Affective Driver-Pedestrian Interaction: Exploring Driver Affective Responses toward Pedestrian Crossing Actions using Camera and Physiological Sensors

Shruti Rao
University of Amsterdam
Amsterdam, The Netherlands
shruti.rao@cwi.nl

Sabrina Wirjopawiro
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
s.wirjopawiro@gmail.com

Gerard Pons Rodriguez
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
gerardponsr@gmail.com

Thomas Röggla
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
t.roggla@cwi.nl

Pablo Cesar
Centrum Wiskunde & Informatica
Delft University of Technology
Amsterdam, The Netherlands
p.s.cesar@cwi.nl

Abdallah El Ali
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
aea@cwi.nl

ABSTRACT

Eliciting and capturing drivers' affective responses in a realistic outdoor setting with pedestrians poses a challenge when designing in-vehicle, *empathic* interfaces. To address this, we designed a controlled, outdoor car driving circuit where drivers (*N=27*) drove and encountered pedestrian confederates who performed non-verbal positive or non-positive road crossing actions towards them. Our findings reveal that drivers reported higher valence upon observing positive, non-verbal crossing actions, and higher arousal upon observing non-positive crossing actions. Drivers' heart signals (BVP, IBI and BPM), skin conductance and facial expressions (brow lowering, eyelid tightening, nose wrinkling, and lip stretching) all varied significantly when observing positive and non-positive actions. Our car driving study, by drawing on realistic driving conditions, further contributes to the development of in-vehicle empathic interfaces that leverage behavioural and physiological sensing. Through automatic inference of driver affect resulting from pedestrian actions, our work can enable novel empathic interfaces for supporting driver emotion self-regulation.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Empirical studies in HCI**.

KEYWORDS

empathic cars, driver emotion recognition, pedestrian non-verbal crossing actions, outdoor driving circuit, physiological sensing, thermal sensing, facial expression analysis

AutomotiveUI '23, September 18–22, 2023, Ingolstadt, Germany

ACM ISBN 979-8-4007-0105-4/23/09.
https://doi.org/10.1145/3580585.3607168

ACM Reference Format:
Shruti Rao, Sabrina Wirjopawiro, Gerard Pons Rodriguez, Thomas Röggla, Pablo Cesar, and Abdallah El Ali. 2023. Affective Driver-Pedestrian Interaction: Exploring Driver Affective Responses toward Pedestrian Crossing Actions using Camera and Physiological Sensors. In *15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '23), September 18–22, 2023, Ingolstadt, Germany*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3580585.3607168

1 INTRODUCTION

Driver emotions, such as anger or stress, or even ecstatic joy, that arise during driving scenarios can significantly impact driving behaviour [24]. Consequently, *Empathic* interfaces [1,2] are being developed to identify and regulate driver emotions for improving road safety [27, 60, 61]. Particularly, these interfaces may help drivers regulate their emotions by providing calming cues (for e.g. music or colours [21]) when signs of stress or anxiety in drivers are detected. Empathic interfaces can also be used to enhance non-verbal communication during critical moments like merging or yielding by communicating drivers' emotional state to other road users through visual interfaces. [19] These driver emotions are detected through collection of physiological data, such as heart rate or skin conductance, facial expressions, vocal tone, and eye movements that in turn inform empathic interfaces.

While environmental and situational factors have been considered in previous research for inferring drivers' emotional states [6, 19, 25], non-verbal interactions between drivers and pedestrians have received less attention [16, 45]. Pedestrian non-verbal behaviour can often elicit different emotional responses from drivers, where inferring driver affective states during driving scenarios can support designing empathic in-vehicle interfaces for improving driving experience and road safety [3, 60].

To capture the impact of non-verbal, pedestrian crossing actions on drivers' affective states in outdoor driving conditions, we designed a driving track and equipped a car with a combination of sensors to capture drivers' affective states (emotion self-reports,

[1] https://www.theguardian.com/business/2018/jan/23/a-car-which-detects-emotions-how-driving-one-made-us-feel
[2] https://www.irishtimes.com/business/transport-and-tourism/researchers-developing-empathic-car-technology-1.3900701

physiological signals, facial data) in response to positive and non-positive, non-verbal pedestrian crossing actions. In this work, we ask **RQ**: How do drivers' affective responses vary toward non-verbal, pedestrian crossing actions in a controlled, outdoor environment? We ran an approximately week-long controlled, outdoor driving study ($N = 27$) with drivers to test this. In this setup, drivers encountered a zebra crossing where pedestrians performed non-verbal crossing actions. We investigated the influence of different pedestrian actions on drivers' affective states using a combination of camera and physiological sensors. We recorded drivers' responses in the form of emotion self-reports (valence and arousal, based on Russell's Circumplex model of emotion [49]). Drivers' physiological signals (heart rate and skin conductance) were gathered using an Empatica E4 wristband, and their facial expressions were collected using a FLIR dual RGB and thermal camera.

The momentary nature of actions such as eye contact or a nod by several pedestrians at a given point in time makes it challenging to capture and understand the resulting driver's emotions [14, 48, 62]. Existing works rely on video stimuli depicting pedestrian crossing actions to induce driver affect which were measured using self-reports [16], or in-lab using sensors on participants with prior driving experience [45]. As an alternative, hybrid driving simulators have also been designed that combine driving in a simulated environment while interacting with real-world pedestrian road crossing actions [46]. Both approaches however lack driving context and realistic driver-pedestrian interactions, which are necessary to understand a driver's behavioural and affective state during such interactions. Contrary to such approaches, we incorporated a greater degree of realism and ran a controlled, outdoor study with confederate pedestrians. We collected drivers' signals throughout the entire duration of the study and manually annotated time segments where pedestrian crossing actions occurred. To identify the exact source of driver affect, our study also ensured that a single pedestrian crossed and performed an action at a time.

Our findings show that our outdoor setup can effectively capture drivers' affective states from observing non-verbal, pedestrian crossing actions. Specifically, we observe that drivers reported higher valence (pleasantness) upon observing positive pedestrian crossing actions, and higher arousal (excitement) upon observing non-positive pedestrian crossing actions. Additionally, drivers' physiological signals corresponding to heart rate (BVP, IBI and BPM) were significantly influenced by the different (positive versus non-positive) non-verbal, pedestrian crossing actions. Heart and skin signals (BVP and GSR) also varied significantly for different levels of drivers' valence (positive versus non-positive) and arousal (high versus non-high) scores. Finally, drivers' facial expressions also varied significantly upon observing both positive and non-positive non-verbal, pedestrian crossing actions.

Our work offers two key contributions: **(1)** We introduce a novel controlled, outdoor car driving setup that leverages a combination of camera and physiological sensors for capturing drivers' affective responses as a result of pedestrian non-verbal crossing actions. **(2)** Empirical findings reveal that non-verbal, pedestrian actions can influence drivers' self-reported emotions (valence and arousal), heart signals and facial expressions. Automotive safety research emphasises identifying and regulating drivers' emotions, particularly high arousal levels, which are associated with risky driving behaviour [8, 50]. Quantitative findings from our study indicate high driver arousal levels in response to the non-positive crossing actions enacted by confederate pedestrians. These actions can serve as cues for in-car affect recognition systems to preemptively detect risky driving behaviour resulting from driver-pedestrian interactions. Empathic in-car interfaces can therefore infer drivers' affective cues identified in our study including physiological responses (heart rate and skin conductance), facial expressions, and self-reported emotions, to automatically infer drivers' affective states during driver-pedestrian interactions. This can be integrated into an emotion self-regulation framework aimed at enhancing road safety [4, 29]

2 RELATED WORK

This section outlines the three main strands of research that are relevant for our work in inferring drivers' emotions arising from different pedestrian actions: (a) on-road driver-pedestrian interactions, (b) emotion models for driver affect recognition, and (c) in-car driver affect recognition.

2.1 On-Road Driver-Pedestrian Interactions

Non-verbal communication greatly influences driving behaviour, such as driver-pedestrian communication through body language [19, 55]. Previous studies found eye contact and facial expressions to be primary means of driver-pedestrian communication [18, 57], with pedestrians' body language serving as important cues of intent. Actions such as hand waving, thumbs-up, and head nodding were commonly used by pedestrians to express positive reactions such as gratitude or acknowledgement [30, 53]. Further, pedestrians in uncontrolled crosswalks also performed hand waves, head nods, and combined gestures involving pointing or indicating the desired direction of travel [64]. On the other hand, driver communication towards pedestrians is primarily via vehicle movement such as speeding up or slowing down or honking [41, 51]. Additional factors such as time taken by pedestrian to arrive at the crossing also played a role in driver attitudes such as giving way [17]. Despite existing research, the impact of non-verbal pedestrian actions at road crossings on drivers' emotional states is still relatively unexplored. This study contributes to a better understanding of the role of different non-verbal pedestrian road crossing actions in influencing drivers' emotions.

2.2 Understanding Emotions - Emotion Measurements and Models

Affect determination in the automotive context draws on research from affective computing [43]. These include two main models of emotions: discrete emotion models (e.g., Ekman's six basic emotions [12] and Plutchik's emotion wheel [44]) and continuous emotion models (e.g., the Circumplex [49] and Pleasure-Arousal-Dominance models [38]). Studies have identified anxiety, anger, and happiness as common discrete emotions experienced by drivers [26, 39]. This study uses the 9-point, discrete Self-Assessment Manikin (SAM) scale to obtain emotion self-reports from driver participants, due to its widespread use and ease of use in emotion-measurement studies.

2.3 In-Car Driver Affect Recognition

Existing literature explores facial, bio-physiological, and driving behaviours for emotion recognition. Facial analysis extracts Regions of Interest (ROIs) and facial action units (AU) from RGB images to identify facial expressions and associated emotions [13, 33]. Bio-physiological signals include electrocardiograph (ECG), heart-rate variability (HRV), heart rate(HR), galvanic skin response (GSR), respiratory, and skin temperature signals [7, 61].

Studies have explored drivers' galvanic skin response (GSR) and heart rate (HR) in driving situations, with results indicating higher autonomic activity during cooperation with pedestrians and the opposite trend during non-cooperative situations [1, 20]. Other works have combined EDA and controller area network (CAN) behaviour signals to study irritation [35], and bio-physiological signals, CAN, and GPS signal to study stress [47]. A recent study developed a novel mood-modulation system to induce different emotional states and measured the impact on participants' electroencephalogram (EEG) and photoplethysmogram (PPG) signals in an idle car environment [28]. Facial expressions, often combined with speech or driving behaviours, have also shown potential to capture overtly negative (e.g., frustration) and positive (e.g., joy) emotions [31, 52].

Few studies have explored driver affective states in outdoor or in-the-wild scenarios. Bethge et al. (2021) combined multiple modalities, including drivers' facial expressions and contextual driving data, to classify drivers' emotions [4]. Bethge et al. (2023) also conducted an in-the-wild car driving study using an unobtrusive setup to collect contextual data, demonstrating the validity of relying on contextual information for driver emotion recognition [3]. Contrary to driver sensing, another study combined driver emotion self-reports with GPS and weather data to infer emotions associated with driving, identifying a complex range of emotions associated with different driving scenarios [11].

Previous studies have concentrated on classifying driver emotions based on contextual factors, with less attention given to identifying specific emotions caused by pedestrian non-verbal actions at road crossings. Our study's novel, controlled outdoor setup captures multi-modal driver physiological and behavioural signals resulting from more realistic pedestrian non-verbal road crossing actions.

3 CAR DRIVING STUDY

Our study extends prior research on the impact of pedestrian crossing actions on drivers' affective states through web-based questionnaires or in-lab video-based settings [16, 45]. To enhance realism, we conducted a controlled, outdoor car driving study to capture drivers' affective states resulting from pedestrian non-verbal actions at road crossings. Here, we provide a detailed description of the study design, apparatus, and procedure.

3.1 Study Design

Our study followed a within-subjects design with a single independent variable (IV1: Performed action type: Positive vs. Non-positive vs. No action). *Performed action type* included 6 different crossing actions (3 positive, 3 non-positive) and 2 no-crossing (control) actions. Drivers drove on a circuit with a zebra crossing and encountered a pedestrian who performed a scripted road-crossing action. After each interaction, the driver verbally reported valence and arousal the 9-point discrete Self-Assessment Manikin (SAM) [5]. Physiological, thermal, and RGB facial information were captured by sensors during the entire duration of the experiment. The study was conducted for eight consecutive weekdays in June (Western Europe), capturing driving under diverse weather conditions including sunny, cloudy, and thundershowers (historical weather information included in supplementary material). The study followed strict guidelines from our institute's ethics and data protection committee, including COVID-19 regulations.

3.2 Study Setup

3.2.1 Driver Stimuli - Enacted Pedestrian Non-Verbal Crossing Actions. We selected non-verbal pedestrian crossing actions enacted by confederate pedestrians as driver emotion-inducing stimuli based on prior works [16, 45]. The selected actions included positive (*handwave, smile, nod*) and non-positive actions (*stay_back, impolite_hand_action, inattentive_with_phone*). We also included *no_pedestrian* and *pedestrian_present_but_not_crossing* as baseline control scenarios (Figures 1g and 1h). All actions and scenarios are included as part of the supplementary material.

3.2.2 Driving Circuit. The driver accompanied by the experimenter drove along the pre-determined driving circuit which was restricted for public use. Figure 2 shows a trial which was defined as the distance travelled from the *start* point and back. The circuit had a zebra-crossing where the driver encountered a pedestrian (*road-crossing marker*). The crossing was placed along the straight segment of the circuit, where drivers have enough time to see the crossing and stop safely for the pedestrian. Moreover, the maximum driving speed was set to 30 km/h. The *stop* marker was used as a reference for the experimenter to ascertain the point after which the driver were verbally asked to report valence and arousal scores.

3.2.3 Driving Trials. The driver completed 32 trials, with 12 trials involving positive interactions, 12 trials involving non-positive interactions, and 8 trials with no interactions. The choice of 32 trials was based on maximizing sensor data while ensuring an engaging driving experience and results from power analysis [3]. The order of actions was determined using a balanced Latin-square approach to ensure counterbalancing, with a pseudo-random generator used to initialise the set of factors while avoiding the occurrence of two *no-action* types together. The occurrence of the pedestrian was random, with both positive and non-positive actions performed equally by pedestrians to avoid driver anticipation and gender bias.

3.3 Study Apparatus

Our complete study apparatus comprises the following components - (a) thermal, RGB and physiological sensors and (b) sensor synchronisation application.

3.3.1 Sensors. Sensors were used to record drivers' real-time affective data, including an **Empatica E4** wristband for physiological signals (heart rate and skin conductance) at 64 Hz, a **GoPro Hero9** high speed camera for facial expressions at 240 fps, a **FLIR Duo Pro R** thermal camera for facial skin temperature and RGB images

[3]For effect size f=0.25 under α = 0.05 and power (1-β) = 0.95, with 24 repeated measurements within factors (discarding no interaction trials), one would need a minimum sample size of 12 participants.

(a) Handwave (b) Nod (c) Smile (d) Stay back

(e) Impolite hand action (f) Inattentive with phone (g) No pedestrian (h) Pedestrian present but not crossing

Figure 1: Non-verbal positive (a - c), non-positive (d - f) and no-action (g, h) pedestrian crossing actions performed during the study

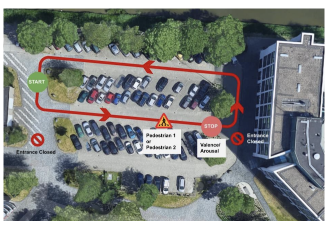


Figure 2: The circuit defined for the car driving study, with the driving path (trial) indicated via the red line. The driver stopped at the zebra crossing for the pedestrian. The *stop* marker served as a cue for the experimenter to ask drivers to verbally report valence and arousal scores.

at 29.97 fps, and a **Samsung Galaxy S10** for recording drivers' view of the road at 60 fps. Videos showing the drivers' view were later human-annotated to identify exact time periods of driver-pedestrian interactions. The car was also equipped with a 400 Watt inverter to power the FLIR thermal camera. The sensors used and their placement are shown in Figures 3 and 4.

3.3.2 Integrating Sensors Module. A modified Electron-based application [4] called SensorSync was used to synchronously start and stop sensor recording. The FLIR camera was modified with a custom ESP8266-ESP-12 micro-controller and connected to a circuit board with a Wi-Fi module that starts an HTTP server. The Empatica E4 wristband was connected to an Android mobile device running the EmpaticaRelay application, which connects to the software running on the micro-controller, and starts a TCP server to fetch data. Wi-Fi was provided via an external high-speed router (Figure 3c). Figure 4c shows the user interface (UI) of the SensorSync application that allows for the selected sensors to be connected. The Empatica data was stored on the EmpaticaRelay application, while all other sensors (FLIR camera, GoPro, Samsung) saved the data onto internal memory cards.

[4]https://github.com/electron/electron

3.4 Study Procedure

The experiment consisted of 4 sessions: pre-driving, driving, post-driving, and post-session. In the pre-driving session, drivers' demographic information and informed consent were obtained, and the car's sensors were set up and calibrated. A demo session was conducted to familiarise the driver with the circuit and relevant parameters. The SensorSync application was used to initiate all sensors to start recording, and a clapperboard was used to record the start and end times across all sensors. These cues were later used during data pre-processing to accurately identify the start and end of the driving session for all modalities. During each of the 32 trials, the driver encountered one of two pedestrians who performed a positive, non-positive, or no-action action, and reported Valence and Arousal scores on a 9-point Self-Assessment Manikin (SAM) scale (upon crossing the stop marker in Figure 2) [5]. An exit interview was conducted in the post-driving procedure, and subsequently the car was sanitised. The study lasted for about 45 to 60 minutes. Figure 5 illustrates the entire experiment and post-experiment session.

3.4.1 Overview of Measured Variables. We gathered qualitative and quantitative data from drivers throughout the duration of the study

(a) Empatica E4 wristband recorded drivers' heart and skin responses.
(b) The FLIR Duo Pro R recorded thermal and RGB camera recorded drivers' facial expressions.
(c) A Router provided Wi-Fi connectivity to the FLIR camera and Empatica.
(d) The Samsung Galaxy S10 captured RGB videos of the scene from the driver's perspective.
(e) The GoPro Hero 9 camera RGB Camera recorded drivers' facial expressions.
(f) A 400 Watt Inverter powered the FLIR camera.

Figure 3: Sensors and accessories used during the car driving study.

(a) Exterior of car used for the study.
(b) Final sensor setup (FLIR, GoPro and Samsung Galaxy) inside the car used to record driver and scene data.
(c) SensorSync application connected the FLIR, GoPro, Samsung S10 and Empatica devices, and was used to start and stop recording of data.

Figure 4: Car and sensor setup during car driving study.

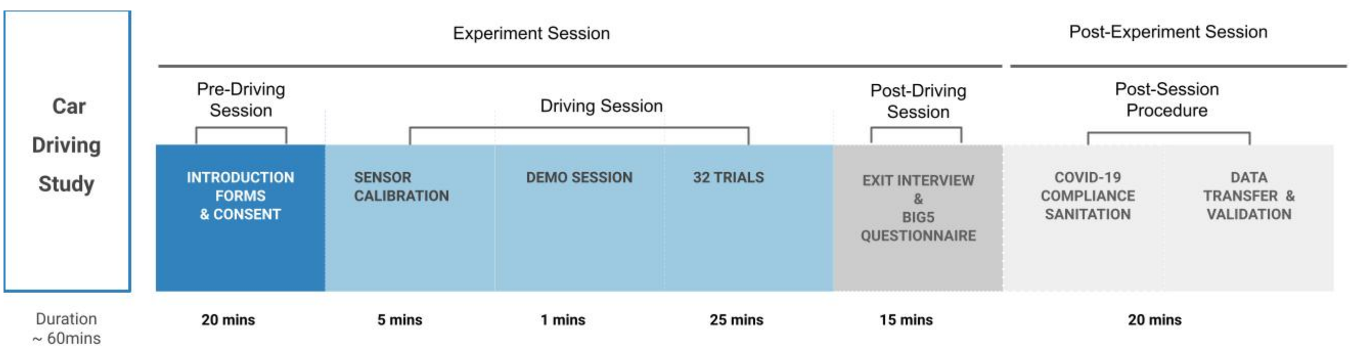


Figure 5: Overview of the car driving study procedure that comprised the experiment session and the post-experiment session.

and these are summarised in Table 1. Additionally, outdoor weather data was gathered from the local weather station and monitored for the duration of the study. These reports can be found as part of supplementary materials.

3.4.2 Covid-19 safety measures. To adhere to Covid-19 safety protocols, the experimenter sat in the rear-right seat of the car to maintain maximum possible distance from the driver. The windows were rolled down at least 10 cm for good air circulation, and in case of rainfall, only the rear-left and front-right windows were opened based on recommended guidelines [37]. All surfaces and sensors were disinfected after each experiment session to ensure safety.

3.4.3 Participants for the Car Driving Study. 40 drivers (age ≥ 21, driving experience ≥ 1 year) were recruited via a Western European agency. After accounting for data quality, 27 drivers were finalised (12 males, 15 females, with ages $21-60$ ($M = 39.93, SD = 13.42$) and driving experience of $1.5-42$ years ($M = 17.65, SD = 12.83$). Two confederates, one male (28) and one female (26), acted as pedestrians (Figure 1). Pedestrians received multiple training sessions

Type	Data	Data Collection Method	Experiment Segment
Subjective	Demographic information	Recorded on paper forms	Pre-driving session
	Study feedback	Recorded verbal interview	Post-driving session
	Valence and arousal self-reports	Verbally reported	Driving session
Objective	Heart rate	Empatica E4	Driving session
	Skin conductance	Empatica E4	
	Facial Thermal Data	FLIR Camera	
	Facial RGB Data	FLIR Camera; GoPro Camera	
	RGB road view from drivers' perspective	Samsung S10 Smartphone	

Table 1: Overview of the different qualitative and quantitative data gathered from drivers during the Car Driving Study.

on performing each action and timing their approach to the zebra-crossing with the approaching car for realism. They were provided with a list of actions, their order, and instructions on which trials to avoid crossing.

4 RESULTS

We report our analysis of drivers' affective responses from the controlled, outdoor driving study. We first explain the data pre-processing steps undertaken and thereafter discuss: (a) emotion self-report analysis, (b) physiological signal analysis, and (c) facial data analysis. Our study also recorded GoPro data of drivers' facial expressions, however these were excluded from analysis.

4.1 Data Pre-processing

Prior to analysis, the raw data for 27 finalised drivers underwent several stages of pre-processing which resulted in 1.79 TB of processed data (summarised in Table 2). The pre-processing steps are described below.

Data Type	Modality	Sensor	Total
Face	RGB	FLIR Dual Camera	1,069,272 frames
	Thermal		1,069,272 frames
Heart	BVP	Empatica E4	241,476 samples
	IBI	Empatica E4	214,631 samples
	BPM	Empatica E4	214,631 samples
Skin	GSR	Empatica E4	241,476 samples

Table 2: Final dataset after all pre-processing, and sampled at 30 Hz/frames per second (fps).

4.1.1 Driver Affective States Before and During Pedestrian Interaction. All driver signals (physiological, and facial) associated with pedestrian crossing actions were selected and processed using two segments: *before action* and *during action*. The *during action* segment was defined as the period from when the pedestrian stepped onto the zebra crossing until they stepped off. The *before action* segment was defined as the 5-second period prior to the pedestrian stepping onto the zebra crossing, based on empirical evaluation and following Schneeman et al. (2016), who found that driver interaction with pedestrians starts about 30 meters before the crosswalk in a 30 km/h zone [51].

4.1.2 Valence-Arousal Ratings Transformation. Valence and arousal self-reports corresponding to each pedestrian interaction were collected from every driver. Following prior work [16, 45], we grouped valence scores into *positive* or *non-positive* categories depending on whether they were ≥ 3 or < 3, respectively. Similarly, arousal scores were categorised as *high* or *non-high* scores.

4.1.3 Signal Cleaning and Processing. Facial data in the form of RGB and thermal videos were recorded at 29.97 fps using a FLIR camera. The thermal and RGB images were aligned using a homography matrix based on 14 selected points from each image [54, 56]. Galvanic Skin Responses (GSR) and Blood Volume Pressure (BVP) were gathered using an Empatica E4 wristband at 64 HZ. All missing values were interpolated linearly [42] and downsampled to 30 Hz (corresponding to the thermal camera), with GSR signals downsampled using a lowpass technique [15] and BVP signals downsampled using Stationary Wavelet Transform 7th level Daubechies mother wavelet [40]. Inter-beat interval (IBI) and heart beats per minute (BPM) were extracted from BVP, with outliers manually (23%) removed by visually examining the peaks in the plotted data [5]. The final dataset for analysis included aligned thermal and RGB frames, GSR, BVP, IBI, BPM signals, and self-reported valence and arousal.

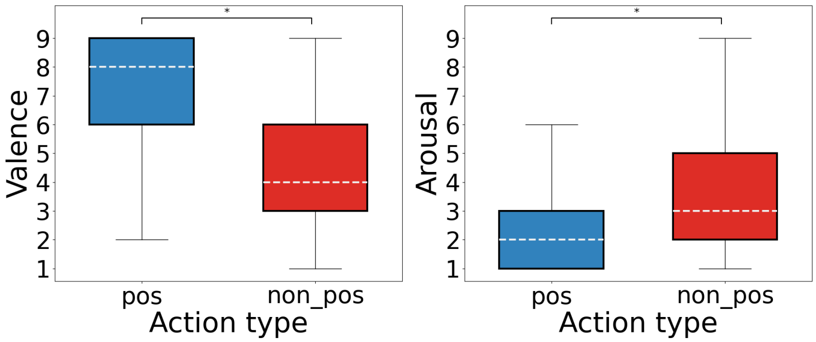
4.2 Emotion Self-report Analysis

Figure 6 shows self-reported valence and arousal scores grouped by positive and non-positive action types. Since the Shapiro-Wilk test revealed that the responses did not follow a normal distribution ($p < 0.05$), we ran a Mann-Whitney U test to evaluate the difference in the responses from the 9-point Self Assessment Manikin (SAM) scale. We see that drivers' self-reported valence ($U = 18991.5, Z = 13.81, p < 0.05, r = 0.771$) and arousal ($U = 32213, Z = -8.17, p < 0.05, r = -0.456$) scores vary significantly for the two action types.

4.3 Physiological Signal Analysis

Following prior work that studied driver affect, skin conductance (GSR), and heart data (BVP, BPM and IBI) were analysed for their variation in standard deviation [34, 45]. Specifically, we examined

[5]https://support.empatica.com/hc/en-us/articles/360030058011-E4-data-IBI-expected-signal



(a) Valence scores across action types **(b) Arousal scores across action types**

Figure 6: Emotion self-report variation grouped by action types for (a) valence and (b) arousal scores reveals a significant effect of action type.

the (a) variation in physiological signals across different action types - *positive, non-positive, no-action* and (b) the relationship between self-reported valence and arousal scores and variation in physiological signals.

We first compared variation in the standard deviation of all physiological signals across the three action types - *positive*, *non-positive* and *no-action.* The Shapiro-Wilk test revealed a non-normal distribution for all signals (BVP: $\chi^2(2) = 7.64, p < 0.05$, IBI: $\chi^2(2) = 24.94, p < 0.001$, BPM: $\chi^2(2) = 18.42, p < 0.001$). For all signals, a Kruskal Wallis test (accommodating the three action types) was used to identify significant differences following which post-hoc Mann-Whitney U test with Bonferroni corrections were employed.

Mann-Whitney U Test revealed significant differences for BVP between *pos* and *no_action* ($U = 35009.0, p < 0.001, r = 0.18$) and, *non_pos* and *no_action* pairs ($U = 19672.0, p < 0.001, r = 0.15$). For IBI, significant differences were found between *pos* and *no_action* ($U = 24551.0, p < 0.001, r = 0.33$) and, *non_pos* and *no_action* pairs ($U = 12018.0, p < 0.001, r = 0.281$). Finally, for BPM, significant differences were noted between *pos* and *no_action* ($U = 24983.0, p < 0.001, r = 0.29$) and, *non_pos* and *no_action* pairs ($U = 11577.5, p < 0.001, r = 25$). These are illustrated in Figure 7.

Physiological signals were also compared for their variation corresponding to two levels of drivers' self reported valence (*pos* and *non-pos*) and arousal (*high* and *non-high*) scores in Figure 8. In both cases, the Shapiro-Wilk test indicated a non-normal distribution ($p < 0.05$). Thereby, Mann-Whitney U tests revealed a significant effect ($p < 0.05$) of valence levels for BVP values ($U = 76041$, $p < 0.001$, $r = 0.132$*)*. Similarly, the two levels of arousal scores had a significant effect on BVP ($U = 107223.0, p < 0.001, r = 0.265$) and GSR ($U = 101254.0, p < 0.05, r = 0.148$) signals.

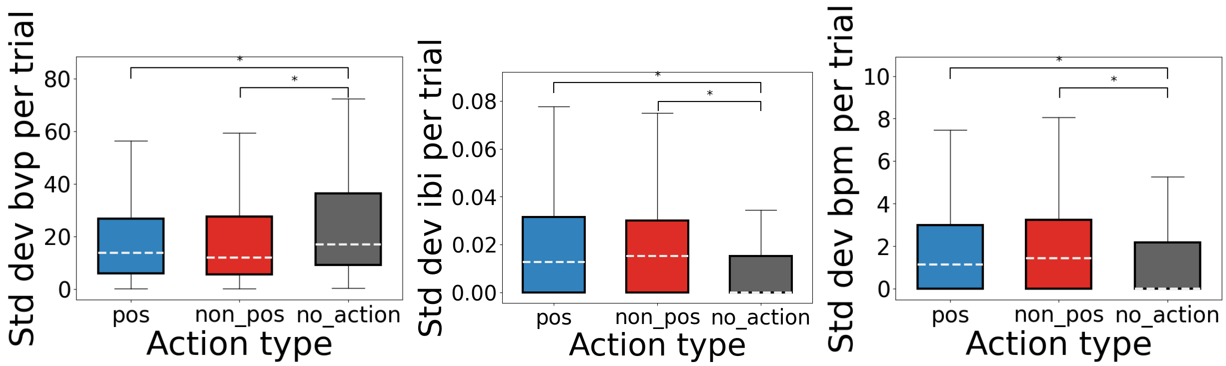
4.4 Facial Analysis

We present results from FLIR thermal data in the form of changes in facial landmarks analysis (a) before and during pedestrian interactions and (b) variation in facial landmarks corresponding to self-reported valence and arousal scores. We then repeat the analysis for the facial action units (FAUs) using the FLIR RGB data.

4.4.1 Facial Landmarks Analysis. We used OpenFace to detect 3D facial landmarks from drivers' aligned thermal images [2]. Handcrafted features were extracted as per Masip et al. (2014) [36]. Averaged landmark coordinates defined each frame, and dispersion was calculated as the average distance to the center. Standard deviation and difference between maximum and minimum dispersion were computed. Wilcoxon signed rank test analyzed non-parametric data [58], and the Paired T-Test assessed parametric data [22]. Drivers' facial landmarks did not show significant changes ($p < 0.05$) in response to positive and non-positive pedestrian actions. Similarly, there were no significant differences ($p < 0.05$) in facial landmarks based on drivers' self-reported valence (pos and non-pos) and arousal (high and non-high) scores.

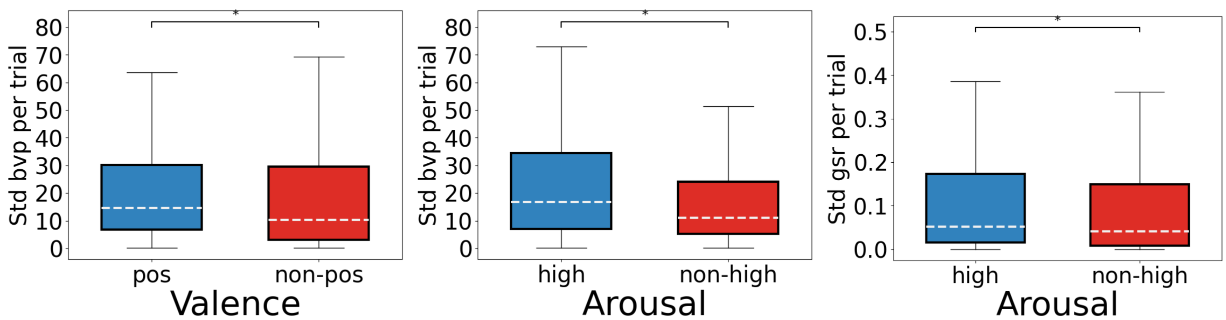
4.4.2 Facial Expressions Analysis. FLIR RGB data was analyzed using Facial Action Coding System (FACS) to identify facial expressions [9]. Activation levels were compared before and during pedestrian action. Standard deviation values compared facial data across the three action types. A Shapiro-Wilk test was performed to determine data distribution ($p < 0.05$), and a Wilcoxon signed-rank or a Paired T Test was used depending on the distribution. Significant differences were observed: **4** for positive trials, **2** for non-positive trials, and **7** for no-action trials [9, 23].

The summary statistics for positive trials are as follows - AU 04 (Brow Lowerer), Before: *Paired T: t(52) = 2.501, p-value = 0.0190*; AU 06 (Cheek Raiser) *Wilcoxon: U=54, Z=0.0012, p<0.05, r = 0.441*; AU 10 (Upper Lip Raiser) *Wilcoxon: U=65, Z=0.0028, p<0.05, r=0.405*; AU 20 (Lip stretcher) *Wilcoxon: U80, Z=0.0088, p<0.05, r=0.356.* The summary statistics for non-positive trials are as follows - AU 09 (Nose Wrinkler) *Wilcoxon: U=91.0, Z=2.354, p < 0.05, r=0.320*; AU 20 (Lip stretcher) *Wilcoxon: U=87.0, Z=2.451, p < 0.05, r=0.0224.* The summary statistics for no-action trials are as follows - AU 01 (Inner Brow Raiser) *Wilcoxon: U=102.0, Z= 2.090, p<0.05, r=0.284*; AU 02 (Outer Brow Raiser) *Paired T: t(52) = -3.76, p-value=0.0009*; AU 06 (Cheek Raiser)*Wilcoxon: U=94.0, Z=2.282, p<0.05, r=0.311*; AU 09 (Nose Wrinkler) *Wilcoxon: U=101.0, Z=2.114, p<0.05, r=0.288*; AU 10 (Upper Lip Raiser) *Wilcoxon: U=76.0, Z=2.714, p<0.05, r=0.369*; AU 20 (Lip stretcher) *Wilcoxon: U=32.0, Z=3.772, p<0.05, r=0.513*; AU



(a) BVP comparison **(b) IBI comparison** **(c) BPM comparison**

Figure 7: Significant variation ($p < 0.05$) in std dev. of trial-wise values across action types for (a) BVP (b) IBI (c) BPM signals.



(a) BVP variation across Valence **(b) BVP variation across Arousal Scores** **(c) GSR variation across Arousal Scores**

Figure 8: Significant variation ($p < 0.05$) in (a) BVP across two levels of valence; (b) and BVP and GSR across two levels of arousal scores.

25 (Lips part) *Wilcoxon: U=23.0, Z=3.988, p<0.05, r=0.543*. Detailed statistics are part of supplementary material.

We also examined the changes in facial expressions of drivers corresponding to self-reported valence and arousal scores. As the Shapiro-Wilk test revealed non-normal data distribution in both cases ($p < 0.05$), the Mann-Whitney U Test indicated that there was no significant effect of self reported valence and arousal scores on any facial action units.

5 DISCUSSION

5.1 Key Findings

We conducted a controlled study to investigate the impact of non-verbal pedestrian crossing actions on drivers' affective states using camera and physiological sensors. Our study showed that non-verbal actions significantly influenced drivers' physiological responses, facial temperature, and self-reported emotions. Our key findings include - (a) drivers reported greater valence scores upon observing positive pedestrian crossing actions and greater arousal scores for non-positive crossing actions (Figure 6). (b) Drivers' heart signals (BVP, IBI, and BPM) showed significant variations when observing positive and non-positive pedestrian crossing actions, compared to no actions (Figure 7). BVP also correlated significantly with self-reported valence and arousal scores, while skin conductance (GSR) varied significantly with arousal scores (Figure 8). (c) Facial landmarks analysis did not show any significant variation upon observing positive, non-positive and no-action types, and emotion self-reports. (d) Facial expressions analysis using facial action units (FAUs) revealed significant differences in brow (lowering), cheek (raising) and lip movements (raising, stretching) upon observing positive pedestrian actions. Nose (wrinkling) and lip (stretching) movements showed significant differences upon observing non-positive pedestrian actions. Brow (inner and outer brow raising), cheek (raising), and lip (raising, stretching, parting) movements varied significantly during the no-action scenarios. Additionally, we found eyebrow lowering (AU_04) to uniquely occur for drivers upon observing positive pedestrian actions. Typically associated with non-positive emotions, these AUs may have resulted from drivers misinterpreting positive actions such as a smile, as sarcastic and therefore non-positive [2], which is key to avoiding algorithmic bias (cf., [59]) in facial emotion expression inference. Next, drivers' brow movements (raising the inner and outer brow (AU_01, AU_02)), and parting of lips (AU_25) were unique to the no-action scenarios. Associated with perplexity, annoyance and frustration, these AUs may be a result of drivers misconstruing *pedestrian_present_but_not_crossing* as non-positive [2] (c.f. 4.4.2).

5.2 Implications

Our study's findings have implications for the development of automated driving systems that aim to enhance safety by predicting driver emotions using driver behavioural and physiological signals. Our work further validates the suitability of our selected positive and non-positive, non-verbal pedestrian actions as driver emotion-inducing stimuli given realistic driving conditions. Our findings expand on in-lab studies that put forth a set of driver emotion-inducing stimuli based on survey-based affective self-reports [16], as well as affective responses toward in-video pedestrian actions [45]. These pedestrian actions can serve as cues for road scene recognition systems that can better anticipate and respond to pedestrian crossing actions to enhance driver assistance features such as blind-spot monitoring and automatic emergency braking for road safety [63, 65]. Furthermore, we contribute an experimental study setup using a driving track equipped with sensors inside the vehicle that is suitable for collecting and inferring driver affect while ensuring the safety of drivers and confederate pedestrians. Our selected sensors captured driver affective cues arising from driver-pedestrian interactions, including heart (BVP, IBI, BPM), skin conductance (GSR), and facial expression changes. We found that non-positive pedestrian crossing actions elicited higher driver arousal associated with risky driving behaviour [8, 50], and our facial expression analysis revealed the role of subtle pedestrian facial cues in causing unintended non-positive emotions in drivers (e.g. positive pedestrian facial expression was misconstrued as non-positive by drivers), akin to real-world scenarios. Combined, these cues can aid researchers in selecting appropriate sensing modalities for detecting driver emotion signals related to non-verbal, pedestrian crossing actions.

5.3 Limitations and Future Work

Our study involved drivers in an outdoor circuit, encountering pedestrians performing road crossing actions. However, it lacked ecological validity due to safety concerns, potentially limiting its representativeness of real-world driving and driver-pedestrian interactions. This is because driver emotions may vary with simultaneous observation of multiple pedestrian actions. Multiple positive actions can evoke stronger positive reactions, while multiple negative actions may amplify non-positive responses. Further, mixed actions can elicit even more complex emotional responses from drivers. Nevertheless, our study contributes to understanding the impact of non-verbal pedestrian crossing actions on driver affective states in a controlled outdoor setting, prioritising participants' and confederates' safety. Future research could explore in-the-wild studies with less obtrusive sensors, enhancing realism of driver reactions.

Physiological signals (IBI, BVP, BPM, GSR) showed significant differences between positive/no-action and non-positive/no-action conditions. However, no significant differences were observed between positive and non-positive actions. This could be due to factors like sensitivity of chosen measures, individual differences, and limited stimulus variability. For example, the study primarily had experienced drivers (mean driving experiences was 18 years), who may have been unaffected by the pedestrian actions. Facial expressions analysis however revealed positive driver expressions (e.g. smiling) for positive actions and aversion expressions (e.g. nose wrinkling) for non-positive actions. Given the role of cultural and individual differences in interpreting facial expressions, multi-modal methods, including driver self-reports, enhance our understanding of emotional responses.

Lastly, while our study revealed variations in driver affective states based on observed positive and non-positive actions, we cannot draw precise conclusions about real-world driver emotions. Moreover, the effectiveness of such signals for just-in-time interventions in an empathic vehicle, enabling drivers to self-regulate their emotions in real-time, remains an open question [8]. Inferring aggressive driving solely from high arousal ([50]) or driver affective states solely from pedestrian actions can lead to incorrect conclusions. Incorporating additional data such as scene understanding, driving characteristics, and user feedback can enhance context and accuracy. User verification and feedback can further improve self-report emotion annotations and strengthen recognition models against classification errors.

6 CONCLUSION

Non-verbal communication between drivers and pedestrians has been found to impact driving behaviour. Therefore, inferring driver affect during on-road driver-pedestrian non-verbal interactions is key for developing empathetic in-car interfaces as part of improved road safety [10, 32, 55]. Drawing on realistic driving conditions, our controlled, outdoor car driving experiment with 27 participants investigated the effect of enacted, non-verbal pedestrian crossing actions on drivers' (*N=27*) affective responses. These include valence and arousal self-reports, heart and skin physiological signals, and facial expressions. Results showed drivers reported higher valence and arousal when observing positive and non-positive pedestrian crossing actions, and their physiological and facial expressions varied significantly. Our study validates our outdoor driving environment and non-verbal pedestrian crossing scenarios using empirical evidence that may serve towards the development of automatic, in-car empathic interfaces. Such interfaces, integrated into real-time emotion recognition systems, can have the ability to infer drivers' affective states based on observed pedestrian actions and facilitate "just-in-time" driver emotion regulation, thereby enhancing road safety.

REFERENCES

[1] Michela Balconi and Adriana Bortolotti. 2012. Empathy in cooperative versus non-cooperative situations: the contribution of self-report measures and autonomic responses. *Applied Psychophysiology and Biofeedback* 37, 3 (2012), 161–169.

[2] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. 2018. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*. IEEE, IEEE, Xi'an, China, 59–66.

[3] David Bethge, Luis Falconeri Coelho, Thomas Kosch, Satiyabooshan Murugaboopathy, Ulrich von Zadow, Albrecht Schmidt, and Tobias Grosse-Puppendahl. 2023. Technical Design Space Analysis for Unobtrusive Driver Emotion Assessment Using Multi-Domain Context. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–30.

[4] David Bethge, Thomas Kosch, Tobias Grosse-Puppendahl, Lewis L Chuang, Mohamed Kari, Alexander Jagaciak, and Albrecht Schmidt. 2021. VEmotion: Using Driving Context for Indirect Emotion Prediction in Real-Time. In *The 34th Annual ACM Symposium on User Interface Software and Technology*. ACM, New York United States, Virtual Event USA, 638–651.

[5] Margaret M Bradley and Peter J Lang. 1994. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry* 25, 1 (1994), 49–59.

[6] Michael Braun, Jonas Schubert, Bastian Pfleging, and Florian Alt. 2019. Improving driver emotions with affective strategies. *Multimodal Technologies and Interaction* 3, 1 (2019), 21.

[7] Michael Braun, Florian Weber, and Florian Alt. 2020. Affective Automotive User Interfaces–Reviewing the State of Emotion Regulation in the Car. *arXiv preprint arXiv:2003.13731* 54, 7 (2020), 23.

[8] Michael Braun, Florian Weber, and Florian Alt. 2021. Affective Automotive User Interfaces?Reviewing the State of Driver Affect Research and Emotion Regulation in the Car. *ACM Comput. Surv.* 54, 7, Article 137 (sep 2021), 26 pages. https://doi.org/10.1145/3460938

[9] Elizabeth A Clark, J'Nai Kessinger, Susan E Duncan, Martha Ann Bell, Jacob Lahne, Daniel L Gallagher, and Sean F O'Keefe. 2020. The facial action coding system for characterization of human affective response to consumer product-based stimuli: a systematic review. *Frontiers in psychology* 11 (2020), 920.

[10] Debargha Dey and Jacques Terken. 2017. Pedestrian interaction with vehicles: roles of explicit and implicit communication. In *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications*. Association for Computing Machinery New York NY United States, New York NY United States, 109–113.

[11] Monique Dittrich. 2021. Why drivers feel the way they do: An on-the-road study using self-reports and geo-tagging. In *13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM New York, NY, USA, Leeds United Kingdom, 116–125.

[12] Paul Ekman. 1992. An argument for basic emotions. *Cognition & emotion* 6, 3-4 (1992), 169–200.

[13] Paul Ekman and Wallace V Friesen. 1978. *Facial action coding system: Investigator's guide.* Consulting Psychologists Press, Palo Alto, California.

[14] Abdallah El Ali, Monica Perusquía-Hernández, Pete Denman, Yomna Abdelrahman, Mariam Hassib, Alexander Meschtscherjakov, Denzil Ferreira, and Niels Henze. 2020. MEEC: First Workshop on Momentary Emotion Elicitation and Capture. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1?8. https://doi.org/10.1145/3334480.3375175

[15] Koray S Erer. 2007. Adaptive usage of the Butterworth digital filter. *Journal of biomechanics* 40, 13 (2007), 2934–2943.

[16] Surjya Ghosh, Gerard Pons Rodriguez, Shruti Rao, Abdallah El Ali, and Pablo Cesar. 2022. Exploring Emotion Responses toward Pedestrian Crossing Actions for Designing In-vehicle Empathic Interfaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (submitted)* (New Orleans, LA, USA) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3491101.3519764

[17] Andrea Gorrini, Luca Crociani, Giuseppe Vizzari, and Stefania Bandini. 2018. Observation results on pedestrian-vehicle interactions at non-signalized intersections towards simulation. *Transportation research part F: traffic psychology and behaviour* 59 (2018), 269–285.

[18] Nicolas Guéguen, Sébastien Meineri, and Chloé Eyssartier. 2015. A pedestrian's stare and drivers' stopping behavior: A field experiment at the pedestrian crossing. *Safety science* 75 (2015), 87–89.

[19] Azra Habibovic. 2018. Communicating intent of automated vehicles to pedestrians. *Frontiers in psychology* 9 (2018), 1336.

[20] Jonathan L Helm, David Sbarra, and Emilio Ferrer. 2012. Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion* 12, 4 (2012), 748.

[21] Javier Hernandez, Daniel McDuff, Xavier Benavides, Judith Amores, Pattie Maes, and Rosalind Picard. 2014. AutoEmotive: bringing empathy to the driving experience to manage stress. In *Proceedings of the 2014 companion publication on Designing interactive systems*. Association for Computing Machinery, Vancouver BC Canada, 53–56.

[22] Henry Hsu and Peter A Lachenbruch. 2014. Paired t test. *Wiley StatsRef: statistics reference online* 2, 1 (2014), 1–3.

[23] Klas Ihme, Anirudh Unni, Meng Zhang, Jochem W Rieger, and Meike Jipp. 2018. Recognizing frustration of drivers from face video recordings and brain activation measurements with functional near-infrared spectroscopy. *Frontiers in human neuroscience* 12 (2018), 327.

[24] Myounghoon Jeon. 2015. Towards affect-integrated driving behaviour research. *Theoretical Issues in Ergonomics Science* 16, 6 (2015), 553–585.

[25] Myounghoon Jeon. 2016. Don't cry while you're driving: Sad driving is as bad as angry driving. *International Journal of Human–Computer Interaction* 32, 10 (2016), 777–790.

[26] Myounghoon Jeon and Bruce N Walker. 2011. 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. SAGE Publications Sage CA: Los Angeles, CA, SAGE Publications, Los Angeles, CA, 1889–1893.

[27] Myounghoon Jeon, Jung-Bin Yim, and Bruce N Walker. 2011. An angry driver is not the same as a fearful driver: effects of specific negative emotions on risk perception, driving performance, and workload. In *Proceedings of the 3rd international conference on automotive user interfaces and interactive vehicular applications*. ACM, ACM NY, NY, 137–142.

[28] Hodam Kim, Suhye Kim, Hongmin Kim, Youngsoo Ji, and Chang-Hwan Im. 2022. Modulation of Driver's Emotional States by Manipulating In-Vehicle Environment: Validation With Biosignals Recorded in An Actual Car Environment. *IEEE Transactions on Affective Computing* 13, 4 (2022), 1783–1792.

[29] Kevin Koch, Varun Mishra, Shu Liu, Thomas Berger, Elgar Fleisch, David Kotz, and Felix Wortmann. 2021. When do drivers interact with in-vehicle well-being interventions? an exploratory analysis of a longitudinal study on public roads. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (2021), 1–30.

[30] Julian Francisco Pieter Kooij, Nicolas Schneider, Fabian Flohr, and Dariu M Gavrila. 2014. Context-based pedestrian path prediction. In *European Conference on Computer Vision*. Springer, LNCS, Zurich, Switzerland, 618–633.

[31] Thomas Kosch, Mariam Hassib, Robin Reutter, and Florian Alt. 2020. Emotions on the go: Mobile emotion assessment in real-time using facial expressions. In *Proceedings of the International Conference on Advanced Visual Interfaces*. Association for Computing Machinery, Rome Italy, 1–9.

[32] Andreas Löcken, Carmen Golling, and Andreas Riener. 2019. How should automated vehicles interact with pedestrians? A comparative analysis of interaction concepts in virtual reality. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. Association for Computing Machinery New York NY United States, Utrecht, Netherlands, 262–274.

[33] Zhiyi Ma, Marwa Mahmoud, Peter Robinson, Eduardo Dias, and Lee Skrypchuk. 2017. Automatic detection of a driver's complex mental states. In *International Conference on Computational Science and Its Applications*. Springer, Springer, Cagliari, Italy, 678–691.

[34] Víctor Corcoba Magaña, Wilhelm Daniel Scherz, Ralf Seepold, Natividad Martínez Madrid, Xabiel García Pañeda, and Roberto Garcia. 2020. The effects of the driver's mental state and passenger compartment conditions on driving performance and driving stress. *Sensors* 20, 18 (2020), 5274.

[35] Lucas Malta, Pongtep Angkititrakul, Chiyomi Miyajima, and Kazuya Takeda. 2008. Multi-modal real-world driving data collection, transcription, and integration using Bayesian network. In *2008 IEEE Intelligent Vehicles Symposium*. IEEE, ACM, Eindhoven, Netherlands, 150–155.

[36] David Masip, Michael S North, Alexander Todorov, and Daniel N Osherson. 2014. Automated prediction of preferences using facial expressions. *PloS one* 9, 2 (2014), 2519–2528.

[37] Varghese Mathai, Asimanshu Das, Jeffrey A Bailey, and Kenneth Breuer. 2021. Airflows inside passenger cars and implications for airborne disease transmission. *Science advances* 7, 1 (2021), eabe0166.

[38] Albert Mehrabian. 1996. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology* 14, 4 (1996), 261–292.

[39] Jolieke Mesken, Marjan P Hagenzieker, Talib Rothengatter, and Dick de Waard. 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 (2007), 458–475.

[40] Guy P Nason and Bernard W Silverman. 1995. The stationary wavelet transform and some statistical applications. In *Wavelets and statistics*. Springer, NY, 281–299.

[41] Dimitris Nathanael, Evangelia Portouli, Vassilis Papakostopoulos, Kostas Gkikas, and Angelos Amditis. 2018. Naturalistic observation of interactions between car drivers and pedestrians in high density urban settings. In *Congress of the International Ergonomics Association*. Springer, Springer, Florence, Italy,, 389–397.

[42] Daniel J Perry, George E Mount, and Chester D Hull. 1959. The effect of varying intramuscular dosages of atropine and Banthine on the galvanic skin response. *The Journal of Psychology* 47, 2 (1959), 219–222.

[43] Rosalind W Picard. 2000. *Affective computing*. MIT press, Cambridge, Massachusetts.

[44] Robert Plutchik. 2001. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist* 89, 4 (2001), 344–350.

[45] Shruti Rao, Surjya Ghosh, Gerard Pons Rodriguez, Thomas Röggla, Abdallah El Ali, and Pablo Cesar. 2022. Investigating Affective Responses toward In-Video Pedestrian Crossing Actions using Camera and Physiological Sensors. In *Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. Association for Computing Machinery, Seoul, South Korea, 226–235.

[46] Shruti Rao, Surjya Ghosh, Gerard Pons Rodriguez, Thomas Röggla, Pablo Cesar, and Abdallah El Ali. 2023. From Video to Hybrid Simulator: Exploring Affective Responses toward Non-Verbal Pedestrian Crossing Actions Using Camera and Physiological Sensors. *International Journal of Human–Computer Interaction* 39 (2023), 1–24.

[47] George Rigas, Yorgos Goletsis, and Dimitrios I Fotiadis. 2011. Real-time driver's stress event detection. *IEEE Transactions on intelligent transportation systems* 13, 1 (2011), 221–234.

[48] Malte Risto, Colleen Emmenegger, Erik Vinkhuyzen, Melissa Cefkin, and Jim Hollan. 2017. Human-vehicle interfaces: The power of vehicle movement gestures in human road user coordination. In *Proceedings of the Ninth International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*. National Advanced Driving Simulator, Vermont, Iowa City, IA, 411–420.

[49] James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.

[50] Susan Raouf Hadadi Sani, Zahra Tabibi, Javad Salehi Fadardi, and Despina Stavrinos. 2017. Aggression, emotional self-regulation, attentional bias, and cognitive inhibition predict risky driving behavior. *Accident Analysis & Prevention* 109 (2017), 78–88.

[51] Friederike Schneemann and Irene Gohl. 2016. Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments. In *2016 IEEE intelligent vehicles symposium (IV)*. IEEE, IEEE, Gothenburg Sweden, 38–43.

[52] Björn Schuller, Matthias Wimmer, Dejan Arsic, Tobias Moosmayr, and Gerhard Rigoll. 2008. Detection of security related affect and behaviour in passenger transport. In *Ninth Annual Conference of the International Speech Communication Association*. IEEE, Brisbane, Australia, 265–268.

[53] Andreas Th Schulz and Rainer Stiefelhagen. 2015. Pedestrian intention recognition using latent-dynamic conditional random fields. In *2015 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, IEEE, Seoul, South Korea, 622–627.

[54] Namrata Srivastava. 2019. Using contactless sensors to estimate learning difficulty in digital learning environments. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*. ACM New York,, London United Kingdom, 399–403.

[55] Matus Sucha, Daniel Dostal, and Ralf Risser. 2017. Pedestrian-driver communication and decision strategies at marked crossings. *Accident Analysis & Prevention* 102 (2017), 41–50.

[56] Toshio Ueshiba and Fumiaki Tomita. 2003. Plane-based calibration algorithm for multi-camera systems via factorization of homography matrices. In *Computer Vision, IEEE International Conference on*, Vol. 3. IEEE Computer Society, IEEE, Nice, France, 966–966.

[57] Tianjiao Wang, Jianping Wu, Pengjun Zheng, and Mike McDonald. 2010. Study of pedestrians' gap acceptance behavior when they jaywalk outside crossing facilities. In *13th International IEEE Conference on Intelligent Transportation Systems*. IEEE, IEEE, Funchal, Portugal, 1295–1300.

[58] Robert F Woolson. 2007. Wilcoxon signed-rank test. *Wiley encyclopedia of clinical trials* 62, 1 (2007), 1–3.

[59] Tian Xu, Jennifer White, Sinan Kalkan, and Hatice Gunes. 2020. Investigating Bias and Fairness in Facial Expression Recognition. In *Computer Vision – ECCV 2020 Workshops*, Adrien Bartoli and Andrea Fusiello (Eds.). Springer International Publishing, Cham, 506–523.

[60] Sebastian Zepf, Monique Dittrich, Javier Hernandez, and Alexander Schmitt. 2019. Towards Empathetic Car Interfaces: Emotional Triggers while Driving. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu, Hawai, US, 1–6.

[61] Sebastian Zepf, Javier Hernandez, Alexander Schmitt, Wolfgang Minker, and Rosalind W Picard. 2020. Driver Emotion Recognition for Intelligent Vehicles: A Survey. *ACM Computing Surveys (CSUR)* 53, 3 (2020), 1–30.

[62] Tianyi Zhang, Abdallah El Ali, Chen Wang, Alan Hanjalic, and Pablo Cesar. 2020. RCEA: Real-Time, Continuous Emotion Annotation for Collecting Precise Mobile Video Ground Truth Labels. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1?15. https://doi.org/10.1145/3313831.3376808

[63] Zhipeng Zhang, Feiyu Wang, and Yunchao Wei. 2020. Pedestrian and Vehicle Detection using Convolutional Neural Networks. In *2020 IEEE 3rd International Conference on Information and Computer Technologies (ICICT)*. IEEE, IEEE, San Jose, CA, 105–110.

[64] Xiangling Zhuang and Changxu Wu. 2014. Pedestrian gestures increase driver yielding at uncontrolled mid-block road crossings. *Accident Analysis & Prevention* 70 (2014), 235–244.

[65] Xin Zou, Yao Huang, Qing Xu, and Zhitong Zhang. 2022. Road scene recognition based on a deep learning method: A review. *IEEE Access* 10 (2022), 16608–16628.