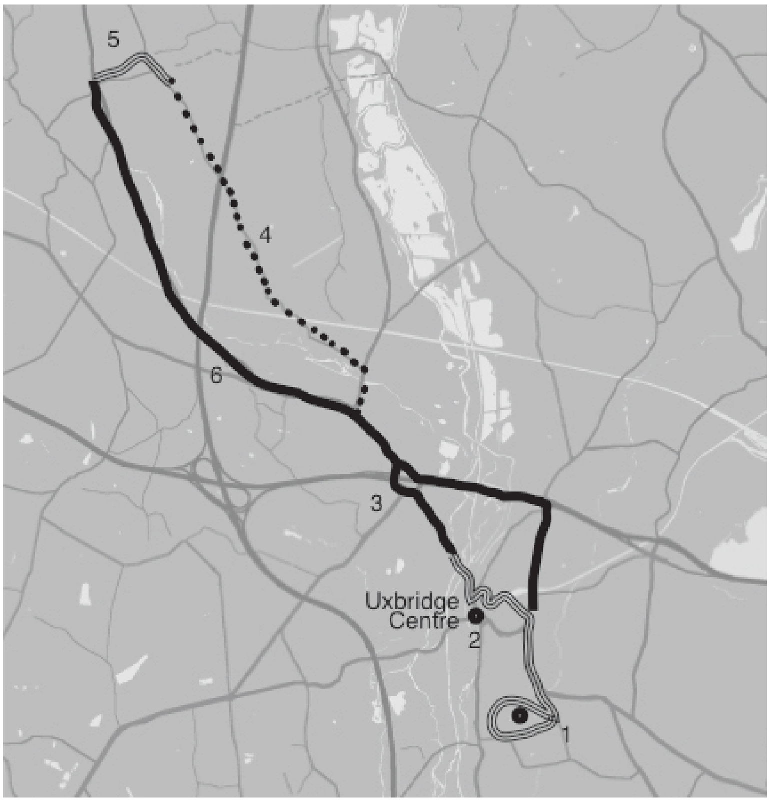


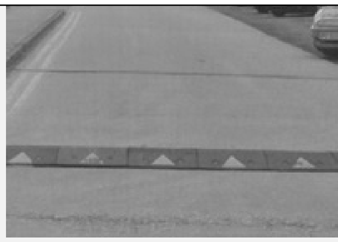
Figure 3 Map indicating road types (triple line – urban roads, line – major roads, dotted – rural roads) and road circuit numbers (see Table 2)

In order to include driving situations which may have an impact on the drivers' emotional experience (Roidl et al., 2013) literature investigating emotions during driving and their influences was reviewed (Argandar, Gil and Berlanga, 2016; Cœugnet et al., 2013; Deffenbacher et al., 1994; Lee and Winston, 2016; Pau and Angius, 2001; Roidl et al., 2013)

The number of driving and road situations, known to have an emotional impact on the driver were covered in the planned road circuit (Table 2). These include roundabouts and large challenging junctions (Funke et al., 2007; Lee and Winston, 2016; Roidl et al., 2013;), poor road conditions (e.g. potholes, eroded roads) (Argandar, Gil and Berlanga, 2016; Roidl et al., 2013), limited visual field (e.g. dense vegetation, winding road) (Roidl et al., 2013), speed bumps (Argandar, Gil and Berlanga, 2016; Pau and Angius, 2001) and bus stops and pedestrians crossing the road (Deffenbacher et al., 1994).

Table 2 Detailed explanation of the road circuit

Number (see Figure 3)	**Explanation**	**Image**

1 (Start)	A private/urban road leading over 11 speed bumps, leaving the university through 3 roundabouts. Possible impact: Stress (Argandar, Gil and Berlanga, 2016), anger (Pau and Angius, 2001)	
2	An urban road leading towards and through the town center, with high traffic density, pedestrians crossing and buses stopping. Possible impact: Stress (Argandar, Gil and Berlanga, 2016), annoyance (Cœugnet et al., 2013), anger (Mesken et al., 2007)	
3	A major road towards a large junction. Possible impact: Stress (Lee and Winston, 2016), frustration and anger (Roidl et al., 2013)	
4	A rural road with poor road conditions and a limited visual field due to dense vegetation and a winding road lay-out. Possible impact: Stress (Argandar, Gil and Berlanga, 2016), surprise (Roidl et al., 2013)	
5	An urban road with very poor road conditions and a narrow road often blocked by parked vehicles. Possible impact: Stress (Argandar, Gil and Berlanga, 2016), anger (Pau and Angius, 2001; Deffenbacher et al., 1994)	
6	Major roads leading back to university with no major challenges	

2.4 Participant Selection and Recruitment

To ensure a high quality of data the participant selection and recruitment was conducted following a purposive sampling strategy. Factors (age, gender and driver type) identified in previous research as affecting driving behavior, performance and attitude (Gwyther and Holland, 2012; Turner and McClure, 2003) were therefore controlled. To identify driver types and ensure the participation of all types, participants were asked to complete the

Multidimensional Driving Style Inventory, a standard driving style assessment tool (Taubman-Ben-Ari, Mikulincer and Gillath et al., 2004). All five driver types (angry, anxious, dissociative, distress-reduction, careful driver) were represented in the study.

To identify a suitable sampling size, research suggesting sampling sizes for qualitative, quantitative and mixed method research approaches, and literature considering validity of sampling size for data analysis, was reviewed (Creswell and Poth, 2017; Guo et al., 2013; Morse, 1994; Teddlie and Yu, 2007; VanVoorhis, Wilson and Betsy, 2007). When following a purposive sampling strategy in mixed method studies, 20-30 participants has been suggested as an appropriate sampling size (Creswell and Poth, 2017; Teddlie and Yu, 2007). For stable data analysis, sample sizes of 8-20 have been identified as sufficient (Morse, 1994).

Based on the reviewed literature 21 participants, including 10 female and 11 male drivers between the ages of 18-55 (M= 31.5, SD=11.2) were recruited for the study. They had an average 13.6 (SD= 12.2) years driving experience with an average of 10.000-15.000 miles driven per year. The selection of participants and all phases of the study were performed in accordance with the University's ethics policy.

2.5 Data Analysis Approach

The study data was analyzed following a multimethod approach.

2.5.1 Quantitative Data Analysis

Statistical analysis was performed on the collected FEA data. All facial expressions above threshold were collated for all participants and separated for the three different road types. The total average frequency (i.e. the average number of emotions registered by the FEA tool per minute) of all facial expressions was calculated. Next, the individual expressions and their frequencies for each road type were collated and the percentage differences from the total average of emotion frequency were compared. To investigate the statistical significance of the study results the frequencies of emotions a chi- squared test was performed using the road type data sets.

2.5.2 Qualitative Data Analysis

In an observational analysis during and after the study, causes (i.e. short textual description of the cause of the emotion) were assigned to the facial expressions by the researcher. All causes assigned during the study were revised afterwards, through reviewing the FEA and video data. If a cause could not be assigned during the study due to the high rate of incoming data, causes were assigned afterwards. If no obvious cause could be identified the expression was categorized as *no cause assigned* (NCA). The assigned causes were separated into the three road types.

To minimize research bias and ensure validity of the assignment of causes an inter-observer reliability test was conducted (Marques and McCall, 2005). Two independent researchers were asked to complete the same observational analysis with the purpose of cause assignment to the measured expressions for 10% of the total sample (Armstrong et al., 1997). The degree of agreement between all three researchers was then evaluated by calculating Fleiss' Kappa.

3 Results

A total of 21 participants, including 10 female and 11 male drivers in four age groups (18-25, 26-34, 36-45, 46-55) took part in the driving study. Video and emotion data was collected for each individual participant and categorized by road type. Due to durations of travel on each road type varying by participant, the frequency of emotions was considered, that is the average number of emotions registered by the FEA tool per minute. The results are summarized in Table 3, where the percentage difference from the total average was calculated from $100\,\frac{Total\ average - Road\ type\ average}{Total\ average}$.

Table 3 Frequencies of facial expressions on different road types

Road type	**Total time (minutes)**	**Total facial expressions measured**	**Average emotion frequency (emotions per minute)**	**SD**	**% difference from overall average**
URBAN	350	210	0.605	0.564	- 6.09%
MAJOR	300	229	0.777	1.140	+11.15%
RURAL	189	120	0.617	0.823	- 4.88%
Total	839	559	0.666	0.861	

In a total study time of 839 minutes, 559 emotional responses were measured, the total average frequency was calculated as 0.666 emotions per minute (SD=0.861). The comparison of the individual road frequencies to the total average showed -6.09% below average frequencies for urban roads, +11.15% above average frequencies for major roads and +4.88% above average frequencies for rural roads.

3.1 Expressions, frequencies and causes on urban roads

The tables below describe the frequencies of facial expressions as well as the most frequently assigned causes (assigned at least 5 times) for urban roads (Table 4).

Table 4 Frequencies of basic emotions on urban roads and their most frequently assigned causes

Basic emotion	**n**	**% of all basic emotions measured (total=210)**	**Causes most frequently assigned (total≥5)**
JOY	50	24	Enjoying driving the car (total=21)
			Personal interaction (total=11)
			No cause assigned (total=8)
ANGER	39	18	Navigation alert (total=8)
			Checking navigation (total=6)

			High traffic density (total=6)
SURPRISE	50	24	Navigation alert (total=8)
FEAR	6	3	
DISGUST	46	22	Navigation alert (total=6) Checking navigation (total=6)
SADNESS	19	9	

3.2 Expressions, frequencies and causes on major roads

The tables below describe the frequencies of facial expressions as well as the most frequently assigned causes (assigned at least 5 times) for major roads (Table 5).

Table 5 Frequencies of basic emotions on major roads and their most frequently assigned causes

Basic emotion	**n**	**% of all basic emotions measured (total=229)**	**Causes most frequently assigned (total≥5)**
JOY	50	22	Enjoying driving the car (total=28) Personal interaction (total=8) No cause assigned (total=6)
ANGER	46	20	Checking navigation (total=15) Navigation alert (total=7) High traffic density (total=6)
SURPRISE	44	19	Checking navigation (total=7) Poor road conditions (total=6)
FEAR	0	0	
DISGUST	71	31	High traffic density (total=20) Poor road conditions (total=12) Checking navigation (total=6)
SADNESS	18	8	

3.3 Expressions, frequencies and causes on rural roads

The tables below describe the frequencies of facial expressions as well as the most frequently assigned causes (assigned at least 5 times) for rural roads (Table 6).

Table 6 Frequencies of basic emotions on rural roads and their most frequently assigned causes

Basic emotion	**Number of emotion occurrence**	**% of all basic emotions measured (total=120)**	**Causes most frequently assigned (total≥5)**
JOY	28	23	Enjoying driving the car (total=19) Personal interaction (total=9)
ANGER	17	14	Checking navigation (total=6)
SURPRISE	35	29	Poor road conditions (total=14)

			Car passing close on narrow road (total=6)
FEAR	1	1	
DISGUST	27	23	Poor road conditions (total=10) High traffic density (total=8)
SADNESS	12	10	

3.4. Results of the Chi-Squared Test

The high standard deviations (Table 3) indicate the wide spread of emotion frequency found between participants. Consequently, the average frequency is a poor indicator of individual performance, but considering the entire data can illuminate the variations in emotion frequency between road types.

A chi-square test of independence was calculated comparing the drivers' emotions on the different road type. A p-value <0.10 was considered as a threshold for statistically significant results for this test. It is worth remarking that this significance level is slightly less strict than the conventional ones ($p < .05$ or $p < .01$). This because the goal of this analysis is to identify trends between the analysed dimensions of the three road type (Fisher 1992). A significant difference was found ($\chi2$ (10) = 16.047, $p = 0.098$), indicating that road type influences the drivers emotions. A bar-chart reported in Fig XX shows the emotion frequency for each road.

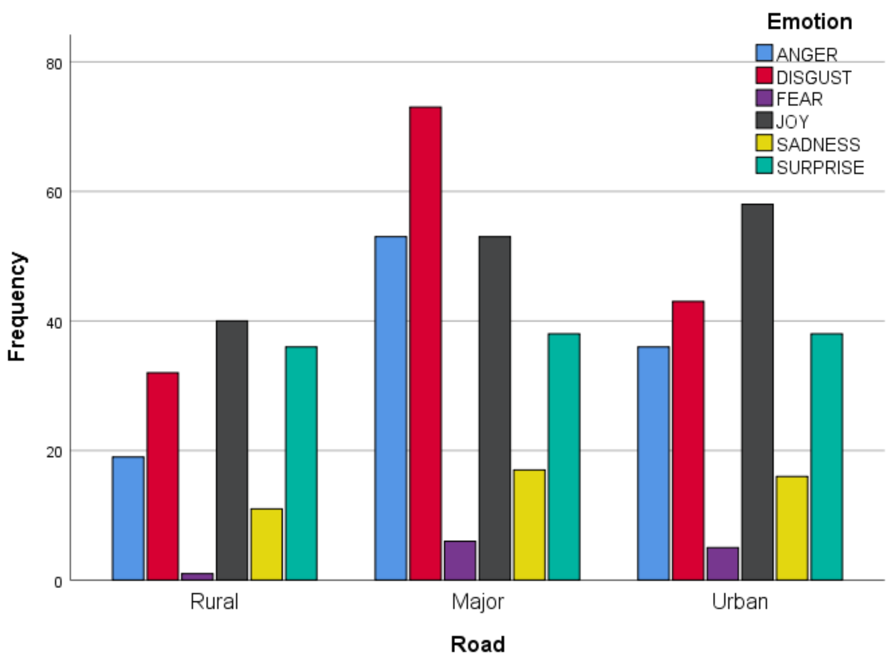


Figure 4 Bar chart indicating the road type influence on emotion frequency for road type.

3.5 Results of the inter-observer reliability test

To ensure validity of the observational analysis results and avoid research bias, an inter-observer reliability test was conducted. Two independent researcher were asked to review 10% of the study data and complete the same cause assignment exercise previously completed by the primary researcher (Armstrong et al., 1997). The degree of agreement between all three researchers was calculated using Fleiss′ Kappa, a standard measure of agreement between observers categorizing items of data and a generalization of Cohen's Kappa to multiple observers. It was calculated as κ = 0.68; this is considered to indicate "substantial" agreement not attributable to chance. As κ ranges from -1 to 1, with 0 indicating purely chance, and 1 perfect agreement, it was interpreted as substantial agreement between the observers (Xie et al., 2017).

4 Discussion

The aim of this research was to investigate the dependency of a driver's emotional experience on road types and driving conditions. A methodology for the investigation of natures, frequencies and causes of emotions during driving was introduced. Knowledge of the statistical frequencies and of the contextual causes could permit the optimization of the testing of new vehicle concepts, and could possibly lead to the redesign of test circuits for purposes of human-centered evaluations.

The research hypothesis that emotional responses depend on road types and driving conditions was supported by the statistical significance of the data collected; it was concluded that the data was indicative of a significant differences between emotion frequencies on each road type, with a low probability that these differences were due to random variations. Comparable studies showed similar results with stress-levels depending on road types and driving conditions (Healey and Picard, 2005; Mesken et al., 2007). When reviewing the planned road circuit, an explanation for the difference in frequencies may be the fact that the major roads in the road circuit included large, multi-lane roundabouts and higher traffic density while challenging situations on selected urban and rural roads were limited.

When reviewing results for the individual road types, additional differences become apparent. These additional observations produce some insight into the underlying causes of the distribution of emotions recorded during the study, however for rigorous interpretation further studies should be conducted which aim at standardizing the triggers assigned to emotion events.

The basic emotions measured most frequently for urban roads were joy and surprise (both 24% of the total), followed by disgust (22%) and anger (18%), with the lowest frequencies measured for sadness (9%) and fear (3%). The measured frequencies of basic emotions are somewhat surprising since the urban road passage included high traffic density, pedestrians crossing and buses stopping, conditions which were previously identified to trigger negative emotions (Argandar, Gil and Berlanga, 2016; Cœugnet et al., 2013; Mesken et al., 2007)

The causes most frequently assigned to joy on urban roads were *enjoying driving the car* (21 out of 48), *personal interaction* (11 out of 48) and *no cause assigned* (8 out of 48), showing a major impact of the type of car on experienced joy. Causes for anger were *navigation alert* (8 out of 36), *checking navigation* (6 out of 36) and *high traffic density* (6 out of 36). *Navigation alert* was also assigned to surprise (8 out of 48). Causes assigned to disgust included *navigation alert* (6 out of 43) and *checking navigation* (6 out of 36). It can be inferred that the type of car, as well as the use of a navigation device has a strong impact on the emotional experience on urban roads.

On major roads, disgust (31% of the total) was most frequently measured, followed by joy (22%), anger (20%) and surprise (19%), infrequent sadness (8%) and the absence of measurements of fear. These results are comparable to previous research were some of the conditions of the planned "major roads" section (e.g. challenging driving situations such as large junctions) were connected to stress and frustration (Funke et al., 2007; Lee and Winston, 2016; Roidl et al., 2013;), closely related to disgust.
The causes most frequently assigned to joy are again *enjoying driving the car* (28 out of 50), *personal interaction* (8 out of 50) and *no cause assigned* (6 out of 50). For anger the most frequent causes include *checking navigation* (15 out of 44), *navigation alert* (7 out of 44) and *high traffic density* (6 out of 44). *Checking navigation* (7 out of 42) and *poor road conditions* (6 out of 42) were assigned to surprise, while *high traffic density* (20 out of 79), *poor road conditions* (12 out of 70) and *checking navigation* (6 out of 50) were assigned to disgust. Similar to urban roads, the navigation device appeared to play an important role in the drivers' emotional experience. It is also notable that joy, the most frequently measured expression on urban roads was replaced by disgust on major roads, possibly due to higher traffic density and road conditions.

For rural roads, surprise (29% of the total of measured emotions) was the most frequently measured expression, followed by disgust and joy (both 23%), with anger and sadness measured less frequently (10-14%) and very few instances of fear (1%). The frequencies of basic emotions are comparable to results of previous research connecting surprise with winding roads and limited visual fields (Roidl et al., 2013).

The most frequently assigned causes of joy, *enjoying driving the car* (19 out of 31) and *personal interaction* (9 out of 31), are shared with urban and major roads. *Checking navigation* (6 out of 19) was most frequently assigned to anger, while *poor road conditions* (14 out of 40) and *car passing close on narrow road* (6 out of 40) were most frequently assigned to surprise. Most frequently assigned to disgust were poor road conditions (10 out of 30) and high traffic density (8 out of 31). The nature of the road (poor road conditions, narrow) seems to have a major impact on emotions experienced on rural roads. Since rural roads did not have the highest measured impact on workload, frustration and stress level in previous research (Miller, 2013; Schweitzer and Green, 2007; Sugiono, Widhayanuriyawan and Andriani, 2017) this should be further investigated in future research.

Low measured responses of fear in this dataset are surprising as fear and anxiety, closely related to fear, were reported to have major impact on driving emotion and behavior in previous research (Mesken et al., 2017; Taylor, Deane and Podd, 2000; Taylor et al., 2010). One possible explanation of the discrepancies of this study and past research could be the reliance on the Facial Action Coding System or potentially a weakness of the Affdex Affectiva emotion algorithm. Another explanation could be that the chosen driving area might not be eliciting fear in participants as they might be used to the surroundings of the university. The scare occurrence of fear should be investigated in future research.

The results display a clear indication of some of the primary causes for both negative and positive emotions on different road types. These insights can aid the development of an affective human-machine interaction through the avoidance of the causes of negative emotions and the enhancement of positive emotions.

The fact that the causes assigned to the facial expressions are often directly linked to the road type (for instance *car passing close on narrow road* as a frequent cause for emotion on a rural road) further supports the hypothesis that the emotional experience does in fact depend on the road type and driving situation. This knowledge can be used for improved, personalized navigation, which takes the driver's individual emotional experience into account when planning a route. In the future knowledge about emotional experiences on different roads could be used to tailor the route choice of self-driving vehicles such that the occupants will have the best emotional experience possible.

The knowledge that the navigation device had a major impact on the emotional experience during this study can be used for the creation of design criteria for coping with stressful driving, for example through avoiding certain road types through an alteration of the navigation route, personalized to the emotional reactions of the driver. Depending on the driver's preference and emotional responses, a more pleasurable driving experience could be created.

The study introduces an appropriate methodology for the real-time investigation of the drivers' emotions and the assignment of their causes through combining FEA and observational analysis. Results of the inter-observer reliability test ensure the validly of the assignment results. Information about the causes of emotions can assist automotive designers in detecting key issues to rectify and identifying opportunities to optimize subsystems or components. These insights could also be applied for the development of user journeys and scenario-creation, tools frequently applied in automotive research (Gkouskos, Normak and Lundgren, 2014).

5 Threats to Validity

Threats to validity in this study are listed and explained in the following.

Limited choice of road types

The choice of road types was limited by the location of the start and end point of the study route and restricted study time. This had an impact on both the road type ratio and the variance of roads (e.g. urban roads in Uxbridge Town Centre being less busy than urban roads in London city center). The ratio of road types in human factors and ergonomics research (Giacomin and Bracco, 1995; Taylor, Lynam and Baruya, 2000) was therefore not exactly met which may have influenced the variety of emotional responses on certain roads due to limited length of driving time on those. Furthermore, a different study location (busier urban roads) may have triggered different emotional responses or caused higher frequencies of emotions. To avoid influences of road type ratio and variance of road on emotional responses of participants a greater variety of roads and a larger participant sample should be considered in future research.

Researcher's presence in the car

The Hawthorne effect is an alteration of behavior when participants are aware they are under observation (Jackson and Cox, 2013; Oswald, 2014). While previous research has debated the existence and significance of the effect (Franke and Kaul, 1978; Jones 1992), all efforts were made to avoid any potential bias attributable to the presence of the observer in the car during the study. In order to achieve this, steps were taken to mitigate the effect (Jackson and Cox, 2013; Oswald, 2014): unobtrusive, naturalistic observation of the participant's behavior (researcher seated in the back and no interruption of the study); creation of a nonthreatening perception by generating a comfortable environment (giving the participant time to get used to the car, choosing a route around the participants' work or study place); application of triangulation (combination of qualitative and quantitative measurement techniques). To fully avoid any potential influences of the Hawthorne effect in future studies all data could be sent to a control room in real-time to complete the observation without the need to be present in the automobile.

Technology

The choice of emotion recognition technology and configuration may have impacted the results. For instance, the use of a single camera restricted the range of head movement that allows FEA and requires placement which impacts the participant's visual field. To achieve more reliable results multiple cameras should be used. Furthermore, the combination of different emotion measurement techniques must be considered in the future. It has been suggested, for instance, that a combination of behavioral and observational measures with physiological measures (e.g. galvanic-skin-response, heart rate measurement) will yield a superior result (Mesken et al., 2007).

Facial Action Coding System (FACS)

The use of the FACS has been criticized by numerous researchers (Essa and Pentland, 1997; Sayette et al., 2001; Wolf, 2015) for various reasons, such as the controversial opinions about FACS in science, its lack of temporal and detailed spatial information, the underlying assumption that facial expressions and emotion have an exact correspondence and the fact that its application has proven difficult to adapt for machine recognition of facial expression. While the FACS is still widely used and the most comprehensive facial-coding taxonomy (McDuff et al., 2016) the use or addition of other emotion taxonomies should be considered in future research.

Assignment of causes

A cause could not be assigned to all facial expressions (see NCA). Causes were not assigned if no obvious cause could be identified. This is a limitation which could be avoided by using more cameras to provide more information about the driving environment or by questioning the participant. Both suggestions should be considered in future research.

6 Conclusion

For this research, a mixed-method approach was applied, combining both quantitative and qualitative methods for the investigation of emotions, their natures, frequencies and causes on different road types. The results helped gain a better understanding of emotions during driving on different road types and in different driving conditions, as well as which specific causes trigger certain reactions on rural, major and urban roads. Frequencies of facial expressions were compared between the different road types and analyzed in detail for each type. Causes were examined to determine what the most significant influences on emotions are during driving on different road types. Results of this research reinforce the notion that emotions play a significant role during automobile driving and provide knowledge on causes for the emotional influences.

This study provides an appropriate methodology for the real-time investigation of emotions during driving, as well as the assignment of their causes through a combination of FEA and observational analysis. This will allow future research to improve automotive design by addressing the highlighted issues, and expand the body of knowledge addressing emotions during driving. Knowledge of the natures, frequencies and causes of emotions can assist automotive designers in identifying issues and components to analyze and modify. Results of this research may be applied to the design of standardized road tests intended to investigate emotional responses during driving. While outcomes could be used for the formulation of automotive design criteria, notice that, although very promising, some of the results should be interpreted with caution due to effect size and participants number as shown by the chi-square test in section 3.4.

Furthermore, knowledge acquired in this research could see further application in personalizing and tailoring the driving experience, allowing causes of positive emotions to be emphasized, and those of negative emotions to be prevented. This could lead to prediction of emotional responses to a given situation, and personalization of the driving experience based on the knowledge collected about the occupants' emotions during driving. The methodology presented, and the knowledge that its application can provide, may be utilized to improve both the current generation of automobiles, and to ensure the optimal integration and implementation of new technologies in the next generation of autonomous automobiles.

Bibliography:

Argandar, G.D., Gil, F.T. and Berlanga, J.F., 2016. Measuring situations that stress Mexicans while driving. *Transportation research part F: traffic psychology and behaviour*, *37*, pp.154-161.

Armstrong, D., Gosling, A., Weinman, J. and Marteau, T., 1997. The place of inter-rater reliability in qualitative research: an empirical study. Sociology, 31(3), pp.597-606.

Bullis, K., 2011. How vehicle automation will cut fuel consumption. MIT's Technology Review. October, 24.

Butler, E.A. and Strayer, J., 1998. The many faces of empathy. In *Poster presented at the annual meeting of the Canadian Psychological Association, Edmonton, Alberta, Canada*.

Carmona, J., García, F., de Miguel, M.Á., de la Escalera, A. and Armingol, J.M., 2016. Analysis of Aggressive Driver Behaviour using Data Fusion. In *VEHITS* (pp. 85-90).

Cerin, E., Szabo, A. and Williams, C., 2001. Is the experience sampling method (ESM) appropriate for studying pre-competitive emotions?. *Psychology of Sport and Exercise*, *2*(1), pp.27-45.

Cienki, A. and Mittelberg, I., 2013. Creativity in the forms and functions of spontaneous gestures with speech. The Agile Mind: A Multidisciplinary Study of a Multifaceted Phenomenon. Berlin, Germany: De Gruyter Mouton, pp.231-252.

Cœugnet, S., Naveteur, J., Antoine, P. and Anceaux, F., 2013. Time pressure and driving: Work, emotions and risks. *Transportation research part F: traffic psychology and behaviour*, *20*, pp.39-51.

Creswell, J.W. and Poth, C.N., 2017. Qualitative inquiry and research design: Choosing among five approaches. Sage publications.

Deffenbacher, J.L., Oetting, E.R. and Lynch, R.S., 1994. Development of a driving anger scale. *Psychological reports*, *74*(1), pp.83-91.

Desmet, P., 2003. Measuring emotion: Development and application of an instrument to measure emotional responses to products. In *Funology* (pp. 111-123). Springer Netherlands.

DFT (Department of Transport), 2017a. Road length notes definitions. Available at: http://www.englandhighways.co.uk/wp-content/uploads/2017/03/road-length-notes-definitions.pdf (Accessed: 26 June 2017)

DFT (Department of Transport) 2017b: Road traffic estimates: Great Britain 2016. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/611304/annual-road-traffic-estimates-2016.pdf (Accessed: 28 June 2017)

Du, X., Shen, Y., Chang, R. and Ma, J., 2018. The exceptionists of Chinese roads: The effect of road situations and ethical positions on driver aggression. *Transportation Research Part F: Traffic Psychology and Behaviour*, *58*, pp.719-729.

Duenwald, M., 2005. The physiology of facial expressions. Retrieved September, 19, p.2007.

Dula, C.S. and Geller, E.S., 2003. Risky, aggressive, or emotional driving: Addressing the need for consistent communication in research. Journal of safety research, 34(5), pp.559-566.

Dumaine, B., 2012. The driverless revolution rolls on. Availabe at http://fortune.com/2012/11/12/the-driverless-revolution-rolls-on/ (Accessed: 3 September 2017)

Ekman, P., Friesen, W.V. and Ellsworth, P., 2013. Emotion in the human face: Guidelines for research and an integration of findings. Elsevier.

Elliott, E.A. and Jacobs, A.M., 2013. Facial expressions, emotions, and sign languages. Frontiers in psychology, 4.

Elliott, M.A., Armitage, C.J. and Baughan, C.J., 2007. Using the theory of planned behaviour to predict observed driving behaviour. *British Journal of Social Psychology*, *46*(1), pp.69-90.

Escanés, G. and Poó, F.M., 2018. Driving anger in Argentina. *Safety science*, *105*, pp.228-237.

Essa, I.A. and Pentland, A.P., 1997. Coding, analysis, interpretation, and recognition of facial expressions. IEEE transactions on pattern analysis and machine intelligence, 19(7), pp.757-763.

Eyben, F., Wöllmer, M., Poitschke, T., Schuller, B., Blaschke, C., Färber, B. and Nguyen-Thien, N., 2010. Emotion on the road—necessity, acceptance, and feasibility of affective computing in the car. Advances in human-computer interaction, 2010.

Fisher, Ronald Aylmer. 1992. "Statistical methods for research workers." In Breakthroughs in statistics, 66-70. Springer.

Franke, R.H. and Kaul, J.D., 1978. The Hawthorne experiments: First statistical interpretation. American sociological review, pp.623-643.

Funke, G., Matthews, G., Warm, J.S. and Emo, A.K., 2007. Vehicle automation: A remedy for driver stress?. *Ergonomics*, *50*(8), pp.1302-1323.

Gao, H., Yüce, A. and Thiran, J.P., 2014, October. Detecting emotional stress from facial expressions for driving safety. In Image Processing (ICIP), 2014 IEEE International Conference on(pp. 5961-5965). IEEE.

Gao, P.,Kaas, H., Mohr, D., Wee, D., 2016. Automotive revolution: perspective towards 2030: how the convergence of disruptive technology-driven trends could transform the auto industry. Available at: http://www.mckinsey.com/industries/high-tech/our-insights/disruptive-trends-that-will-transform-the-auto-industry (Accessed: 05 January 2017)

Giacomin, J. and Bracco, R., 1995. An experimental approach for the vibration optimisation of automotive seats. ATA Third International, 7.

Giuliano, L., Germak, C. and Giacomin, J., 2017. Effect of Driving Context On Design Dialogue.

Gkatzidou, V., Giacomin, J. and Skrypchuk, L., 2016. Automotive Habitat Laboratory: a facility for automotive co-design.Proceedings of the 7th International Conference on Applied Human Factors and Ergonomics, Orlando, Florida, USA. July 27-31.

Gkouskos, D., Normark, C.J. and Lundgren, S., 2014. What drivers really want: Investigating dimensions in automobile user needs. *International Journal of Design*, *8*(1).

Grimm, M., Kroschel, K., Harris, H., Nass, C., Schuller, B., Rigoll, G. and Moosmayr, T., 2007. On the necessity and feasibility of detecting a driver's emotional state while driving. Affective computing and intelligent interaction, pp.126-138.

Guo, Y., Logan, H.L., Glueck, D.H. and Muller, K.E., 2013. Selecting a sample size for studies with repeated measures. BMC medical research methodology, 13(1), p.100.
Gwyther, H. and Holland, C., 2012. The effect of age, gender and attitudes on self-regulation in driving. Accident Analysis & Prevention, 45, pp.19-28.

Healey, J.A. and Picard, R.W., 2005. Detecting stress during real-world driving tasks using physiological sensors. IEEE Transactions on intelligent transportation systems, 6(2), pp.156-166.

Healey, J.A., 2000. Wearable and automotive systems for affect recognition from physiology (Doctoral dissertation, Massachusetts Institute of Technology).

Hoch, S., Althoff, F., McGlaun, G. and Rigoll, G., 2005, March. Bimodal fusion of emotional data in an automotive environment. In Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). IEEE International Conference on (Vol. 2, pp. ii-1085). IEEE.

Hou, X., Liu, Y., Sourina, O. and Mueller-Wittig, W., 2015, October. CogniMeter: EEG-based emotion, mental workload and stress visual monitoring. In *Cyberworlds (CW), 2015 International Conference on* (pp. 153-160). IEEE.

iMotions, 2013. Attention Tool Guide. Available at: http://imotionsglobal.com/wpcontent/uploads/2013/08/Guide.pdf (Accessed 25 September 2015)

Jackson, M. and Cox, D.R., 2013. The principles of experimental design and their application in sociology. Annual Review of Sociology, 39, pp.27-49.

Jacques, C., 2014. Self-driving Cars an $87 Billion Opportunity in 2030, Though None Reach Full Autonomy. Lux Research. Available at: http://www.luxresearchinc.com/news-and-events/press-releases/read/self-driving-cars-87-billion-opportunity-2030-though-none-reach (Accessed 31 October 2017)

Jeon, M. and Walker, B.N., 2011, September. What to detect? Analyzing factor structures of affect in driving contexts for an emotion detection and regulation system. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*(Vol. 55, No. 1, pp. 1889-1893). Sage CA: Los Angeles, CA: Sage Publications.

Jeon, M., 2015. Towards affect-integrated driving behaviour research. Theoretical Issues in Ergonomics Science, 16(6), pp.553-585.

Jeon, M., Walker, B.N. and Yim, J.B., 2014. Effects of specific emotions on subjective judgment, driving performance, and perceived workload. Transportation research part F: traffic psychology and behaviour, 24, pp.197-209.

Jones, C. and Jonsson, I.M., 2008. Using paralinguistic cues in speech to recognise emotions in older car drivers. Affect and Emotion in Human-Computer Interaction, 4868, pp.229-240.

Jones, S.R., 1992. Was there a Hawthorne effect?. American Journal of sociology, 98(3), pp.451-468.

Kapoor, A., Qi, Y. and Picard, R.W., 2003, October. Fully automatic upper facial action recognition. In Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on (pp. 195-202). IEEE.

Klauer, S.G., Neale, V.L., Dingus, T.A., Ramsey, D. and Sudweeks, J., 2005, September. Driver inattention: A contributing factor to crashes and near-crashes. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 49, No. 22, pp. 1922-1926). Sage CA: Los Angeles, CA: SAGE Publications.

Ko, B.C., 2018. A Brief Review of Facial Emotion Recognition Based on Visual Information. *sensors*, *18*(2), p.401.

Kuniecki, M., Wołoszyn, K.B., Domagalik, A. and Pilarczyk, J., 2017. Effects of Scene Properties and Emotional Valence on Brain Activations: A Fixation-Related fMRI Study. *Frontiers in human neuroscience*, *11*, p.429.

Lee, Y.C. and Winston, F.K., 2016. Stress induction techniques in a driving simulator and reactions from newly licensed drivers. *Transportation research part F: traffic psychology and behaviour*, *42*, pp.44-55.

Lee, Y.C., 2010, September. Measuring drivers' frustration in a driving simulator. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 54, No. 19, pp. 1531-1535). Sage CA: Los Angeles, CA: Sage Publications.

Lisetti, C.L. and Nasoz, F., 2005, July. Affective intelligent car interfaces with emotion recognition. In Proceedings of 11th International Conference on Human Computer Interaction, Las Vegas, NV, USA.

Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z. and Matthews, I., 2010, June. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on* (pp. 94-101). IEEE.

Lupton, D., 2002. Road rage: drivers' understandings and experiences. *Journal of Sociology*, *38*(3), pp.275-290.

Manyika, J., Chui, M., Bughin, J., Dobbs, R., Bisson, P. and Marrs, A., 2013. Disruptive technologies: Advances that will transform life, business, and the global economy (Vol. 180). San Francisco, CA: McKinsey Global Institute.

Marques, J.F. and McCall, C., 2005. The application of interrater reliability as a solidification instrument in a phenomenological study. The Qualitative Report, 10(3), pp.439-462.

McDuff, D., Mahmoud, A., Mavadati, M., Amr, M., Turcot, J. and Kaliouby, R.E., 2016, May. AFFDEX SDK: a cross-platform real-time multi-face expression recognition toolkit. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (pp. 3723-3726). ACM.

Mesken, J., 2002. *Measuring emotions in traffic* (No. D-2002-3). Leidschendam: SWOV Institute for Road Safety Research.

Mesken, J., Hagenzieker, M.P., Rothengatter, T. and de Waard, D., 2007. Frequency, determinants, and consequences of different drivers' emotions: An on-the-road study using self-reports, (observed) behaviour, and physiology. Transportation research part F: traffic psychology and behaviour, 10(6), pp.458-475.

Miller, D., 2001. Driven societies. in D. Miller (ed.) Automobile Cultures. Oxford: Berg.

Miller, E.E., 2013. Effects of Roadway on Driver Stress: An On-Road Study using Physiological Measures (Doctoral dissertation).

Morse, J.M. ed., 1994. Critical issues in qualitative research methods. Sage, pp. 281–297.

Namba, S., Kabir, R.S., Miyatani, M. and Nakao, T., 2017. Spontaneous Facial Actions Map onto Emotional Experiences in a Non-social Context: Toward a Component-Based Approach. Frontiers in Psychology, 8.

Noldus, L.P., Spink, A.J., Bollen, R. and Heffelaar, T., 2017. Smart Mobility: Driver State Estimation and Advanced Driver-Vehicle Interfaces. In *Mobility Engineering* (pp. 11-18). Springer, Singapore.

Oswald, D., Sherratt, F. and Smith, S., 2014. Handling the Hawthorne effect: The challenges surrounding a participant observer. *Review of social studies*, *1*(1), pp.53-73.

Pau, M. and Angius, S., 2001. Do speed bumps really decrease traffic speed? An Italian experience. *Accident Analysis & Prevention*, *33*(5), pp.585-597.

Picard, R.W., 2003. Affective computing: challenges. International Journal of Human-Computer Studies, 59(1), pp.55-64.

RAC Foundation, 2009. Accident Trends by Road Type. Available at https://www.racfoundation.org/assets/rac_foundation/content/downloadables/roads%20and%20reality%20-%20bayliss%20-%20accident%20trends%20by%20road%20type%20-%20160309%20-%20background%20paper%209.pdf (accessed 24 March 2018)

Roidl, E., Frehse, B., Oehl, M. and Höger, R., 2013. The emotional spectrum in traffic situations: Results of two online-studies. *Transportation research part F: traffic psychology and behaviour*, *18*, pp.168-188.

Roidl, E., Siebert, F.W., Oehl, M. and Höger, R., 2013. Introducing a multivariate model for predicting driving performance: The role of driving anger and personal characteristics. *Journal of safety research*, *47*, pp.47-56.

Rubino, L., Bonnel, P., Hummel, R., Krasenbrink, A. and Manfredi, U., 2007. Mobile measurement of pollutant emissions and fuel consumption of road vehicles in real-world driving situations using portable emission measurement systems (PEMS). *Final report. Eur. Commission, Ispra.*

Russell, J.A. and Fernández-Dols, J.M. eds., 1997. The psychology of facial expression. Cambridge university press.

Sam Sheehan, 2017. New UK real-world emissions tests start today. Available at: https://www.autocar.co.uk/car-news/industry/new-uk-real-world-emissions-tests-start-today (Accessed: 26 June 2017)

Sayette, M.A., Cohn, J.F., Wertz, J.M., Perrott, M.A. and Parrott, D.J., 2001. A psychometric evaluation of the facial action coding system for assessing spontaneous expression. Journal of Nonverbal Behavior, 25(3), pp.167-185.

Schweitzer, J. and Green, P.E., 2007. Task acceptability and workload of driving city streets, rural roads, and expressways: Ratings from video clips.

Sheller, M., 2004. Automotive emotions: feeling the car. Theory, culture & society, 21(4-5), pp.221-242.

Sugiono, S., Widhayanuriyawan, D. and Andriani, D.P., 2017. Investigating the Impact of Road Condition Complexity on Driving Workload Based on Subjective Measurement using NASA TLX. In *MATEC Web of Conferences* (Vol. 136, p. 02007). EDP Sciences.

Taubman-Ben-Ari, O., Mikulincer, M. and Gillath, O., 2004. The multidimensional driving style inventory—scale construct and validation. Accident Analysis & Prevention, 36(3), pp.323-332.

Taylor, J.E., Alpass, F., Stephens, C. and Towers, A., 2010. Driving anxiety and fear in young older adults in New Zealand. *Age and ageing*, *40*(1), pp.62-66.

Taylor, M.C., Lynam, D.A. and Baruya, A., 2000. *The effects of drivers' speed on the frequency of road accidents*. Crowthorne: Transport Research Laboratory.

Taylor, J.E., Deane, F.P. and Podd, J.V., 1999. Stability of driving fear acquisition pathways over one year. *Behaviour Research and Therapy*, *37*(10), pp.927-939.

Teddlie, C. and Yu, F., 2007. Mixed methods sampling: A typology with examples. Journal of mixed methods research, 1(1), pp.77-100.

Tischler, M.A., Peter, C., Wimmer, M. and Voskamp, J., 2007, September. Application of emotion recognition methods in automotive research. In Proceedings of the 2nd Workshop on Emotion and Computing—Current Research and Future Impact(Vol. 1, pp. 55-60).

Turner, C. and McClure, R., 2003. Age and gender differences in risk-taking behaviour as an explanation for high incidence of motor vehicle crashes as a driver in young males. Injury control and safety promotion, 10(3), pp.123-130.

Uchiyama, Y., Kojima, S.I., Hongo, T., Terashima, R. and Wakita, T., 2002, September. Voice information system adapted to driver's mental workload. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 46, No. 22, pp. 1871-1875). Sage CA: Los Angeles, CA: SAGE Publications.

VanVoorhis, CR Wilson, and Betsy L. Morgan. "Understanding power and rules of thumb for determining sample sizes." Tutorials in Quantitative Methods for Psychology 3, no. 2 (2007): 43-50.

Weber, M., 2018. Automotive emotions: a human-centred approach towards the measurement and understanding of drivers' emotions and their triggers (Doctoral dissertation, Brunel University London).

Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J. and Kissler, J., 2017. Mapping the emotional face. How individual face parts contribute to successful emotion recognition. PloS one, 12(5), p.e0177239.

Wells-Parker, E., Ceminsky, J., Hallberg, V., Snow, R.W., Dunaway, G., Guiling, S., Williams, M. and Anderson, B., 2002. An exploratory study of the relationship between road rage and crash experience in a representative sample of US drivers. Accident Analysis & Prevention, 34(3), pp.271-278.

Wolf, K., 2015. Measuring facial expression of emotion. Dialogues in clinical neuroscience, 17(4), p.457.

Xie, Z., Gadepalli, C., Jalalinajafabadi, F., Cheetham, B.M. and Homer, J.J., 2017, October. Measurement of rater consistency and its application in voice quality assessments. In *Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2017 10th International Congress on*(pp. 1-6). IEEE.

Zeng, Z., Pantic, M., Roisman, G.I. and Huang, T.S., 2009. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. IEEE transactions on pattern analysis and machine intelligence, 31(1), pp.39-58.