This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 814761.

Deliverable 2.1
State of the art on monitoring the driver state and task demand

Safe tolerance zone calculation and interventions for driver-vehicle-environment interactions under challenging conditions

iDREAMS

Project identification

Grant Agreement No	**814761**
Acronym	i-DREAMS
Project Title	Safety tolerance zone calculation and interventions for driver-vehicle-environment interactions under challenging conditions
Start Date	01/05/2019
End-Date	30/04/2022
Project URL	www.idreamsproject.eu

Document summary

Deliverable No	2.1
Deliverable Title	State of the art on monitoring driver state and task demand
Work Package	2
Contractual due date	29.02.2020
Actual submission date	28.02.2020
Nature	Revised after review
Dissemination level	Public
Lead Beneficiary	Austrian Road Safety Board (KFV)
Responsible Author	Susanne Kaiser
Contributions from	Susanne Kaiser, Anita Eichhorn, Eva Aigner-Breuss, Carina Pracherstorfer (Austrian Road Safety Board) Christos Katrakazas, Eva Michelaraki, George Yannis (NTUA) Fran Pilkington-Cheney, Rachel Talbot, Graham Hancox (University Loughborough) Evelien Polders, Veerle Ross, Tom Brijs, Kris Brijs (UHasselt) Chiara Gruden, Matjaž Šraml, Vlasta Rodošek, Tomaž Tollazzi (University Maribor) Petros Fortsakis (OSeven) Bart De Vos (DriveSimSolutions) Eleonora Papadimitriou (University Delft) André Lourenco, Carlos Carreiras (CardioID)

Revision history

Version	Issue date	% Complete	Changes	Contributor(s)
v1.0	23.12.2019	70%	Draft for review	Susanne Kaiser
V2.0	26.02.2020	95%	Revised after peer review / Content finalised	Susanne Kaiser
V3.0	27.02.2020	100%	Last details	Susanne Kaiser

Disclaimer

Copyright

Table of contents

List of Figures

List of Tables

Abbreviations

Abbreviation	Description
ANS	Autonomous Nervous System
AOI	Area Of Interest
APS	Average Pupil Size
BD	Blink Duration
BPM	Beat Per Minute
BR	Blink Rate
CFA	Confirmatory Factor Analysis
DRT	Detection Response Task
DSM	Driver State Monitoring
DWM	Driver Workload Metrics
ECG, EKG	Electrocardiogram
EDA	Electrodermal Activity
EEG	Electroencephalogram
EFA	Exploratory Factor Analysis
EOG	Electrooculogram
FIT	Fitness Impairment Tracker
fNIRS	Functional Near-Infrared Spectroscopy
GDPR	General Data Protection Regulation
GPS	Global Positioning System
GSR	Galvanic Skin Response
HFE	Human Factors Engineering
HR	Heart Rate
HRV	Heart Rate Variability
IBI	Inter-Beat Interval
ICA	Index of Cognitive Activity
ISA	Instantaneous Self-Assessment
IVIS	In-Vehicle Information System
KSI	Killed and seriously injured
KSS	Karolinska Sleepiness Scale
LCD	Lane Change Delay
MANOVA	Multivariate Analysis of Variance
MEG	Magnetoencephalography
MLFA	Maximum Likelihood Factor Analysis
MPDCR	Mean Pupil Diameter Change Rate
MWL	Mental Workload
NIR	Near-Infrared
NS.SCR	Non-Specific Skin Conductance Response
OBD	On-Board Diagnostic
OBU	On-Board Unit
PDT	Peripheral Detection Task
PERCLOS	Percentage of Eyelid Closure
PPG	Photoplethysmography
PVT	Psychomotor Vigilance Task

RT	Reaction Time
RMSSD	Variability of the Inter-beat interval
RSME	Rating Scale Mental Effort
SC	Skin Conductance or Conductivity
SCL	Skin Conductance Level
SCR	Skin Conductance Response
STZ	Safety Tolerance Zone
VRU	Vulnerable Road Users

Executive summary

The i-DREAMS project aims at setting up a framework for the definition, development, testing and validation of a context-aware safety envelope for driving called the ‘Safety Tolerance Zone’. Taking into account driver background factors and real-time risk indicators associated with the driving performance as well as the driver state and driving task complexity indicators, a continuous real-time assessment will be made to monitor and determine if a driver is within acceptable boundaries of safe operation. Moreover, safety-oriented interventions will be developed to inform or warn the driver in real-time in an effective way as well as on an aggregated level after driving, through an app- and web-based gamified coaching platform (post-trip intervention). Furthermore, a user-license Human Factors database with anonymized data from the simulator and field experiments will be developed.

The conceptual framework of the i-DREAMS platform integrates aspects of monitoring (such as context, operator, vehicle, task complexity and coping capacity), to develop a safety tolerance zone for driving. In-vehicle interventions and post trip interventions will help to maintain the safety tolerance zone as well as provide feedback to the driver. This conceptual framework will be tested in simulator studies and three stages of on-road trials in Belgium, Germany, Greece, Portugal and the United Kingdom with a total of 600 participants representing car, bus, truck and train drivers. In relation to driver-related factors, operator state and trait factors will also be considered and measured as part of this platform.

The documented work was guided by the following objectives:

- Identification of measurable factors contributing to the overall risk level at a given time and documentation of corresponding indicators for the i-DREAMS modes (car, bus, truck, train)
- Review and assessment of state of the art (in-vehicle) technology suitable to track those indicators and combine them to get a real-time indication of risk
- Give recommendations on set of relevant systems suitable for test case implementation as a next step in the project

Currently, there is no standard procedure for measuring the driver’s coping capacity and task complexity, with a plethora of methods, indicators and algorithms, each with strengths and drawbacks. Capturing real-time workload assessment by monitoring driver state and driving context evoked task demand is the main objective of the work documented in this report. Moreover, the conclusions drawn from this review serve as the base for selecting appropriate measuring systems and devices for the future project work and for building the theoretical and mathematical model which are the backbone of the development of the i-DREAMS platform. Constructs to be measured are task demand, the driver’s cognitive and affectional state (mental state) in terms of attention and distraction, fatigue and sleepiness, and emotional states, and driving behaviour, as well as more stable characteristics which are known to impact safe driving.

A systematic search of relevant scientific and grey literature was conducted for each of the three key driver mental states (attention and distraction, fatigue and sleepiness, and emotion), with the purpose of identifying approaches to measure the various driver states. Search terms were generated for each of the mental states and entered in combination in various well-established databases. The findings were screened by title and then abstract, and relevant literature was documented and summarised. This report contains a dedicated section in chapter five for each of the four main driver mental states. The literature predominantly concerned car driving, however the extent of the transferability of the findings to the other i-DREAMS modes (truck, bus, train and tram), is discussed in each section.

Chapter four of this report focuses on measuring task demand and the indicators, methods and technology for the real-time monitoring of driving context. Measures indicating cognitive workload as well as task demand impacted by traffic environment factors were considered. The three main classifications of measuring task demand were subjective, performance and physiological measures. In terms of subjective measures, it was found that the Rating Scale for Mental Effort (RSME) and the NASA Task Load Index (TLX) were reliable and sensitive indicators for workload measurement. Performance measures focused on techniques that registered driver ability to perform driving tasks at acceptable or safe levels, with lateral position deviation on if the most important indicators. Speed, longitudinal control and reaction time were other important indicators to assess driving performance. Physiological measures included heart rate, EOG, EEG, EDA, head movements and evaluation of eye movement, with the most frequent and reliable measures as eye fixations and ECG signals which could potentially be measured through steering wheel sensors and eye tracking systems.

The indicators, methods and technology for the three key driver mental states is reviewed in chapter five. In terms of attention and distraction, in most cases driver distraction was measured in terms of impact on attention, behaviour and accident risk. Real time eye tracking, radars for physiological measurement and cameras were the most frequent detection techniques. Driver distraction could also be measured by lateral/longitudinal measurement, safety measures such as reaction time or gap acceptance, and eye or workload measures. However, the literature indicated a diverse range of methods and experimental design. The majority of studies reviewed in relation to fatigue focused on ‘sleepiness’ rather than task related fatigue, with most research conducted in simulators. EEG was the most frequently used measure, however, could be impractical for use with the proposed system. HRV shows potential and can be developed into unobtrusive measures, however this method has not been used much in operational settings and needs further development and validation. Ocular measures are reliable and utilised mostly in commercial sleepiness detection technology, with blink duration and PERCLOS being the most robust indicators. Many studies used multiple measures and indicators to detect sleepiness, which could aid in the detection reliability. The review of literature focusing on measuring emotion, anger, frustration, aggression, stress, fear and anxiety were the most frequently studied constructs. In terms of potential indicators of emotional states, EDA and heart-based measures were the most frequently used indicators, with the majority of studies using more than one measure. With regards to the measurement of substance impairment, driving under the influence of drugs and medicines is less well understood than drink driving. There is limited universal agreement on the most reliable way to measure impairment from drugs or medicines. Wearable technologies have recently been developed to monitor substance impairment; however, these are insufficiently validated. The majority of available technology focuses on monitoring impairment from alcohol, for example wrist worn transdermal alcohol sensors.

Indicators and methods of measuring driver characteristics is detailed in chapter six. Many variables important for assessing a driver’s capability are not suitable for real-time measuring since they are more robust over time and less sensitive to situational influences. The one-time measurement of cognitive capabilities and competences such as attention regulation and reactivity as well as personal factors such as personality, experience, age, gender, cultural identity and health status is advised. Many factors can be measured through surveying; however, some require commercially available performance test equipment. Collecting driver characteristics about the i-DREAMS participants serves various goals in the project: populating the i-DREAMS research data base, customizing interventions, accounting for covariates and possibly introducing stable factors into the Safety Tolerance Zone model as correction factors.

Chapter seven details the indicators, methods and technology for the real-time monitoring of driver behaviour. Reviewed tools included cameras, smartphones, OBD-II, GPS, radar, lidar, laser, steering angle sensors, distance sensors, brake and gas pedal sensors, speed sensors, yaw rate sensors, thermal radiation sensors, infrared sensors, digital tachograph, potentiometer, inertial sensors. Both direct and indirect measures could be used to monitoring driving behaviour, including speed, trajectory, acceleration and time to collision. However, the majority of the literature focused on on-road vehicles and possibly would not be applicable to trains and trams.

In terms of the implications for i-DREAMS and the relevance for the project's technology, an overall conclusion that can be drawn is that two physiological/behavioural measurement methods should be used for the continuous driver monitoring. This insight applies to measuring all of the single constructs: task demand attention and distraction, fatigue and sleepiness as well as emotions and related constructs. Thereby, drawbacks of a single measurement method can be compensated for, assuring validity is facilitated. While heart rate measures show promise, heart rate and heart rate variability are sensitive to inter-individual differences and confounding factors, which need to be considered. ECG indicators can be recorded through sensors on the steering wheel, with the potential to measure several driver states, a complementary eye tracker seems beneficial. An additional camera can be used to track head and eye features and changes in facial features as an indication of emotional states. Wrist worn measures may also be beneficial, and could be used to heart rate, EDA, as well as alcohol impairment. In addition, advanced driver-assistance systems, which utilize a forward-facing camera and provide warnings for collision prevention and mitigation, as well as smartphone applications which can measures lateral and longitudinal acceleration, can be utilized as surrogate safety measures.

The table below summarises the operator states, and the recommended measures, technology and thresholds for use when monitoring driver task complexity and coping capacity.

Operator state	**Optimal measure**	**Ideal technology**	**Influence on coping capacity/ task demand**	**Safety critical threshold**	**Frequency of measure (real time or one-off)**
Attention and distraction	- PERCLOS - PERLOOK - Glance duration - Head movement - driver behaviour (lateral and longitudinal measures, reaction time, gap acceptance)	- Eye tracker (glasses / system) - Driver facing camera - Forward facing camera and collision avoidance system (Mobileye)	Increased PERCLOS, PERLOOK, glance duration, head movements = increased distraction and reduced coping capacity.	- PERCLOS and PERLOOK > 35% - Glace duration of 2 seconds - Head turns > 5 seconds	Real time
Alertness (fatigue / sleepiness)	- Blink rate - PERCLOS - Heart rate variability (HRV)	- Eye tracker (glasses / system) - Driver facing camera - Heart rate sensors embedded in steering wheel (CardioWheel)	Slowed blink rate, increased PERCLOS = increased sleepiness and reduced coping capacity. HRV data mixed findings	Various thresholds reported	Real time

		- Wearable heart rate monitor			
Emotion, stress	- ECG (heart rate) - EDA	- ECG sensors (CardioWheel) - EDA wearable device - Driver facing camera - Eye tracker (glasses / system)	Increased heart rate and EDA = increased emotional response and reduced coping capacity	Unsure	Real time
Substance impairment	- Blood and Urine samples - Tissue readings -Breathalysers - EDA	- Wearable sensors (TruTouch)	Increased reading of impairment = reduced coping capacity	Unsure	Real time and one-off
Driving behaviour	- Speed - Braking - Lateral and longitudinal movement - Trajectory - Acceleration - Time to collision	- Forward facing camera and collision avoidance system (Mobileye) - Smart phones - Various driving sensors	Increased variables = reduced coping capacity	Various thresholds reported	Real time, post trip

Overall recommendations for i-DREAMS include:

- Most of the evidence is available for car drivers. The transferability of some of the findings to trucks, busses, trams and trains may partly be determined in an iterative process and by actual trial and error
- ‘Mental state’, ‘emotions’, ‘distraction’ etc. are theoretical constructs that ask for deciding on one of the plethora of definitions and theoretical concepts.
- Using at least two approaches for driver state monitoring will be beneficial for assuring validity and reliability
- Majority of driver mental state variables could be measured with cameras, eye tracking, and heart rate sensors either embedded in the steering wheel or incorporated into It should be considered that the use of devices that have to be put on or activated by the participant before driving may compromise the naturalistic driving character of the trials.
- The potential to consider the drivers’ traits and characteristics in the calculation of the safety tolerance zone should be explored.
- Thoroughly testing indicators and measures at the simulator stage is indispensable

1 Introduction

1.1 About the i-DREAMS project

The overall objective of the i-DREAMS project is to setup a framework for the definition, development, testing and validation of a context-aware safety envelope for driving ('Safety Tolerance Zone'), within a smart Driver, Vehicle & Environment Assessment and Monitoring System (*i*-DREAMS). Taking into account driver background factors and real-time risk indicators associated with the driving performance as well as the driver state and driving task complexity indicators, a continuous real-time assessment will be made to monitor and determine if a driver is within acceptable boundaries of safe operation. Moreover, safety-oriented interventions will be developed to inform or warn the driver real-time in an effective way as well as on an aggregated level after driving through an app- and web-based gamified coaching platform. Figure 1 summarizes the conceptual framework, which will be tested in a simulator study and three stages of on-road trials in Belgium, Germany, Greece, Portugal and the United Kingdom with a total of 600 participants representing car, bus, truck and train drivers.

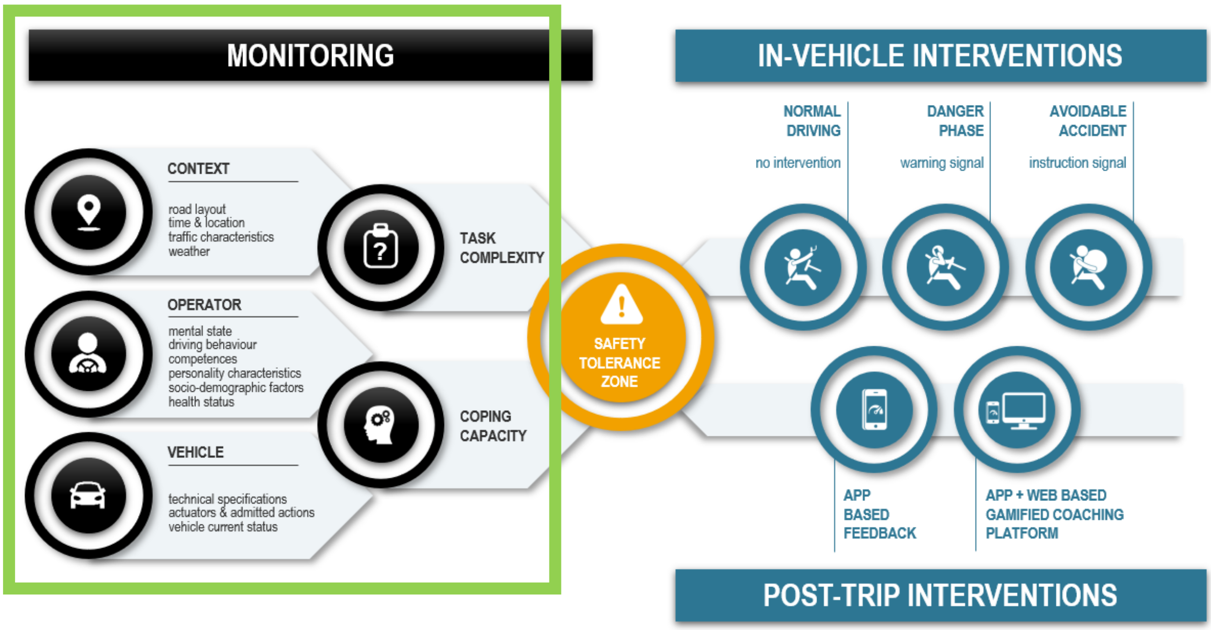

Figure 1: Conceptual framework of the i-DREAMS platform. The green dotted frame indicates the thematic scope of this deliverable (see section 1.2)

Expected by the end of the project in 2022, the key output of the project will be an integrated set of monitoring and communication tools for intervention and support, including i.e. in-vehicle assistance and feedback and notification tools as well as a gamified platform for self-determined goal setting working with incentive schemes, training and community building tools[1]. Furthermore, a user-license Human Factors database with anonymized data from the simulator and field experiments will be developed.[2]

[1] A state-of-the-art assessment of real-time and post-trip intervention approaches are documented in Deliverable 2.2 of this project.

[2] Further general project information can be found on the website: https://idreamsproject.eu

1.2 About this report

The work presented in this deliverable addresses the left half of Figure 1, the monitoring of the task demand and complexity based on contextual factors of the driver environment and the monitoring of the driver (car, truck, bus and train) in real-time; both with the aim of eventually determining whether an individual operates within or without a safe zone.

In recent years, considerable research has been conducted in relation to these topics and the development of associated technologies is progressing fast – especially in view of the autonomous driving boom – which has resulted in an enormous variety of approaches, data collection methods, monitoring and warning equipment etc. This variety and ever-changing technology maturity, however, also reflects the circumstance that there is not a standard approach but quite the contrary, a multitude of tools, measurement methods, indicators, algorithms etc. with specific advantages and drawbacks, delivering their full potential in specific contexts. Therefore, capturing the state of the art of monitoring techniques and subsequently assessing the applicability for i-DREAMS' endeavour is indispensable.

The documented work was guided by the following objectives:

- Identification of measurable factors contributing to the overall risk level at a given time and documentation of corresponding indicators for the i-DREAMS modes (car, bus, truck, train)
- Review and assessment of state of the art (in-vehicle) technology suitable to track those indicators and combine them to get a real-time indication of risk
- Give recommendations on set of relevant systems suitable for test case implementation as a next step in the project

To achieve those objectives a comprehensive literature search (scientific as well as grey literature) was conducted and identified measurement methods and associated technologies were assessed based on pre-defined criteria such as intrusiveness, validity etc. The review started with the transportation mode which is covered most extensively in literature: the car. Following this, the transferability of the results to the other three i-DREAMS modes was assessed and if necessary, a dedicated further search for a certain mode was carried out. Where applicable, the circumstances of professional drivers versus non-professional drivers were considered, as this is an integral part of the i-DREAMS objectives.

While the context factors such as road layout or weather are somewhat self-explanatory, many driver-related factors are not. When discussing the *mental state* of a driver, the constructs attention vs. distraction, sleepiness and fatigue, emotions, arousal, stress and substance impairment (alcohol, illegal and prescribed drugs) are referred to. Driver behaviour patterns (longitudinal and lateral control etc.) can also be a result of the driver's state and measured in real time. Therefore, driver behaviour indicators were considered in state-of-the-art analysis as well (Figure 2) and which ultimately will determine the safety tolerance zone together with the driver state and the driving context. The state factors change continuously during driving while the operator's or driver's trait factors are more stable over time but nevertheless, affect driver behaviour and driver safety. Figure 2 provides an overview of exemplary operator state and trait factors that are considered in this state-of-the-art analysis.

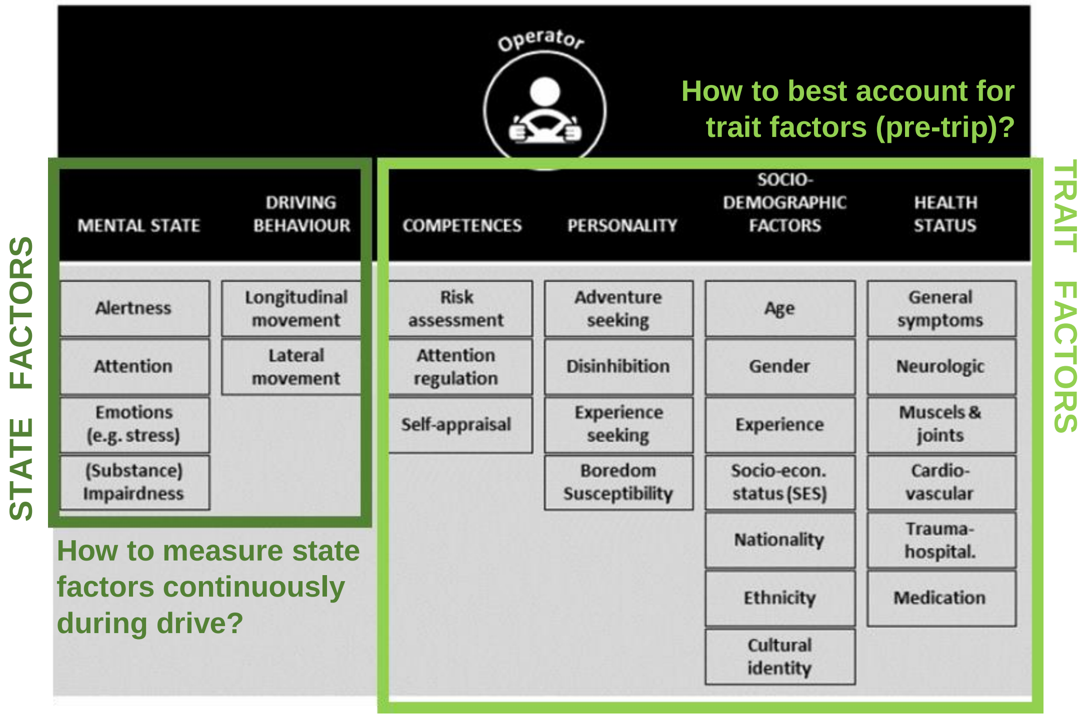

Figure 2: Exemplary driver state and trait factors, which will be measured before the drive (categorial) and during the drive (continuously)

Although, the relevant factors and associated measurement techniques and tools – in regard to the environment as well as the operator – are elaborated in sequential, dedicated chapters in this report, this does not mean they are independent from each other. Mental state factors can limit the available coping capacities and together with the factors defining the task complexity at a given moment, they determine workload in a complex, non-linear way. For example, a certain driving behaviour, like compromised lateral control, can be an expression of fatigue. Health factors like mild cognitive impairment can influence the mental state and the driving behaviour – to some extent independently and to some extent in a chained sequence. Where possible, these intercorrelations and pitfalls have been considered throughout this review. However, the main purpose of this report is to document the factors which should be considered throughout the project and the review as well as assessment of measuring methods, both in a simulator and in the real-world driving environment. This results in the recommendations as reported in chapter 7.

2 Theoretical considerations and introduction of concepts and terminology

Designing and validating a holistic driver monitoring system ideally requires an exhaustive list of factors contributing to the dynamic safety envelope, the safety tolerance zone (STZ). That is, broadly speaking, a combination of factors determining the driving task demand and the individual capacities to cope with the task. The dynamic combination of those two concepts can also be referred to as workload and is partly expressed in driver performance.[3]

The driving task can be characterised as the '*dynamic control task in which the driver has to select relevant information from a vast array of mainly visual inputs to make decisions and execute appropriate control responses'* and '*drivers execute planned actions which are shaped by their expectations of the unfolding road, pedestrian and traffic scenario in front of them and the reality that they actually observe*' '(Shinar & Oppenheim, 2011, p.216). Thus, it is partly determined by exogenous factors of the driving environment and partly by the driver's perception, planning and execution abilities. The latter in turn, is influenced by a plethora of situational and continuous driver characteristics. The mental state, consisting of cognitive and affective state, is considered dynamic, can change constantly and is aimed at being measured in real-time through physiological indicators while driving. Driver characteristics, such as personality and experience, are more stable over time and thus, do not require real-time assessment but are still important to consider.

In an attempt to be as comprehensive as possible with single influencing factors, the most obvious source is accident statistics and systematic road safety risk assessments (e.g. the SafetyCube DSS[4]). Psychological fitness to drive regulations or the Goals for Driver Education (Hatakka et al., 2002) are valuable additional sources.

2.1 Driver state monitoring

As the real-time measurement of physiological and behavioural indicators is crucial for the i-DREAMS concept. In this report the most important ones will be introduced below with definitions and descriptions.

In general, physiological measures refer to the activity of the autonomous nervous system, for example the heartbeat. This activity cannot (or hardly) be controlled by an individual whereas behavioural measures refer to the movement of body parts that can actively be controlled, such as the eyes or facial expression, and which, however, are not always controlled consciously.

Among the physiological indicators, the individual's level of arousal is a central concept which is linked to attention, alertness, stress and emotions (Borghini et al., 2014). Arousal is largely affected by the autonomous nervous system (ANS), which includes sympathetic and parasympathetic branches. The sympathetic branch generates an alerting response in stressful situations, which can be recognized by increased breathing rate, accelerated heart rate (HR), sweaty palms and dilated pupils. The parasympathetic branch is mostly activated during relaxed situations, such as sleep periods and leads to decreased breathing rate, HR and blood pressure. Therefore, breathing rate, HR and skin moisture are examples of indicators of the ANS's activity, which in turn can indicate the driver's arousal level and alertness.

[3] While the current state of the vehicle plays a role as well, it is not subject to the considerations of this state-of-the-art review.

[4] https://www.roadsafety-dss.eu/

Significant advancements have been made recently in the domain of signal processing and in developing signal acquisition and processing methods for driver monitoring systems, typical referred to as Driver State Monitoring (DSM). Many DSM systems combine various sources of information, including sensors that measure in-vehicle indicators (e.g. the steering angle), sensors measure a driver's physiological signals (e.g. heartbeat or changes in blood volume) and camera(s) installed in the passenger cabin to detect behavioural indicators such as the head position or frequency of eye lid closure.

Table 1 summarizes a list of the most important physiological indicators and measurement methods related to DSM. No claim is made to completeness. Further indicators are introduced in the subsequent chapters.

Table 1: Short descriptions of physiological measures which are commonly used as indicators for attention, fatigue, stress and emotional states

Physiological and behavioural indicators and measures of the driver state	
Blood oxygen saturation	Blood oxygen saturation can be measured with a pulse oximeter device, which is worn on the finger. The ratio of oxygenated versus deoxygenated haemoglobin in the blood and blood volume results in the saturation. The oxygen saturation impacts brain functions such as memory, decision making and attention. (Mabry et al., 2019)
Blood pressure	Blood pressure is the force on the blood vessels and is dependent on the stage of the cardiac cycle (low to high) and is affected by various factors such as age, stress or environmental influences. It is described as a pair of the systolic (highest level) and diastolic (lowest level) value. Arousal during mental effort leads to greater cardiovascular reactivity and thereby to increased blood pressure. The most commonly used non-invasive measuring method is auscultatory measurement recording the sound of the blood flow. (Balters & Steinert, 2015; Schmidt, 2017; Lohani et al., 2019)
EDA, Electrodermal activity GSR, Galvanic skin response SC, Skin conductance	EDA, GSR or SC is the continuous variation in the electrical characteristics of the skin. Measuring tools consist mostly of two electrodes applicated on fingers or toes to measure the electrical conductance between two points on the skin. It is not steady but varies in relation to other factors such as the moisture level (sweat) (Gonzalez-Sanchez et al., 2017;)
EEG, Electro-encephalogram	Neuronal, electrical activity of the brain and brain waves can be detected with the help of EEG by applying electrodes on the scalp. The neuronal activity exists between positive and negative potentials of the electrodes. A neutral electrode is needed for reference. Electroencephalography can detect several types of waves, for example alpha, beta, delta and theta waves which each are specific for different states such as a conscious state of mind or cognitive processes and sleep phases (Balters & Steinert, 2015; Madry et al., 2019).
EOG, Electrooculography	Electrooculography is a technique to measure the movement and position of the eyes, such as saccadic movements and fixations, blinking (frequency, duration, velocity, amplitude) and share of time the eyes are closed (PERCLOS). Electrodes are placed above and below the eyes for measuring electric potentials between them (Jia & Tyler, 2019).
ECG, Electrocardiogram	ECG is a non-invasive measurement method to measure heart rate and heart rate variability. The contraction of the heart is based on depolarization. ECG uses the electrical potential difference between two electrodes which is caused by cardiac potential differences (electrical activity of the heart). At least three electrodes are necessary for an ECG,

	whereas one electrode is used as a neural electrode. (Stockburger & Möckel, 2016; Balters & Steinert, 2015)
fNIRS, functional near-infrared spectroscopy	fNIRS is a neuro imaging technique that measures changes in concentration of oxygenated and deoxygenated haemoglobin in the cortex by means of near-infrared light. This is associated with neuronal activity. Compared to other neuro imaging methods, it is more robust in relation to movement. (Sangani et al., 2015; Mabry, 2019)
Head movement	By means of videography or head tracking devices, the head's position and movement can be measured, which used to identify fatigue/sleepiness and distraction (Mabry, 2019).
HR, Heart rate HRV, Heart rate variability	Heart rate is simply referring to the number of heartbeats per minute, which is between 60 and 90 beats for the average adult. ECG is most frequently used to monitor the cardiac activity in the laboratory and controlled environments. HR increases with increasing activity of sympathetic nervous system as well as with the decrease of the parasympathetic activity (and vice versa). HRV is the variation of the time between heartbeats and is linked to other physiological information such as respiration, temperature and vasomotor activity. (Shaffer & Ginsberg, 2017; Mabry, 2019)
IBI, Inter beat interval	The inter-beat interval is also called beat to beat interval and is the interval between the heart beats.
Ocular measures (other than EOG)	Movement of eyes, eye-blinking, pupil diameter and eyelid closure can not only be measured by means of EOG but also with videography or eye tracking. The indicator PERCLOS represents the percentage of time during which the pupils are covered by more than 80%. PERCLOS is mainly used for detecting fatigue and sleepiness. There are devices which directly measure and feedback on PERCLOS. However, ocular measures are also very important for the indication of attention and distraction (scanning patterns, direction of views etc.). (Mabry, 2019)
PPG, Photoplethysmography	Changes of blood volume can be detected with the optical method of photoplethysmography. It makes use of infrared light and can measure on the skin surface. The retrieved waveform is associated with the cardiac synchronous changes in the blood volume. (Allen, 2007; Rabe & Gerlach, 2005)
Skin temperature, thermography	Changes in skin temperature can serve as indicators for e.g. emotional states. Thermal cameras are used to record radiation emitted in the mid/long wavelength, which allows to detect the surface temperature of individuals. Facial thermography captures the heat distribution on the forehead or nose which varies depending on the sympathetic activity. (Lohani et al., 2019)

3 Task demand – indicators, methods and technology for real-time monitoring of driving context

Learning to drive demands a lot of practice before expert levels are reached. To begin with, task demand is determined by goals that have to be reached by performance (de Waard, 1996; Fairclough et al., 2005; Paxion et al., 2014). The driving task is partly determined by the demands of the road environment, traffic restrictions, weather conditions and time of the day or location (European Commission, 2019). However, the complexity of the driving task is also associated with driver performance, such as harsh events, driving speeds, or following distances. In order to capture all possible mechanisms of driving context and their influence on the driving task, task demand was investigated in terms of both cognitive workload (section 4.1), as well as the impact of exogenous factors on road safety (section 4.2). As the perception of the driver is also considered in the i-Dreams project, section 4.3 discusses the potential of using physiological measurement to assess subjective task demand. Furthermore, because the majority of the studies with regards to task demand are concerned with passenger cars, section 4.4 discusses the transferability of task demand monitoring for trains, buses and trucks. Finally, conclusions and recommendations are drawn in section 4.5.

3.1 Task demand measured as cognitive workload

Technologies for monitoring task demand are developed in order to mitigate the contextual effect on driving and contribute to driver behaviour traffic road safety enhancement (Sowmya, 2014; Sezgin & Lin, 2019). In order to identify the most relevant factors contributing to driving task demand, a literature search was initiated to correlate the aforementioned four factors (road layout, traffic environment, weather and daytime) with task demand and complexity.

An initial examination of the identified studies demonstrated that the state-of-the-art research deals mostly with the effects of road layout, traffic conditions and weather on driver's task demand. On the contrary, limited evidence of studies investigating the relationship between driver's task complexity and time or location was found. As mentioned before, the included studies are concerned with monitoring the effect of contextual information on task demand and are not involved with the effect of road, traffic, time and weather characteristics on road safety. Literature was searched within popular scientific databases such as Scopus, ScienceDirect and Google Scholar. The key words used per factor, as well as the number of screened and included papers are provided in Table 2. Details of the reviewed studies can be found in Annex A.

Table 2: Key words, screened and included papers per factor analysed

Factor	Key words (without word stem variations)	Screened papers	Included papers
Task Demand	"task demand" AND "driving measures" OR "performance measurements" OR "driver characteristics" OR "driving monitoring" OR "workload" OR "traffic conditions" OR "traffic" OR "weather" OR "road layout" OR "time of day"	413	11

The effect of increased mental and physical demand on the driving task, is most frequently measured by physiological indicators and tools such as ECG for heart-related measurements (de Waard and Brookhuis, 1991; Schwarze et al., 2014; Stuiver et al., 2014; Marquart et al., 2015; Stojmenova and Sodnik, 2015; Bongiorno et al., 2017), EEG or functional near infrared spectroscopy (fNIRS) for brain activity measurements (de Waard and Brookhuis, 1991; Stojmenova and Sodnik, 2015; Bongiorno et al., 2017), skin conductance (Stojmenova and Sodnik, 2015; Bongiorno et al., 2017; Foy and Chapman, 2018) or eye tracking measures

(Brookhuis and de Waard, 2010; Benedetto et al., 2011; Auflick, 2015; Marquart et al., 2015; Stojmenova and Sodnik, 2015; Foy and Chapman, 2018). Other successful indicators of additional workload posed on the driver during difficult driving tasks include vehicle kinematics (de Waard et al., 2008; Auflick, 2015; Foy and Chapman, 2018).

In order to quantify the effects, researchers mostly conduct a driving simulator experiment, while only two of the eleven studies also tested their research questions on an open-field driving experiment (de Waard and Brookhuis, 1991; Patten et al., 2006).

With regards to the results of the reviewed studies, it was observed that there is a decrease in heart rate when traffic is dense under adverse weather e.g. fog; (Stuiver et al., 2014), which is also evident when transitioning to a motorway from urban traffic (de Waard and Brookhuis, 1991). Moreover, systolic blood pressure variability leads to a stronger increase of workload with the appearance of fog, especially in the high traffic conditions. An increase in heart rate has been documented during lane changing events (de Waard et al., 2008) and when drivers join the urban traffic from a quiet motorway (de Waard and Brookhuis, 1991). Additionally, an increase in HGV vehicle composition was found to increase mental effort of drivers and led to larger speed variation and shorter time safety margins (de Waard et al., 2008). In Marquart et al. (2015), eye blink rates decreases with sharper road curves, as the driving task becomes more demanding. It was found that a shorter blink duration increases both mental and visual task demands and blink rate. Furthermore, PERCLOS, fixation duration and pupil dilation also increase for all cognitive tasks during hazardous moments such as listening, talking, or calculating, indicating increased mental workload compared to the control condition.

In addition, it was identified that road geometry or traffic contribute to mental workload functions. For instance, when road geometry is more complicated, drivers become more stressful and they are required to have more concentration on the driving task (Bongiorno et al., 2017). Also, an increase of visibility leads to Galvanic Skin Response (GSR) increase. According to Schwarze et al. (2014), who investigated the driving difficulties for elderly drivers, darkness and rain increase the mental workload and it was shown that a higher workload is required in difficult weather situations for both age groups. Furthermore, increases in subjective ratings of mental workload caused by changes in road type were accompanied by increases in skin conductance, acceleration signatures and horizontal spread. Such changes were also associated with increases in the concentration of oxygenated haemoglobin in the prefrontal cortex (Foy and Chapman, 2018). The main disadvantage of the aforementioned studies was that there are no thresholds given for detecting a significant effect of context on the difficulty level of the driving task and that most of the studies were concerned only with car drivers.

3.2 Task demand as an indirect result of exogenous factors

Task demand systematically identifies variables which influence the level of individual effort in a given traffic scenario. Indicators for exogenous factors[5], identified as relevant for the overall driver state and thus risk level, are reviewed and compared in a systematic way. The majority of the reviewed studies in this section present an indirect effect on task demand as they were mostly concerned with accident frequency, risky driving behaviour or the probability of an accident. The assumption is that infrastructure, traffic or external conditions leading to decreased safety levels, play a negative role with regards to task demand as well. Studies were considered to describe a negative effect on task demand if risk or difficulties were found to increase, and a positive effect if a decrease of these factors was observed. Detailed results

[5] Exogenous factors are ones that emerge outside of the vehicle, thus, not related to the driver or the vehicle but the driving envorionment

on the effect of road layout, weather and traffic conditions, as well as time on task demand are presented in the following sections. Details on the reviewed studies can be found in Annex B.

4.2.1 Road layout

With regards to road layout, it was observed that there appears to be an increase in task demand when the number of lanes is higher (Chenqye et al., 2013; Rangel et al., 2013). Furthermore, narrow lanes (Russo et al., 2014; Da Costa et al., 2018) and wider lanes with high traffic volumes (Rangel et al., 2013) increase task demand and subsequently increase crash risk, accidents, injuries or fatalities. Moreover, it was found that deceleration lane lengths (Chen et al., 2009) and spirals, highway curves or geometric design (Zegeer et al., 1990) make driving performance more difficult, increasing the risk and frequency of crashes. Likewise, minor right-turn lanes, main and secondary roads or motorways increase the task demand and risk or crash frequency (Bergel et al., 2011; Pulugurtha and Nujjetty, 2011). According to Valent et al. (2002) driving in a provincial or state road within an urban area results in increased task demand and leads to an intense anxiety for risk of fatal and non-fatal injuries. Similarly, driving in a major artery road for heavy tractor trailers and also in a primary state arterial road (Stephan and Newstead, 2014) requires greater concentration (Blower et al., 1993).

4.2.2 Traffic

Through-traffic per lane on minor roads (Guo et al., 2010) was found to decrease the task demand. All studies indicated that congestion is associated with increased driving difficulty and crash occurrence. In particular, Shi et al. (2016) found congestion to be detrimental for crash frequency during peak hours. Furthermore, it was revealed that congestion, and to a lesser degree transition, increased the driving complexity and therefore the odds of a crash (Zheng, 2012). According to Golob et al. (2008), when the entire road is congested, there was an increase of task demand because crashes were more likely to be caused. Finally, Wang et al. (2013) showed that a 1% increase in traffic delay per kilometre, where vehicles move very slowly, increased KSI[6] crashes by about 0.1%.

Trains regularly share tracks with other trains, with points where tracks have to be crossed, increasing task demand. Train drivers need to attend to signal information, clearing them to cross or proceed along tracks. Therefore, passing signals at danger can increase the risk of collisions. Although trains do not have to contend with traffic in the same way as road transport, trains still have to share the tracks with other trains, sometimes overtaking slower moving trains. Trams on the other hand often share the road with multiple road users, navigating differing road layouts and traffic, all of which increasing task demand.

4.2.3 Weather

Weather conditions have been found to have a significant effect on task demand. For example, rain intensity or duration (Brijs et al., 2008) as well as rainfall height (Fridstrøm and Ingebrigtsen, 1991, Elvik et al., 2013) increase driving complexity. Martensen et al. (2016) investigated the effect of frost and found a noticeable increase of task demand on motorways by a percentage of 71%. Rainfall has been associated with a higher accident risk, which appears to be confirmed by the higher number of victims among car occupants under rainy conditions. It must be noted however, that among motorcyclists all risk factors associated with rain such as bad vision, visibility, or friction, are applied even more strongly than for car users. In addition, snow was found to be connected with a decrease in the total number of injuries or fatal crashes and crashes involving a two-wheeler and car crashes, but it was found that the

[6] Killed and seriously injured

impact of snow was not significant in other cases. Likewise, fog has been revealed to negatively affect task demand in Abdel-Aty et al. (2011) and Sabir (2011) where it was indicated that there was an increased likelihood of injury accident risk under fog conditions. In general, precipitation, sun, wind and frost/snow days, had an increase on driving complexity (Martensen et al., 2016).

It is worth mentioning that weather conditions constitute a less important factor for trains and trams, apart from extreme weather such as floods or deep snow which may restrict passage on tracks. Heavy weather like fog, heavy rain or snow could impact task demand in relation to making it difficult to see and monitor signals.

4.2.4 Time of the day

With regards to the effect of time, it was revealed that the dark, the absence of street lighting and twilight influences task difficulty for drivers (Olszewski et al., 2015). Similarly, darkness was shown to increase the task demand and crash risk by 30% in urban areas, 50% in rural areas and 40% in both rural and urban areas (Johansson et al., 2009). In relation to time of the day, higher task complexity and driving risk was found to occur in the early morning hours from 05:00 to 06:00 and in the evening hours from 17:00 to 19:00 for both national and regional roads (Gaca and Kiec, 2013).

It should be noted that time of day was not found to have a strong impact on train drivers. Tracks are not always lit, however, stations are. Darkness may increase task demand generally.

3.3 Perceived task demand

As indicated in the sections above, task complexity and task demand can be measured by various means. It is, however, also something that can be perceived and subjectively assessed. The correct assessment of the task demand is also linked to risk taking. If a driver underestimates the demand of a complex and dense junction for example, in favour of allocating cognitive resources to other tasks such as anticipating upcoming routing decisions, risk can increase. Many decisions in traffic have to be made quickly and not all relevant information may be accessible to the driver or assessed relevant to make a rational and the safest decision. Decisions are often made quickly and intuitively, which is referred to as the 'experiential system' (Slovic et al., 2004) or 'system 1' in Kahneman's famous dual process theory (2011). Similarly, the Somatic Marker Hypothesis emphasizes the biasing role of emotions and feelings on the process of decision making (Reimann & Bechara, 2010).

In this context, the concept of task difficulty homeostasis is noteworthy. It postulates that drivers dynamically maintain the perceived task difficulty within certain boundaries that conform with their corresponding preferences. Perceived task difficulty results from the driver's perceived capability in conjunction with task demand. The main mechanism for adjustment when task difficulty is outside of the preferred margins is reducing or increasing speed (Fuller, 2011).[7] There is recent research that tries to account for perceived capability and demand in warning systems. Wang et al. (2018) aimed at identifying thresholds for lane change warning systems in a simulator study that concurs with the driver's perception rather than with absolute risk assessment. Furthermore, there is some indication that perceived risk can be inferred from physiological measures such as galvanic skin response (Vaa, 2014). However, this does not seem assured knowledge (yet). Moreover, the driver's assessment of task demand does not necessarily correspond with the actual demand. A recent study by Stapel et al. (2019)

[7] i-DREAMS' deliverable 3.1 further explores theoretical driver behaviour models and the link to the Safety Tolerance Zone.

compared objective and subjective task demand of handling an automated system. While drivers unexperienced with the system assessed the demand as equal to conventional driving, performance of a secondary task indicated increased task demand.

3.4 Applicability to other modes

Trains It is evident that train driving combines the need for prolonged sustained attention. In addition, train drivers also have to withstand the monotony of the driving task as well as the monotony of the environment (Dunn and Williamson, 2011). It's a fact that train driving requires expertise and experience and is also influenced and partly determined by other demands, such as weather conditions, speed restrictions, signals, platforms or passengers. Furthermore, train drivers have the complexities of managing speed, responding to signals, sometimes dealing with passengers and station stops, as well as unexpected obstructions on tracks, all of which could be considered high workload and task demand. However, train drivers also face the complexity of stretches of driving that consist of low workload with increased monitoring of speed and signals, requiring alertness and vigilance. Therefore, measures to identify the effect of driving task, for example psychophysiological indicators (ECG, EEG, skin conductance, vehicle kinematics, eye tracking) could/would be relevant for use in train drivers.

General

In the information processing and task performance literature search, the most representative and relatively more developed tools for measuring task demand were found through Galvanic Skin Response (GSR) sensors, in-vehicle information systems, OBD ports, cameras or eye tracking devices. Moreover, electroencephalography (EEG) or electrocardiogram (ECG) were also used, using electrodes to record brain's and heart's activity, respectively. These can be easily applied to other transport modes, i.e. not only for cars but also for trains, buses or trucks. For example, Torsvall and Åkerstedt (1987) measured EEG and EOG changes continuously in train drivers and Keckluno and Åkerstedt (1993) recorded truck drivers' on-going EEG activity during a night or evening of driving. The advantages of EEG and ECG recording methods are they could potentially be used in all modes, are typically non-invasive, and due to drivers being seated for periods of time when driving, distortion from physical movements should be limited. In addition, the Detection Response Task (i.e. a method for assessing the attentional effects of cognitive load in a driving environment; Stojmenova and Sodnik, 2018) was a method which can be transferable to all the transport modes in order to objectively obtain values for the mental workload of a task, or to assess the attentional effects of cognitive load in a driving environment.

3.5 Conclusions on measuring task demand

As demonstrated in the previous subsections, there were three main classifications of measuring task demand: subjective or self-report, performance and physiological measures (O'Donnel and Eggemeier, 1986). According to the conducted literature search, it was revealed that the most frequent measurement of task demand was derived through psychophysiological indicators and tools.

With regards to self-report measures different dimensions of task demand, such as effort, individual differences, operator state and attitude were taken into account. De Waard (1996) claimed that no one was able to provide a more accurate judgement with respect to experienced mental workload than the person concerned. It was found that the Rating Scale for Mental Effort (RSME) and NASA Task Load Index (TLX) were reliable and sensitive indicators for subjective workload measurement.

In addition, performance measures rely on techniques of direct registration of driver ability to perform the driving task at a level considered acceptable and safe, and properly maintain the vehicle on the road without colliding with other road users (da Silva, 2014). It was found that lateral position deviation was one of the most important indicators of deficiencies in task demand and it can be translated by the possibility of the driver leaving the road centreline and getting involved in an accident. Furthermore, the results indicated that driving performance measures such as speed, longitudinal control or reaction time were some of the major indicators considered for assessing driver performance.

Concerning physiological measures, the most reliable indicators were found to be heart rate, heart rate variability, assessment of electrooculogram (EOG), electroencephalogram (EEG), electrocardiogram (ECG), magnetoencephalography (MEG), electrodermal activity (EDA), head movements, evaluation of eye movement or pupil dilation and blood pressure evaluation.

With regards to the i-Dreams project, psychophysiological measurements are preferred in order to detect task demand. The number and duration of eye fixations as well as ECG measures are the most reliable indicators. Steering wheel sensors for the aim of biometric recognition on electrocardiogram (ECG) signals from the driver’s hands and abnormal cardiac health problems can be utilized to that purpose and driving performance measurements can be detected through smartphone-based technologies. A supplementary eye tracking system to detect eye fixations could also prove beneficial for the detection of task demand within the project.

It is worth mentioning that the majority of the studies reviewed, were conducted in driving simulators with limited studies using open field driving experiments with real road conditions within a specific transport mode. In driving simulators, a particular technology, device or navigation system that was connected directly to a specific transport mode has not been used for task demand monitoring. For instance, no technology or product was able to discriminate between cars' or trains' interior. The results obtained in driving simulators are applicable to real-world and on-road driving. Consequently, all methods that were developed from driving simulator experiments in order to measure task demand, are easily transferable to different transport modes. Indeed, technologies and sensors that were examined, are available and relevant for all modes of transport. Da Silva (2014) claimed that driving simulators were identified as the most widely used methodological environment in research and allowed the creation of real conditions without any objective risk. Finally, it was revealed that most of the measurement tools that were used to monitor task demand and complexity were adapted to all vehicles and no technological device or system was found to be adapted to a specific fleet's specifications. This could be very important for the i-DREAMS project, providing flexibility, meaning that the system does not need to be redesigned for each mode of transport.

4 Driver mental state – indicators, methods and technology for real-time monitoring

The term 'mental state' refers to the cognitive state (attention, fatigue, workload) but also the emotional state of a driver. The 'dichotomisation of the mind' into cognitive and emotional components of the mind facilitates operating with the terms. However, this division is not as definitive as it may seem, and many complex interrelations exist. Emotions, for example, can shift our attention and be distracting (Chan & Singhal, 2015; Cunningham & Regan, 2017; Lafont et al., 2018; Zimasa et al., 2019). The drivers' mental state can be classified on a spectrum of arousal states (low, passive, ideal, hyper), which can be measured with the help of physiological indicators (Lohani et al., 2019). Within the i-DREAMS project, the single cognitive or emotional mental states and, as well as the source of an emerging risk, are of importance since interventions in real time or after driving will be targeted at specific triggers. Therefore, the review of the state of the art of measuring the driver's mental state was structured by the distinct topics of attention and distraction, fatigue and sleepiness as well as emotions and stress.

4.1 Attention and distraction

4.1.1 Definitions & measurements

According to Cunningham and Regan (2008), distraction can be defined as "a diversion of attention away from activities critical for safe driving toward a competing activity". Following that definition, the review focused on identifying the ways in which distraction can be monitored during trips or experiments rather than the relationship between road safety and distraction. For example, in Cunningham and Regan (2008) and Papantoniou et al. (2017) critical driving parameters on distraction are explicitly described. Among those parameters lateral and longitudinal control measurements, surrogate safety measures (e.g. reaction times, gap acceptances) and eye or workload measures are deemed to be the most crucial to identify driver distraction. However, in those two frequently cited studies, the aim is describing the effect or distraction on safety performance parameters, without accurately pointing out the monitoring procedure of distinguishing between attentive and distracted driving.

4.1.2 Review of studies

Literature was searched within popular scientific databases such as Scopus, ScienceDirect and Google Scholar. The key words used, as well as the number of screened and included papers are provided in Table 3.

Table 3: Key words, screened and included papers per factor analysed

Factor	Key words (without word stem variations)	Screened papers	Included papers
Distraction	"distraction" OR "distracted" OR "inattention" OR "inattentive" AND "driver monitoring" OR "driver measure"	417	32

When reviewing the 32 identified studies on monitoring driver distraction, following the aforementioned focus areas, an observable distinction was that very few papers considered all the major types of distraction (i.e. visual, cognitive and manual) as described in Cunningham and Regan (2018) and Costa et al. (2019). The majority of the studies (17 out of 32) were concerned solely with visual distraction, probably because of the advances in eye tracking and camera technologies, while 6 were concerned specifically with cognitive distraction and only 4 with manual distraction. With regards to visual distraction, the phenomenon is usually identified through saccades (Costa et al., 2019), glances and blinks (Seppelt et al., 2017; Dumitru et al., 2018; Li and Seignez, 2018; Bakhit et al., 2019; Costa et al., 2019; Kanaan et al., 2019), or general eye position tracking (Hammoud et al., 2008; Botta et al., 2019). Two of the most crucial indicator for detecting distraction and inattention have found to be PERCLOS (percentage of time that the eyelid covers 80% or more of the pupil) and PERLOOK (percentage of time spent not looking ahead during a certain time interval) with a value of more than 35% indicating distraction (Costa et al., 2019). Glance duration has also been demonstrated in most of the studies on visual distraction as an important indicator. In Seppelt et al. (2017) it was found that distracted drivers were identified during near crashes with an average glance duration of 12.39 seconds (SD 8.02), and 9.58 seconds (s.d. 5.08) during crashes, when observations were made 10-25 seconds before incidents. In Botta et al. (2019) 2 seconds was the critical value of glances away from the road, which was also validated in Kanaan et al. (2019).

Regardless of eye metrics, head position monitoring has also been extensively utilized in order to identify general distraction scenarios or has been linked with visual and manual distraction. More specifically, in Huang et al. (2019) head turns longer than 5 seconds in duration are considered a precursor of distraction. Similarly in Hammoud et al. (2008) a head movement of 20 degrees or more to the left or right, has been also linked with distracted driving, which comes in agreement with the thresholds indicated by Ali and Hassan (2018).

A different approach was followed in Botta et al. (2019) and McDonald et al. (2019), where driver kinematics were used post-trip to distinguish between distracted and undistracted drivers. Features utilized for that distinction include the standard deviation of lane offset and the steering quartiles in McDonald et al. (2019) as well as speed, yaw, steering rate values and road geometry in Botta et al. (2019). However, in both studies no thresholds for detecting distraction are mentioned.

Regarding technologies used to monitor distraction in real-time, most of the studies utilized eye trackers or eye movement encoders (e.g. Hammoud et al., 2008; Dumitru et al., 2018; Botta et al., 2019; Costa et al., 2019), or analysed video and image feeds from cameras (Hari and Sankaran, 2017; Ali and Hassan, 2018; Li and Seignez, 2018; Koohestani et al., 2019). Less frequent approaches include EEG (Costa et al., 2019; Khan and Lee, 2019), hand sensors and magnetic glasses (Huang et al., 2019). Table 5.2 describes the indicators, methods and technologies that were developed to monitor real-time distraction and inattention.

4.1.3 Transferability to other modes

Attention monitoring systems, including head, gaze, eye trackers, CAN bus integration with a lane tracker for measurement of latitudinal control performance, dashboard cameras, smartphone applications, wearables, and radars were used to investigate distraction. Non-intrusive methods were strongly preferred for monitoring distraction, and vision-based systems have appeared to be attractive for drivers. It was revealed that all methods, technological devices and systems mentioned above, which measure driver distraction or attention, can be easily transferred to all transport modes, such as cars, trains, buses or trucks and there was no mode-specific technology. Hence, it can be concluded that attention monitoring systems are easily transferrable to all four modes of i-DREAMS.

4.1.4 Conclusions and recommendations

A thorough literature review was carried out in order to identify which indicators can be used to detect, monitor or measure driver's distraction or attention. In most cases, driver distraction was measured in terms of its impact to driver attention, driver behaviour and driver accident risk (Papantoniou et al., 2017). Real time eye tracking systems, radars for physiological measurements and cameras were found to be the most frequent devices to monitor and detect driver's distraction. On the other hand, less frequent approaches included EEG, hand magnetic rings and magnetic glasses.

Among all trackable parameters, lateral and longitudinal control measurements, surrogate safety measures such as reaction times or gap acceptances, and eye or workload measures are deemed to be the most crucial to identify driver distraction. However, the diversity in the measures used, in combination with the diversity in the design of the experiments (i.e. road and traffic factors examined), often complicated the synthesis of the results, especially for less commonly examined distraction factors.

With regards to the i-DREAMS system, it is very important to identify the way of monitoring driver distraction or inattention. A dashboard camera, which can continuously record eye or workload measures should prove beneficial for the project. In addition, advanced driver-assistance systems like Mobileye, which utilize a forward-facing camera and provides warnings for collision prevention and mitigation, as well as a smartphone application which can provide measures such as lateral and longitudinal acceleration can be utilized for surrogate safety measures capturing observed distraction and inattention. In that way, the most crucial parameters indicating distraction can be efficiently captured in real-time. To measure distraction due to mobile phone use, smartphone sensors which detect the movement of the phone can also be beneficial and are non-intrusive

4.2 Fatigue

Fatigue and sleepiness are significant risk factors for road traffic accidents, injuries and deaths within various transport operations (Williamson et al., 2011; Bioulac et al., 2017; Zhang, Yan, Wu, & Qiu, 2014). Driver sleepiness is thought to contribute to approximately 15-30% of road traffic crashes globally (Connor et al., 2002; Horne & Reyner, 1995; Phillip et al., 2014). Certain features such as increased subjective sleepiness, changes to physiological state, performance decrements, reduced alertness and difficulties sustaining attention can be used to categorise fatigue (Williamson, 2007), despite its various definitions. Several factors can impact fatigue and sleepiness, including insufficient or lack of sleep, prolonged wakefulness, circadian rhythm disruptions and sleep disorders (Zhang et al., 2014), as well as time spent on task (Williamson, 2007).

One of the main causes of fatigue related transport incidents are attentional lapses due to insufficient sleep (Philip & Åkerstedt, 2006; Philip et al., 2005; Schwarz et al., 2016). Driver sleepiness results in decrements in performance, increased reaction time, impaired attention and loss of conscious awareness whilst behind the wheel (Williamson et al., 2011). Frequently observed driving impairments due to fatigue and sleepiness also include higher frequency of lane crossings (Hallvig, Anund, Fors, Kecklund & Åkerstedt, 2014), reductions in hazard perception (Smith, Horswill, Chambers & Wetton, 2009) and increased distractibility (Anderson & Horne, 2013). Fatigue and sleepiness impair performance substantially, with research indicating that following 17h of wakefulness (Dawson & Reid, 1997), or two hours of nocturnal driving (Verster, Taillard, Sagaspe, Olivier & Phillip, 2011), driving performance is the equivalent to a BAC of 0.05% For reference, the legal driving limit BAC is 0.08% for the UK and 0.05% in most EU countries and Scotland.

Links between driving incidents and time of day have been highlighted. The highest number of incidents and accidents occur at times when alertness is reduced due to the body's circadian rhythms (Åkerstedt, Connor, Gray & Kecklund, 2008; Connor et al., 2002; Garbarino, Lino, Beelke, Carli, & Ferrillo, 2001; Horne & Reyner, 1995; Milter et al., 1988). The lowest points of these rhythms typically occur between 02:00 and 04:00, and then a smaller dip in the afternoon approximately between 13:00 and 15:00. At these times, the drive to sleep would be the strongest.

It could be considered that professional drivers are more capable of staying alert and vigilant compared to non-professional, private drivers. However, this is not the case (Anund, Ahlström, Fors & Åkerstedt, 2018). Certain aspects of professional driving can be considered risk factors for fatigue, for example, sedentary occupation, restricted seating, long driving hours, irregular shift patterns, and a unique work environment (Bunn. Slavova, Struttmann & Browning, 2005; Chaiard, Deeluea, Suksatit & Songkham, 2019; Öz, Özkan & Lajunen, 2010). Despite professional drivers reporting that they subjectively feel more alert, objective measures such as increased lane crossings indicated greater levels of sleepiness compared to non-professional drivers (Anund et al., 2018).

However, measuring the effects of sleepiness can be difficult. The impact of sleep loss on performance shows large inter-individual differences. Studies have reported large differences between individuals even in known risk groups (Ingre, Åkerstedt, Peters, Anund & Kecklund, 2006). In addition, differences between sleepy and alert drivers can sometimes be very small or non-significant, and sleep research employs a range of different methodologies, making comparisons between studies difficult (Talbot & Filtness, 2017).

4.2.1 Definitions

Within occupational settings and within literature, the terms 'fatigue' and 'sleepiness' can often be used interchangeably, despite the causal factors contributing to state of the driver may differ

(May & Baldwin, 2009). It is likely that the two states are interlinked, which can make them difficult to isolate them from one another. They are also likely influenced by other mental states, such as stress. In the transportation industry the term ‘fatigue’ is generally used, however, within scientific literature, fatigue and sleepiness have distinct definitions. Sleepiness is defined as the physiological urge to fall asleep, usually resulting from sleep loss (Dement & Carskadon, 1982). This can also be referred to as drowsiness or tiredness. However, fatigue can be more difficult to define, despite being a related concept to sleepiness. It has previously been defined as the inability to continue a task which has been going on too long (Bartley & Chute, 1947), and can be due to factors such as monotony, workload (including underload and overload) and task duration (Di Milia et al., 2011). Despite the differences in definitions, sleepiness and fatigue impair driver attention, vigilance and performance.

4.2.2 Measuring fatigue and sleepiness

Within the scientific literature, fatigue is measured in a variety of different ways. It can be assessed in terms of variables associated with sleep habits, including sleep duration, sleep quality, and prior wakefulness, as well as variables related to assessing how sleepy an individual is while driving, such as subjective and observed sleepiness and sleep events. Self-report assessments can measure fatigue either by asking direct questions relating to sleep, or by using an established tool. Subjective tools can be used to measure state sleepiness (how sleepy an individual is at a particular point), or trait sleepiness (how sleepy an individual is in general). Objectively, sleepiness at the wheel can also be studied by measuring and monitoring physical signs of sleepiness. As such, a range of sleepiness detection measures have been developed to assess sleepiness, both physiological measures, vehicle and behavioural measures, each with different strengths and weaknesses.

Brain activity

EEG is often considered the gold standard method of measuring sleepiness in the laboratory or in simulator settings. EEG captures changes in neuronal activity in the brain, indicating levels of sleepiness. EEG can measure sleepiness due to homeostatic and circadian processes passively, continuously and objectively. The focus of EEG analysis is on the dominant brain wave frequencies present in the EEG. The spectral power or strength of the EEG signal can gradually move from the beta band (12-30 Hz), to the alpha band (8-12 Hz), and through the theta band (4-8 Hz) as sleepiness increases (Aeschbach et al., 1997). The delta band (1-4 Hz) follows after falling asleep (Rechtschaffen & Kales, 1968). Microsleep episodes can also be detected through EEG. Previously EEG has been used to detect sleepiness in train drivers (e.g. Torsvall & Åkersted, 1987), bus drivers (e.g. Anund, Fors, Ihlstöm & Kecklund, 2018) in car simulators (e.g. Åkerstedt et al., 2013; Barua, Ahmed, Ahlström & Begum, 2019; Hallvig et al., 2013) and in real-road driving (e.g. Hallvig et al., 2013; 2014; Liang et al., 2019).

However, EEG is usually measured using electrodes fixed to the scalp, which may not be realistic in terms of real time driver state monitoring. Attaching electrodes can take time, as well as being subject to signal loss or loss of electrical contact. More recently, dry electrodes and headbands with embedded electrodes have been developed (e.g. Lees et al., 2018; Wang et al., 2018; Zhang et al., 2018), although these measures may need further validation. EEG also appears not to show the longer-term build-up of chronic sleep restriction (Van Dongen et al., 2003b). EEG can be used to detect short term changes in sleepiness and microsleeps and actual instances of falling asleep, however the latter two occur rather late in terms of sleepiness development, which may be too late in terms of safety and driving. Efforts have been made to develop EEG based sleepiness indicators, however, their predictive value is too low for use in driver sleepiness devices (Åkerstedt et al., 2010).

Ocular measures

Ocular indicators are popular measures used to monitor driver sleepiness. It has been shown that the homeostatic and circadian processes which influences sleepiness cause changes in ocular variables (Ftouni et al., 2013; Jackson et al., 2016). Many ocular measures are associated with sleepiness including blink frequency and duration, eyelid closure, pupil dimension and slow eye movement (Ahlström, Anund, Fors & Åkerstedt, 2018; Aidman et al., 2018; Barua et al., 2019; Cori, Anderson, Soleimanloo, Jackson & Howard, 2019; Filtness et al., 2014; Ftouni et al., 2013).

EOG can be used to detect changes in ocular measures, using electrodes which are attached near the eyes to record electrical activity. This provides a passive, continuous and objective method for recording changes in eye movements and has been shown to correlate with changes in EEG (Torsvall & Åkerstedt, 1987; Ftouni et al., 2013). In terms of driving performance, EOG has been associated with a range of indicators including lane drifting (Åkerstedt et al., 2013; Ingre et al., 2006), and hitting the rumble strip (Anund, Kecklund, Vadeby, Hjälmdahl & Åkerstedt, 2008). Similar to EEG, EOG electrodes can be difficult to use in terms of a real time driver state monitoring device and can be subject to movement and interference and data loss. EOG measures are also not reliably consistent with other EOG measures, even in the same person (Dinges & Grace, 1998).

Cameras can also be used to detect changes in ocular parameters, and typically focus on changes in percentage of eyelid closure (PERCLOS), which has the highest reliability and predictive validity (Sparrow et al., 2019). As it doesn't rely on electrodes, it is also less intrusive and potentially easier to incorporate into a driver state detection monitor. However, in a commercial and operational sense, it also raises issues relating to the recording of individuals and can be met with resistance. There may be the issue of head movements, and poor lighting affecting accuracy. Measures of ocular indicators have also been incorporated into glasses (Ftouni et al., 2013; He et al., 2017; Ma, Gu, Jia, Yao & Chang, 2018), however this requires the driver to tolerate wearing the device, which could interfere with the use of prescription glasses.

Overall, using ocular measures as an indicator of sleepiness is useful to detect high levels of sleepiness (Anderson, Chang, Sullivan, Ronda & Czeisler, 2013), and is one of the more popular measures being incorporated into driver detection technology. In terms of ocular parameters, blink duration and PERCLOS have previously been shown to be the most robust (Cori et al., 2019). However, research indicates that state detection algorithms need to take other factors into account, for example traffic and surrounding road users, and that algorithms that are estimating a drivers state may need to be personalised (Ahlström, Anund, & Kjellman, 2018).

Cardiac measures

Cardiac activity can be affected by sleepiness. Heart rate and heart rate variability (HRV) are indicators that can be used to capture any changes in this activity following sleepiness. Previously, ECG recordings have been used to detect sleepiness, using electrodes attached to the body (e.g. Buendia et al., 2019; Vicente et al., 2016), however the use of electrodes may be difficult in terms of applicability to the real world. Even though research has noted that circadian processes appear to influence cardiac measures systematically (Burgess, Trinder, Kim & Luke, 1997), there have been inconsistent findings in terms of the homeostatic process (Holmes, Burgess & Dawson, 2002; Zhong et al., 2005). There is also the issue of heart rate being affected by other influences, for example stress (Thayer, Ahs, Frerikson, Sollers & Wager, 2012). It may be that in controlled environments HRV predicts impairments due to sleepiness, although a more recent study analysed motorway driving data and found associations between heart rate, HRV and subjective sleepiness (Beundia et al., 2019), indicating promise as a method of sleepiness detection.

Currently, ECG recordings are not widely used in operational settings to measure sleepiness. However, using heart rate signal as a sleepiness indicator may create possibilities of non-intrusive measurement devices being used to detect driver sleepiness. Rather than ECG electrodes, heart rate could be measured by steering wheel sensors, wristbands or seat sensors (Ariansyah, Caruso, Ruscio & Bordegoni, 2018; Balasubramanian & Bhardwaj, 2018; Macias et al., 2013; Wartzek et al., 2011).

It is important that effective pre-processing algorithms are utilised in order to obtain good estimations of HRV indicators and detecting and removing outliers is essential (Lippman et al., 1994), as they may lead to biased results. The choice of spectral transformation method applied to HRV indices may also have a large influence (Clifford & Tarassenko, 2005). Therefore, before applying algorithms for sleepiness detection based on HRV analysis, the impact of different pre-processing methods on sleepiness assessment needs to be considered (Forcolin et al., 2018).

Performance measures

Performance measures can also be a good indicator of sleepiness. Sleepiness can result in lapses of attention, which become more frequent and last for longer as sleepiness increases (Doran, Van Dongen & Dinges, 2001). Vigilance tasks are accurate and sensitive ways to measure impairments in performance, focusing on lapses of attention (Lim & Dinges, 2008). Lapses of attention have also been associated with microsleeps and correlate with PERCLOS (Dinges & Grace, 1998). However, using such tasks such as the Psychomotor Vigilance Test (PVT), can be unrealistic in terms of operational settings, usually requiring the individual to concentrate on the test for approximately 10 minutes, with repeated measures being required. Shorter versions on the PVT have been developed, however this would still require the individual to perform the test for three minutes, which drivers would be unable to do while driving.

Driving performance is a useful indicator of sleepiness. A variety of indicators can be recorded through sensors, providing information relating to acceleration, steering wheel movement, and braking. The benefit of measuring driving performance is that it requires no additional input from the driver and is unobtrusive. Several indicators have been shown to be sensitive to the effects of sleepiness, including lane deviation, speed variability, steering wheel movements and following distance (e.g. Anund et al., 2008; Forsman, Vila, Short, Mott & Van Dongen, 2013; Ingre et al., 2006; Otmani, Pebayle, Roge,& Muzet, 2005). However, in terms of sensitivity, driving performance measures of sleepiness can vary considerably, and they may indicate increased sleepiness when safety is already an issue. Many of these measures can also be impacted by other factors, such as weather and traffic. Despite these limitations, several car manufacturers have developed systems that incorporate several of these measures to detect sleepiness.

Subjective measures

The question of whether individuals are aware of their sleepiness levels has been debated, with laboratory evidence indicating that when asked, drivers can give responses associated with objective sleepiness (e.g. Åkerstedt et al., 2013; Watling et al., 2016b). However, there are questions about whether drivers acknowledge the risks associated with being sleepy while driving (Watling et al., 2016a), and whether if drivers are unprompted, are they as aware of their sleepiness. Drivers including professional drivers also face additional pressures such as schedules or the desire to reach the destination, which may override their awareness of sleepiness. However, in many laboratory and field studies, self-report measures have been used effectively.

The most extensively used self-report measure is the Karolinska Sleepiness Scale (KSS; Åkerstedt & Gillberg, 1990), a 9-point, one dimensional scale ranging from 1 – extremely alert, to 9 – very sleepy, great effort to stay awake, fighting sleep. The KSS has been validated against EEG variables and performance (Kaida et al., 2006; Sagaspe et al., 2008), and is considered a reliable tool for evaluating sleepiness, both in a laboratory environment and in field studies (Åkerstedt, Anund, Axelsson & Kecklund, 2014). However, there is the question of what level is acceptable to drive at? There is an exponential relationship between the KSS and physiological/behavioural measures (e.g. eye strain, slow eye movements, blink durations and lateral control). KSS scores of less than 7 show little signs of sleepiness in physiological and behavioural indicators, whereas level 8 and particularly level 9 show a strong increase of occurrence in these measures, with KSS 8 and 9 being related to crash risk and instances of sleep intrusions as shown by EEG and EOG (Åkerstedt et al., 2014). In terms of driving and safety, this is too late and driving at this level is dangerous.

Situational context may also be important in terms of subjective sleepiness. Ratings of sleepiness can vary depending on the preceding context (Åkerstedt et al., 2014). While time on task can induce sleepiness, a boring, low stimulus environment can also increase subjective sleepiness, for example a train driver driving through long stretches of the same landscape (Ingre et al., 2004). Manual work may result in lower reported sleepiness levels due to the physical nature of the work, compared to more sedentary, monotonous work, for example in the transportation industry.

The KSS shows good correspondence with high levels of sleepiness as well as with performance and physiological measures at group level, however this may not be the case at an individual level (Sparrow et al., 2016). In terms of utilising subjective scales in an operational context, this can be difficult, as subjective responses can be influenced and manipulated due to social pressure and demand characteristics.

4.2.3 Review of studies

A systematic search of the scientific literature was conducted in July 2019. The search was conducted in three databases (*Web of Science, SCOPUS, and PubMed*). Search terms were generated for each of the key search phrases – fatigue (related to both performance and physiological fatigue), the mode of transport, and the study design. The exact search terms used are shown in Table 4.

Table 4: Terms used for systematic search of the literature

Key search phrase	Search terms
Fatigue	"fatigue*" OR "sleep*" OR "tired*" OR "drowsy" OR "drowsiness" OR "alert*" OR "monoton*" OR "mental* fatigue*" OR "weariness" OR "bored*" OR "time on task" OR "mental* tired*"
Performance measures	"driving ability*" OR "driving behav*" OR "lane crossing*" OR "lane maintenance" OR "lane deviat*" OR "steer* movement*" OR "steering wheel variability" OR "speed*" OR "decision making" OR "situational aware*" OR "miss* traffic signal*" OR "miss* check*" OR "longitudinal move*" OR "lateral move*" OR "event detect*" OR "SPAD*" OR "subjective sleep*" OR "collision avoid* warning system*" OR "pedal use" OR "violation" OR "secondary task engagement" OR "eyes off road" OR "eyes off target" OR "braking" OR "harsh braking" OR "headway"

Physiological measures	"blink rate" OR "heart rate" OR "EEG" OR "eye* clos*" OR "PERCLOS" OR "yawn*" OR "head nod*" OR "eye move*" OR "heart rate variability" OR "ECG" OR "eye track*" OR "pulse" OR "galvanic skin respon*" OR "fNIRS" OR "EMG" OR "respiratory" OR "blood pressure" OR "skin conduct*" OR "cortical activity" OR "biochemical mark*" OR "driver monitoring" OR "driver state monitoring" OR "electrodermal activity"
Mode of transport	"car" OR "bus" OR "coach" OR "truck*" OR "lorry" OR "train" OR "tram" OR "rail*" OR "driver" OR "professional driver" OR "commercial driver" OR "vehicle" OR "automobile"
Study design	"simulat*" OR "real-world driv*" OR "instrumented vehicle*" OR "natural* driving stud*" OR "field operational test*" OR "field operational trail*"

The key terms were then entered into the databases in combination: "Fatigue" AND ("Performance" Or "Physiological") And "Mode" AND "Study", with the following inclusion criteria:

- Published between 2009-2019
- Search term included in title, abstract or key words
- Language as English
- Document type as journal or review
- Source type journals

The search was conducted in the three databases, the results were downloaded into the reference manager Mendeley, and deduplicated. The results were then screened by title, and then by abstract. Additional key references were also included. Table 5 details the screening process and number of hits.

Table 5: Final systematic search figures

Search	**Documents**
Web of Science	772
SCOPUS	770
PubMed	3
Combined	1545
After deduplication	1063
After initial screening by title	282
After abstract screening	177
Additional key references	213
Final screening	**144**

Due to the volume of the literature, this is not a comprehensive review of the literature, but rather a summary of how fatigue has been, and is currently being, measured and/or monitored in relation to driving. A summary table of relevant literature is available in the annex.

4.2.4 Conclusion for measuring fatigue and sleepiness in i-DREAMS

Overall, the majority of studies within the literature focus on 'sleepiness'; that is when a driver is sleepy, what indicators can be used to detect, monitor, or measure this state, and usually in relation to car drivers. In addition, most of the research has been conducted in simulators. This is mainly due to the ethical issue of driving on real roads with sleepy drivers and the associated safety risks, as well as the control factor in experimental terms. However, it should be noted that simulator studies are associated with higher levels of subjective and physiological sleepiness in comparison to real driving (Fors, Ahlström & Anund, 2018). The relative validity of simulators is acceptable for many variables, however in absolute terms, there is an increase in sleepiness levels in simulators (Hallvig et al., 2013).

In terms of measurements and indicators suitable for measuring fatigue and sleepiness in a driver context, the literature provides some useful insights. HRV shows promise and can be developed into unobtrusive measures, however this method hasn't been used much in operational settings and may need further development and validation. Heart rate and HRV can also be influenced by stress, and individuals have been shown to have different reactions to sleepiness as measured by HRV (Abtahi et al., 2017). The majority of commercial fatigue detection technology currently utilises ocular measures, and research shows blink duration and PERCLOS to be the most robust indicators. It is important that in terms of driver state detection and estimation algorithms, other contextual factors such as traffic and other road users are taken into consideration, and that the algorithms may need to be personalised to the driver (Ahlström, Anund, & Kjellman, 2018). Similar to EEG, ocular measures may also be limited in potential for lower levels of sleepiness. EEG can be difficult and unrealistic to implement in terms of driver state monitoring, much like vigilance performance tasks, which require driver involvement and attention, and are not suitable to be used whilst driving. Finally, measures of driving performance offer high validity in an operational context, however indicators can be influenced by other factors, and may indicate levels of sleepiness too late, that is when an individual is already in an unsafe state to drive.

Most measures appear to be sensitive to variations in levels of wakefulness, however they all suffer from limitations, for example inter-individual differences (Ingre et al., 2006), as well as being influenced by external factors possibly not related to sleepiness. Therefore, it is important to ensure that the i-DREAMS system includes a context component in addition to the monitoring component.

In terms of individual differences, research has shown variability among individuals as to the effects of sleepiness, which potentially has implications for fatigue and sleepiness monitoring devices. Assessing changes of individuals over time may not be too problematic, however issues may become apparent if a measure is being used to compare against a set threshold (Van Dongen & Belenky, 2012), or between individuals.

It was apparent from the literature that several studies used a combination of measures and indicators to detect sleepiness. Sleepiness is a complex, multidimensional state, and therefore it may be difficult to assess by one single indicator; it has previously been noted that the reliability of detection systems may be improved by combining and incorporating multiple measures (Balkin et al., 2011). This therefore should be taken into consideration when designing the fatigue and sleepiness detection component of the i-DREAMS system.

4.2.5 Conclusions and recommendations

In relation to suitable measures of sleepiness to be incorporated into the i-DREAMS system, it is recommended that two measures are utilised. The literature indicates that all measures currently used to monitor and detect sleepiness have benefits and drawbacks to them, and that further research or validation may need to be conducted, particularly in relation to inter-individual variability. Utilising multiple measures and indicators could help to improve the reliability of sleepiness detection.

It is apparent from the literature, that eye tracking is the most commonly used measure of sleepiness in experimental, and also commercial, sleepiness detection. Heart rate and heart rate variability also shows potential and can be developed as a minimally invasive technique. Therefore, of the reviewed measures and indicators, it is recommended that measures of eye tracking (blink rate and PERCLOS have been shown to be the most robust) and heart rate are incorporated into the i-DREAMS system to potentially monitor and detect sleepiness.

4.3 Emotions and related constructs

Other than the topic of distraction, emotion research has long been neglected in road safety as well as in human factor engineering and is only trending since the early 2000s (Jeon, 2017). However, technology aiming at the detection of emotions is progressing rapidly. The tools are becoming simpler and more widespread. Eichhorn and Pilgerstorfer (2017) come to the conclusion that drivers experiencing emotions (e.g. anger) while driving are slightly more at risk in traffic, which is supporting the relevance of emotions in the road safety domain.

However, the matter is more complicated than that. First of all, “emotion”, “arousal”, “mood”, “affect” or “stress” are distinct constructs (although there are overlaps) but are often used as one single construct which is seen as complementary and opposite to cognition (Jeon, 2017). And secondly, there is no agreed upon standard definition of “emotions” as a psychological construct.

4.3.1 Definitions

In psychology, ‘emotion’ is often used as an umbrella term for an integrated, complex reaction of the organism to a situation or stimulus which can be interpreted as positive or negative (Maderthaner, 2008). Damasio (2001) sees emotions as the physiological response of the nervous system to a stimulus from inside or outside of the body. Opposed to this rather brief physiological reaction, ‘feeling’ is the subjective interpretation and experience of that reaction. The concept of ‘affect’ is a short-term and undifferentiated emotion, while ‘mood’ can be seen as longer lasting but weaker or less pronounced (Maderthaner, 2008).

Constructs that are clearly different from emotions and feelings are ‘stress’ and ‘aggression’. However, at times those constructs are used in the context of emotion research without paying too much attention to clear-cut definitions. Consequently, aggression and stress were also included in this work. ‘Aggression’ can be defined as the motivation to physiologically or psychologically harm others (or objects) and is a reaction to frustration or anger (Maderthaner, 2008; Shinar, 1998). According to Butler (1993), ‘stress’ is either a result of pressure, a response to noxious or aversive stimuli or a dynamic process.

What many researchers agree upon, is that emotions basically have two qualities: valence, which indicates whether an emotion is perceived as positive or negative, and arousal, which indicates how calming or exiting the stimulus is perceived (Eichhorn & Pilgerstorfer, 2017).

The ‘circumplex model of affect’ hypothesizes that two different neurophysiological systems are responsible for different neural sensations which are cognitively interpreted. One neurophysiological system is related to valence and the other to arousal. Each interpreted emotion can be located in a two-dimensional circular space representing arousal (vertical axis) and valence (horizontal axis) as depicted in Figure 3. For example, fear or anxiety are positioned in the top left quadrant with high arousal and negative valence. The four broad categories (quadrants in the figure), however, do not have very clear boundaries. The combination of high arousal and high valence can be called ‘excitement’, low arousal and high valence ‘serenity’, high arousal and low valence ‘distress’ and low arousal and low valence ‘depression’ (Widen & Russell, 2008). A similar two-dimensional concept was introduced by Watson and Tellegen (1985).

A second prominent – and somewhat competing – school of emotion theories is a categorial approach to emotions as opposed to the aforementioned dimensional one. Ekman (1992, quoted by Balters & Steinert, 2015) proposed ‘basic emotions’, a set of categories of distinct emotions which are seen as universal. These basic emotions are happiness, surprise, anger, sadness, disgust and fear. While either of the two approaches can be argued for, it depends on the research question as to which is 'the right one'.

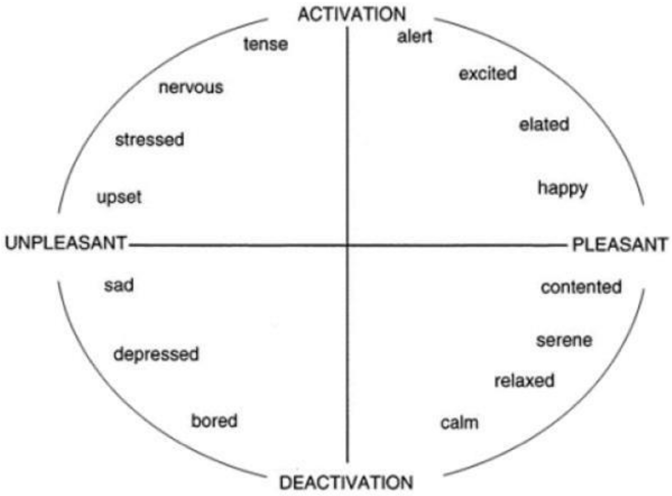

Figure 3: 'Circumplex model of affect' with the horizontal axis representing valence and the vertical representing arousal (Posner et al., 2005, p.716)

Lagarde et al. (2004) researched the impact of 'emotional stress' due to separation or divorce in a cohort study and conclude that these kind of life events are associated with an increase in serious road accidents. Another study similarly found that stress increases at-fault crash risk significantly. Both personal stress triggers (family finances, sickness, work stress etc.) as well as driving context induced stress (traffic density, weather condition, road layout etc.) were considered (Legree et al., 2003). Thus, stress can be induced from non-driving related events and driving itself. Stress is, furthermore, associated with other discussed constructs such as facilitating aggression (e.g. Maderthaner, 2008). Furthermore, time pressure is a common trigger of stress, which can be especially relevant for professional drivers.

4.3.2 Measuring emotions

On a broad level, three classes of emotion measures can be distinguished: subjective measures where individuals assess and communicate their own emotional state, behavioural measures such as facial expressions or posture and physiological measures which are assessing changes in the autonomous nervous system (Balters & Steinert, 2015). Since i-DREAMS aims at objectively measuring the individual mental state while driving in real time, subjective measures are not further explored.

Regarding physiological measures, the autonomous nervous system plays a significant role since emotional reactions have autonomic specificity (distinct change in the autonomous nervous system). There is no consensus, however, on autonomic specificity of categories of emotions (e.g. fear or happiness). Consequently, a universal and reliable concept of interpreting physiological signals is missing. Data collection should be based on clear definitions of emotion or categories of emotions and should be interpreted with care (Balters & Steinert, 2015). Among the behavioural measures, voice characteristics such as amplitude and pitch can be used to detect emotions. Furthermore, facial expression features and body posture are used to infer emotional states.

4.3.3 Impact of emotions on driving performance

Emotions can have detrimental as well as beneficial effects on the driver and the driving task. While fear is to a certain extent a prerequisite to adapt one's own driving behaviour to a manageable scale (see also risk control theory by Summala, 1988), usually emotions are studied in view of the potential increase of risk. Some of the evidence regarding risk increasing the impact of emotions is discussed below.

In the large-scale naturalistic driving study SHRP 2, Dingus et al. (2016) observed a 9.8 times higher crash risk while driving in an emotionally elevated state (including anger, sadness, crying, and/or emotional agitation) compared to normal driving (prevalence of about 0.2% of the total driving time). Studies based on the self-reported accident history of drivers show inconsistent results. Regarding the influence on safety performance indicators, Cyders & Smith (2008) find that emotions and risky behaviours correlate with drivers being more prone to rash actions.

Negative emotions in general (e.g. anger, fear) are assumed to potentially impair drivers, for example through a decrease of the available attentional resources due to an increase of emotion related thoughts (Ellis & Moore, 1999). Then again, positive emotions such as happiness or joy are more frequently discussed in view of their impact on driving behaviour as well (Cunningham & Regan, 2017).

A meta-analysis by Nesbit et al., 2007 indicated a relationship between aggressive driving and accidents as well as between anger and accidents, however, the effect is rather small. However, there are also studies reporting no association of aggression and anger with crash risk (Eichhorn & Pilgerstorfer, 2017). Anger seems a well-studied construct that is associated with speeding and other risk behaviours (Eichhorn & Pilgerstorfer, 2007). Furthermore, anger is associated with a more superficial processing of potential hazardous information and a longer time for corrective actions (Stephens et al., 2013).

Similarly, attentional decrements due to anxiety were identified in previous research. High levels of state anxiety were found to cause distraction by irrelevant stimuli and attentional narrowing (Gotardi et al., 2018; Jeon et al., 2015).

4.3.4 Review of studies

In order to draw conclusions on the state-of-the-art emotion measurement techniques, a systematic search of the scientific literature was carried out mid 2019 using the search terms documented in Table 6.

Table 6: Search terms for systematic literature search on measuring emotions.

Key terms	Search terms
Emotion	"emotion*" OR "ang*" OR "rage*" OR "upset" OR "worr*" OR "self-regulat*" OR "anxi*" OR "panic*" OR "happ*" OR "excite*" OR "aggressi*" OR "arous*" OR "pressure" OR "affect*" OR "feel*" OR "fear" OR "nervous" OR "stress*" OR "*joy*" OR "pleasure" OR "amuse*" OR "grief" OR "surprise*" OR "sad*" OR "euphor*" OR "distress*" OR "depress*"
Performance measures	"perform*" OR "driv* ability*" OR "driv* behavio*" OR "capability" OR "lane deviation" OR "lane keeping" OR "steering*" OR "speed*" OR "decision*" OR "situational awareness" OR "reacti*" OR "longitudinal*" OR "lateral*" OR "*detecti*" OR "SPAD*" OR "violation" OR "hazard*" OR "incident*" OR "error*" OR "crash*" OR "accident*" OR "near crash" OR "near miss" OR "collision*" OR "critical" OR "task*" OR "critical event" OR "TTC" OR "time to collision" OR "safety gap" OR "tailgat*" OR "time headway"
Physiological measures	"blink*" OR "heart rate*" OR "EEG" OR "eye movement" OR "ECG" OR "eye track*" OR "eye-track*" OR "pulse" OR "galvanic skin response" OR "skin conductance*" OR "*NIR*" OR "respirat*" OR "EMG" OR "blood pressure" OR "cortical*" OR "biochemical mark*" OR "bio mark*" OR "monitoring" OR "driver state" OR "mental*" OR "workload" OR "video*" OR "respir*" OR "temperature"

Mode of transport	"car" OR "bus" OR "coach" OR "truck*" OR "lorry" OR "train" OR "tram" OR "rail*" OR "driver" OR "professional driver" OR "commercial driver" OR "vehicle" OR "automobile"
Study design	"simulat*" OR "real-world driv*" OR "instrumented vehicle*" OR "natural* driving stud*" OR "field operational test*" OR "field operational trail*" OR "on road" OR "FOT"

The key terms were then entered into the databases in combination: "Emotion" AND ("Performance" Or "Physiological") AND "Mode" AND "Study", with the following inclusion criteria:

- Published between 2005-2019
- Search term included in title, abstract or key words
- Language as English or German
- Document type as journal or review
- Source type journals

The search was conducted in the databases Scopus and Google Scholar. Publications were deduplicated, screened by title (403 publications) and then by abstract. Although the limitation was set to publications after 2005, six papers which met the criteria had to be excluded due to outdated methods. On the other hand, other publications from similar years were included as they appeared relevant. Additional key references were also added. Eventually, 30 publications were screened thoroughly.

While there is a range of studies investigating the impact of emotions on road safety, this literature search and review explicitly focuses on research relating to objectively measuring emotions, preferably in the context of road safety. A summary table of the reviewed studies can be found in Annex E, including an overview of the constructs measured (e.g. fear, stress, negative emotions etc.), the measuring method(s), the study design, main results and our conclusions, especially in view of the next steps in the i-DREAMS project.

Table 7 represents counts of reviewed studies which reported on measuring a certain emotion category or qualitative aspect of emotions. While (emotional) stress is not an emotion by definition, this construct was included in the search and was most often one of the research topics followed by anxiety or fear and anger, aggression and frustration. The relative importance of those constructs is also reflected in the outcome of recent work by Bosch et al. (2019), who had experts in the automotive industry assess the type of emotions that will be most relevant to detect in vehicles. They found that anger, stress and fear are among the most relevant emotions for the industry.

Table 7: Number of studies reviewed per measured emotion or related construct

Anger, frustration, aggression	(Emotional) stress	Anxiety, fear	Arousal, valence; pos./neg. emotions; affective state	Happiness, euphoria, amusement	Sadness, disappoint-ment	Disgust	Surprise
13	9	7	6	6	6	2	1

Table 8 shows the number of reviewed studies which used a certain measurement method or indicator to determine emotions or similar constructs. Since the table shows a mere quantification, no conclusion for the applicability in i-DREAMS can be drawn from it. However, a tendency of methods used in recent years becomes apparent. Heart related measures (pulse, heart rate, inter-beat interval etc.) and measures using electrodermal activity (using 'skin conductance', also called 'galvanic skin response') were used the most often within the

30 reviewed studies. The top four categories are all physiological measures. The first behaviour measure category is only ranked fifth. However, not all of the publications stem from a driving context. Most of the studies used more than one method to capture the emotion or related construct.

Table 8: Number of studies reviewed that used measurement method

ECG, other cardiac meas.	EDA (SC, SCL)	EEG	Skin temp, NIR	Ocular meas.	Facial expres-sion	EMG	Respir-ation	Speech recog-nition	Blood pressure
17	15	7	7	4	4	3	3	2	1

It has to be noted that the majority of the 30 studies were conducted within the realm of road safety research, a few in the context of emotion research per se or Human Machine Interaction (HMI) other than car manufacturing. Furthermore, most of the road safety studies report driving simulator experiments. The experimental laboratory setting of a simulator allows for great flexibility compared to naturalistic driving studies, when it comes to applying measurement equipment. Speech recognition, for example, will not be a useful technique since in a naturalistic setting, participants cannot be encouraged to speak during their drives.

4.3.5 Conclusions and recommendations

A systematic search of the scientific literature on measuring emotions was conducted to identify the most valid, reliable and appropriate behavioural and physiological indicators and respective measurement methods. Although measuring emotions in the context of transportation research was the focus, studies from the domains of health and HMI were also included in the search. Regarding emotions while driving, most of the reviewed studies were driving simulator experiments, with limited studies conducted in real-world naturalistic driving scenarios. This result is plausible, given the ethical constraints that come with inducing emotions. Also, a manipulation check is easier to conduct in the controlled environment of a simulator.

The research designs of the studies are very heterogeneous with a broad variety of underlying theoretical assumptions regarding the operationalization of ‘emotions’; some studies did not provide details on their definition of emotion at all. However, anger, frustration, aggression, stress as well as fear and anxiety appeared to be the most frequently studied emotional categories. A two-dimensional approach with the combination of arousal and valence levels is very common and can also be recommended for the i-DREAMS project.

With regards to potential indicators of emotional states, EDA and heart-based measures are most frequently used in the reviewed studies. EEG, as the third leading measure in the studies, can be neglected, since the set-up and support from i-DREAMS staff is assumed to be too time-consuming, especially in the trial stage. Furthermore, the majority of studies described using more than one physiological or behavioural measure.

Although the synthesis of results is complicated due to the broad spectrum of research designs, a few recommendations can be drawn:

- Focus on stress, anger and fear, operationalised with arousal/valence approach since these are currently the best understood constructs in terms of real-time measuring.
- Use of more than one measurement method: complementing the ‘CardioWheel’ (ECG) with an EDA measuring device such as a wristband or a (thermal) camera facing the participant is advised. The complementary method may provide evidence for validity.

- Use of simulator phase to validate constructs including the induction of emotions and a manipulation check through self-assessment of experienced emotions since this appears to be the standard in scientific literature.

4.4 Assessment of technology for measuring the driver's mental state

Devices and equipment used in the reviewed studies measuring attention, fatigue/sleepiness and emotional states were separately reviewed and assessed in terms of intrusiveness, validity and reliability (if available) and overall applicability for i-DREAMS purposes. Some of the devices had to be excluded from the review, for example, if the vendor does no longer operate or the product is not commercially available. Furthermore, EEG equipment was also excluded due to the level of intrusiveness. The details of the review can be found in Annex F. Table 9 summarizes the assessment of available technology, focusing on the theoretical suitability of single devices or technologies for measuring the driver state constructs in question and the applicability in two settings 'simulator' and 'on-road trial'. Intrusiveness is reason for a negative assessment of a device for the on-road setting. The table/review does not consider financial feasibility and no prioritisation of positive assessed devices was made in terms of validity.

Table 9: Overall assessment of devices and technology for measuring driver state, cells highlighted in bright green indicate feasibility to operationalize construct or in setting, red indicates the contrary. Lines are highlighted in dark green if device can be used for all constructs and settings.

Product	Indicator	Constructs A=Attention, distraction F=Fatigue, sleepiness E=Emotions, stress			Simulator +/-	On-road test +/-
BioRadio 150 by Great Lakes NeuroTechnologies	EDA	A	F	E	+	–
BIOPAC Systems for (ECG)	EDA, HRV, Temperature	A	F	E	+	–
Cardio Wheel	HRV	A	F	E	+	+
Empatica E4	EDA	A	F	E	+	+
Eye tracking glasses	Fixations, head position, PERCLOS, blink parameters	A	F	E	+	–
FlexComp from Thought Technology	EDA, HRV	A	F	E	+	–
Optalert	Eye and facial features/position	A	F	E	+	+
Seeing Machines	Eye and head features/ position	A	F	E	+	+
Shimmer 3	HRV	A	F	E	+	–
SMI, Smart Eye Eye tracker	Eye and head features/ position	A	F	E	+	–
Texas instruments biometric steering wheel	HRV	A	F	E	+	+
Vigo, eye tracking	Eye and head features/ position	A	F	E	+	+
Vital jacket	EDA, HRV	A	F	E	+	+
Zephyr BioModule	HRV	A	F	E	+	+

4.5 Monitoring mental state of operators in aviation and maritime

This sub-chapter explores methods for monitoring driver state in transport modes other than those considered in i-DREAMS, namely in the aviation and maritime sectors. This aims at potentially learning from the research conducted specifically for those modes. A literature search was carried out with key search terms concerning the driver mental states targeted in the project[8].

The results showed that real-time monitoring of the operator's mental state is not common in the aviation and maritime sectors. This is not surprising for the maritime sector, as the relatively low speed and density of maritime traffic leaves quite large reaction time margins for the navigating officers; therefore, emphasis is put on alerting the operator for risks in the environment rather than his/her own steering behaviour. In the aviation sector too, approaches for real-time monitoring of pilots' mental state are limited, although automation has been deployed for decades, and human-machine interaction is a considerable part of piloting. The fact that fatigue, stress and anxiety, cognitive disorders and poor situational awareness (largely due to human out-of-the-loop problems related to automation) are persistently among the key contributory factors in aviation crashes (Jones & Endseley, 1996), monitoring pilot's mental state is mostly carried out within standard training, re-training and fitness screening processes by means of medical evaluations, neuropsychological tools, simulator sessions etc.

A few recent studies were found dealing with real-time monitoring pilot mental state, most of them testing new sensors and unobtrusive methods for measuring physiological indicators. Lehrer et al. (2010) compared heart rate metrics with expert ratings of task load and stress in a simulator experiment and found a robust matching of increased cardiac data and self report / expert rating. Dehais et al. (2019) tested a portable six dry-electrode EEG system aiming to distinguish light- from heavy-workload conditions in a real-world scenario and concluded that further steps are required so that the sensors can be fine-tuned for everyday flight operations. A single channel EEG system was tested in-flight for 14 pilots during long-haul flights to (successfully) detect low vigilance states (Sauvet et al., 2014).

Majumder et al. (2019) used a data recording system (MP 160, BIOPAC Systems) in order to record ECG, EEG, and PPG (Photoplethysmogram, a physiological measure, which can detect the changes in peripheral blood volume) signals during a simulated flight, aiming to detect drowsiness. It was found that Pulse Arrival Time (PAT, measured as the difference between ECG and PPG peaks) was positively correlated with drowsiness.

Other wearable and portable brain monitoring sensors such as functional near infrared spectroscopy (fNIRS) have also been tested in order to investigate brain activity in various everyday human tasks. Gateau et al. (2018) tested such a system in various tasks within simulator and real-life flying and found that single-trial working memory load could be accurately classified in both experimental conditions.

Peissl et al. (2018) reviewed 76 studies on the use of eye tracking technologies in aviation for measuring fatigue, spatial disorientation, hypoxia and high workload in general, and concluded that studies consistently show great potential of oculomotor measures for making real-time predictions of several risk factors. Thatcher and Kilingaru (2012) described the architecture for an intelligent software agent to assess a pilot's situation awareness through the observation of eye movements. An interesting aspect of this approach is the fact that the eye movement

[8] The search was carried out in the Scopus database, and concerned title, abstract and keywords search. As combined searches (e.g. <fatigue>OR<drowsiness>OR<alertness>) yielded a large number of irrelevant papers, separate searches were carried out for each mental state factor and each transport sector. For instance: <mental state factor> AND <transport mode> AND <pilot/operator> AND <monitoring>. Articles returned were screened by title and abstract in order to select relevant studies.

behaviour is compared to a behaviour database including data on appropriate (i.e. safe) behaviour. Lounis et al. (2020) present a preliminary evaluation of an embedded system ‘FETA’ that evaluates the visual monitoring of the cockpit online. The system compares the current visual scan of the pilot with a database of “standard” visual circuits created on the basis of eye tracking data from 16 airlines pilots and alerts the pilot through an auditory alarm in case there is a large deviation from the database. Preliminary results were promising regarding situation awareness, visual behaviour and overall performance monitoring, however further validation was deemed necessary before proceeding to operational implementation.

There is consensus in several studies dealing with stress monitoring through heart-rate measurements. Socha et al. (2016) used a Garmin commercial chest-strap sensor together with a heartbeat sensor to monitor the level of stress in eight trainee pilots, in simulator and real flight conditions. Luig & Sontacchi (2014) suggested an innovative system combining speech recognition and heart rate measurements for real-time monitoring of pilot stress; the voices and the heart rates of eight airline pilots were recorded while completing an advanced flight simulation programme and found that several heart rate variability parameters correlate with speech features in stress manifestations.

4.5.1 Conclusions

In summary, there are three main methods for monitoring pilot mental state in aviation: (i) ECG and other heart-rate monitoring techniques are considered very reliable for monitoring workload, drowsiness/fatigue and stress, (ii) eye tracking techniques used to monitor fatigue, drowsiness and situational awareness, (iii) speech recognition databases for monitoring stress. Several studies (e.g. Peissl et al., 2018) suggest that a complementary use of unobtrusive sensors would enhance the reliability of monitoring.

4.6 Substance impairment

Driving under the influence of psychoactive substances is one of the main contributing risk factors to road traffic accidents, fatalities and serious injuries (ETSC, 2017; Schulze, Schumacher, Urmeew, & Auerbach, 2012; Talbot et al., 2016). For Europe, the DRUID project estimated the mean prevalence amongst the general driving population for alcohol, illicit drugs, and medicines. According to the results, alcohol was detected in 3.5% (>0.1 g/L) and 1.5% (> 0.5 g/l) of the drivers, illicit drugs (mainly cannabis) in 1.9% and medicines[9] (mainly benzodiazepines and opioids) in 1.4% of the drivers (Houwing et al., 2011). Furthermore, combinations of drugs and medicines were found in 0.39% and alcohol combined with drugs or medicines in 0.37% of the drivers (Houwing et al., 2011). The prevalence of substance use is even higher among injured and fatally injured drivers. The studies performed in the DRUID project concluded that alcohol was present in 24.4% of the seriously injured and 32.8% of the killed drivers, whereas illicit drugs and medicines were detected in 15.2% and 15.6% respectively (Isalberti et al., 2011; Legrand et al., 2014, 2013; Schulze et al., 2012). These figures indicate that drink driving is not only more prevalent among the general driving population but is also a more prevalent cause of road traffic fatalities and injuries compared to driving under the influence of drugs and medicines.

Driving under the influence of psychoactive substances varies according to driver age and gender. Driving under the influence of alcohol is mainly observed among male drivers aged ≥ 35 years during weekday nights and at weekends whereas illicit drug use is more prevalent among young males aged 15-34 years and at all times of the day but mostly during weekends

[9] It should be taken into account that the DRUID-project only tested for a limited number of medicines namely: benzodiazepines, medicinal opioids and Z-hypnotics (zolpidem or zopiclone).

(European Monitoring Centre for Drugs and Drug Addiction, 2012). Furthermore, driving under the influence of medicines is mainly observed among middle-aged and older female drivers (≥ 35 years) during daytime hours (European Monitoring Centre for Drugs and Drug Addiction, 2012).

Substance use influences the psychological state of the driver and impairs driving ability which results in an increased risk of being involved in accidents (Elvik, 2013). As indicated in Table 10, the effect on driving ability varies with the substance type. Substances can be classified into three categories: depressants, stimulants, and hallucinogens. The intake of narcotic substances or depressants has a sedative effect on the user resulting in reduced inhibitions, concentration and overestimation of driving skills (Arnedt, Wilde, Munt, & MacLean, 2001; Marillier & Verstraete, 2019). This results in increased reaction times, higher variability in driving speed, drowsiness and swerving. Stimulants have a different impact on driving skills as they provide an energetic and alert feeling. Drivers become overconfident, adopt a more aggressive driving style and take more risks whereas their vehicle controlling abilities are reduced (Marillier & Verstraete, 2019; Shinar, 2006). Substances with a hallucinogenic effect create feelings of euphoria, relaxation, and drowsiness. The impact on driving skills is reflected by increased reaction times and impaired coordination leading to poor execution of complex driving tasks (i.e. tasks for which the attention has to be divided over various tasks) (Marillier & Verstraete, 2019).

Table 10: Influence of substance use on driving performance (adopted from Marillier & Verstraete, 2019)

Type of substance	Side effects		
	Acute	Chronic	Combination
Depressants (alcohol, benzodiazepines, Z-hypnotics, heroin)	Slow reaction time, poor judgment, impaired vision and hearing, poor coordination and a false sense of confidence, risk-taking behaviour. Signs of impairment are erratic driving (weaving, swerving, ignoring road signs), confusion, slowed reaction time, increased risk-taking, unresponsiveness, tunnel vision, lack of balance, unsteady coordination and varying states of wakefulness	Slowed reflexes, slower information processing, significant loss of cognitive functions. When drinking or taking drugs regularly one will be able to drink larger quantities before feeling or appearing intoxicated (tolerance)	The sedative effects of both alcohol and sedative medications can enhance each other. Various combinations definitely impair the driving performance.
Stimulants (methamphetamine, amphetamine, cocaine)	Lack of coordination, sensory disturbances, disorientation, restlessness, lapses of attention, difficulty reacting appropriately to safely control a vehicle, increased risk-taking, overconfidence in driving skills, drowsiness or rebound fatigue (as the effects wear off). At low concentrations the amphetamines can improve attention etc., at higher doses they cause impairment. They cause significant impairment during high fatigue states after the “high”	Defects in cognitive functions, increased impulsivity and depression.	Alcohol reinforces negative effects and can cause some additional defects. Some effects of alcohol can be diminished.
Hallucinogens (LSD, cannabinoids)	Visual or auditory hallucinations, a feeling of not being in control, or of	After repeated use, users need increasingly larger doses to produce similar effects.	Additive or even synergistic relation with negative effects of alcohol.

	being disconnected from reality, concentration loss, becoming over-confident on the road, more likely to take risks, impaired coordination	Long-term effects of some LSD users include sudden flashbacks, recurrence of certain aspects of a person's experience without the user having taken the drug again	

Since substance use impairs driver performance, it also leads to higher risks of being involved in severe accidents and being seriously injured or killed compared to unimpaired drivers. As illustrated in Table 11, the risk level varies according to the type of substance. The highest risk is associated with a high blood alcohol concentration or using other psychoactive substances in conjunction with alcohol, followed by medium alcohol concentrations, multiple medication use and driving with amphetamines (Schulze et al., 2012). Alcohol concentrations between 0.5 and 0.8 g/L, medicinal opioids, illicit opiates and benzodiazepines, and Z-drugs are associated with a medium increased risk. Nevertheless, accident and injury risk are significantly increased when multiple psychoactive substances are used simultaneously (Hels, et al., 2011; Schulze et al., 2012; SWOV, 2015). Furthermore, it should be mentioned that the results from the DRUID project revealed that driving under the influence of alcohol is the largest road safety problem in nearly all European countries resulting in more road casualties every year compared to driving under the influence of drugs and medicines (Schulze et al., 2012).

Table 11: Relative risk level of involvement in severe accidents and being seriously injured or killed compared to unimpaired drivers (adopted from Schulze et al., 2012)

Risk Level	Relative risk	Substance group
Slightly increased risk	1–3	0.1 g/l ≤ alcohol in blood < 0.5 g/l Cannabis
Medium increased risk	2–10	0.5 g/l ≤ alcohol in blood < 0.8 g/l Benzoylecgonine Cocaine Illicit opioids Benzodiazepines and z-drugs Medicinal opioids
Highly increased risk	5–30	0.8 g/l ≤ alcohol in blood < 1.2 g/l Amphetamines Multiple drugs
Extremely increased risk	20–200	Alcohol in blood ≥ 1.2 g/l Alcohol in combination with drugs
Notes: Cannabis and amphetamines: owing to very different single-country estimates, the risk estimates must be treated with caution. Benzoylecgonine, cocaine and illicit opioids: owing to few positive cases and controls, the risk estimates must be treated with caution.		

It is also worthwhile mentioning that driving under the influence of drugs and medicines is not as well understood as drink driving (ETSC, 2017; Schulze et al., 2012). For alcohol, a clear link is established between blood concentration and accident and injury risk whereas this is not the case for the other substances due to the variety of substances and their varied effects (Schulze et al., 2012). For example, depending on the type of drugs or medicines, a small dose can already result in extreme high-risk levels whereas, for some other types, quantities significantly above the therapeutic range are required to result in extreme high accident and injury risks (Schulze et al., 2012).

4.6.1 Definitions

Substance impairment in this deliverable focuses on alcohol, drugs, and medicines. These substances are often denoted as psychoactive substances because their consumption causes a physiological change in the body (Stedman’s Medical Dictionary, 2005). Because of this psychoactive effect, these substances affect the behaviour, mood, senses, consciousness, and perception of the user, which in turn can have a negative impact on driving skills and higher accident involvement (Talbot et al., 2016). Therefore, driving under the influence of alcohol, drugs, and medicines are regarded as a risk to road safety. Psychoactive substances can be classified into two categories: licit and illicit substances. Some substances that impair driving performance are licit. Examples of licit (addictive) substances are caffeine, nicotine, and alcohol. Alcohol is the most prevalent psychoactive substance among drivers (ETSC, 2017; European Monitoring Centre for Drugs and Drug Addiction, 2012; Schulze et al., 2012; Talbot et al., 2016). Prescription drugs or medicines, used to treat legitimate medical conditions and illnesses, are also regarded as licit substances. Sometimes medicines are even prescribed to reinstate a person's driving skills. Nevertheless, medicines can have harmful consequences for driving skills as they can be consumed incorrectly for instance by taking higher doses than recommended, irregular consumption or by combing multiple medicines. Illicit psychoactive substances are usually denoted as drugs. Examples of drugs are cannabis, cocaine, illegal opiates, heroin etc. Illicit drugs differ from their licit counterparts (medicines or medicinal drugs) in terms of the context in which the substances are used. Medicines are prescribed by a medical practitioner and are part of medical treatment whereas drugs do not have any medical purposes and are merely used recreationally to experience the sought after effects associated with certain illicit drug types (ETSC, 2017).

4.6.2 Measuring substance impairment

Substance impairment can be measured by means of traditional methods, primarily used for law enforcement purposes, and through emerging wearable sensor technologies, which allow continuous monitoring.

Traditional methods

The most common approaches to detect substance impairment are obtained from biological samples by means of blood tests, saliva tests and breathalysers (Mahmud, Fang, Carreiro, Wang, & Boyer, 2019; Marillier & Verstraete, 2019; Veisten, Houwing, Mathijssen, & Akhtar, 2011). Other methods, such as testing hair, urine, and sweat, may prove substance use but have a longer detection window which can give rise to corrupted and inaccurate results (Mahmud et al., 2019; Verstraete, 2004). This implies that when traces of substance use have been found, it cannot be determined with 100% certainty whether the individual is still impaired or has recently consumed psychoactive substances (Mahmud et al., 2019). Furthermore, urine sampling is regarded as a gold standard for detecting drug use, but these tests affect the physical integrity and are difficult to carry out in the context of roadside police detection. This is also applicable to blood tests. Most drugs can be detected in the blood for approximately 24 hours after consumption but are relatively expensive and cannot be conducted during roadside detection (Veisten et al., 2011; Verstraete, 2004). For these reasons, police officials prefer saliva collection to detect drug use because of its ease of use, low invasiveness and low risk of infection (Asbridge & Ogilvie, 2015). Saliva samples are suitable to detect recent drug use because they have the same short detection window as blood samples (Verstraete, 2004). However, some drugs such as cannabis and ecstasy have the side effect of reducing salivation making it challenging to obtain a sufficient sample (Marillier & Verstraete, 2019). Due to recent improvements, the currently available saliva testing devices also have acceptable sensitivity rates. Strano-Rossi et al. (2012) compared four commercial site saliva drug screening devices, namely DDS®, Drugtest 5000®, Drugwipe5+®, and Rapid STAT®. All devices yielded acceptable performance rates for different drug types but the Drugtest 5000® was the only commercially available saliva kit that obtained the highest sensitivity rates (Strano-Rossi et al., 2012).

Alcohol impairment is traditionally measured by means of breathalysers, which indirectly estimate BAC by measuring breath alcohol concentration (Campbell, Kim, & Wang, 2018). This detection method can be easily applied during roadside detection by means of portable breathalysers but the results can be inaccurate due to inconsistent system calibration, contamination from compounds present in the mouth and interference from several factors such as humidity and temperature (Ali et al., 2013; Simpson, 1987; Worner & Prabakaran, 1985).

In addition to testing substance impairment by means of biological samples, behavioural tests can also be used to monitor driver impairment. An example of a behavioural test is the Standardized Field Sobriety Test in which driver impairment is observed by means of a number of coordination exercises such as One-Leg Stand Test, Horizontal Gaze Nystagmus Test and Walk and Turn Test (Marillier & Verstraete, 2019). Behavioural tests can also include monitoring external characteristics of substance use such as wide pupils, bloodshot eyes or excessive chewing movements. These tests are usually performed to provide an extra confirmation of driver impairment (Marillier & Verstraete, 2019).

Another primary information source of substance impairment is self-report measures, which are typically used in clinical trials. However, self-report measurements are very often characterized by inaccurate reporting rates. Consumers also have the tendency to underreport their consumption because substance use is regarded as socially disapproved or stigmatized (Mahmud et al., 2019).

Emerging technologies

As mentioned in the previous section, the traditional methods to detect substance impairment suffer from various weaknesses: the methods are time-intensive and expensive, the detection window can vary significantly according to the substance type leading to inaccurate results and retrospective instead of continuous insights into substance use are provided (Mahmud et al., 2019). These weaknesses stress the necessity for alternative technologies or detection techniques that provide a continuous, robust and accurate way to monitor substance impairment among drivers. Recent advancements in wearable sensor technologies can fill this void. Wearable sensor devices or wearables can register a variety of physiologic measures and are very user-friendly due to their compact size (Mahmud et al., 2019). Additionally, wearables have gained significant research attention during the past years and are also becoming more commercially available (i.e. smartwatches) (Bandodkar, Jeerapan, & Wang, 2016; Heikenfeld et al., 2018; Mahmud et al., 2019; Windmiller & Wang, 2013). Current wearables technologies use biological samples (breath- and transdermal-based measures), touch-based measures, ocular-measures and physiological measures to detect substance impairment.

Transdermal-based wearable sensors

Wearable alcohol sensors have emerged as a valuable technology for non-invasive, objective and continuous monitoring of alcohol consumption and potential intervention (Barnett, 2015; Campbell et al., 2018; Fairbairn & Kang, 2019; Leffingwell et al., 2013; Wang, Fridberg, Leeman, Cook, & Porges, 2018). The majority of wearable alcohol sensors use transdermal measurements to monitor alcohol impairment. Transdermal sensors estimate the level of blood alcohol concentration by examining the amount of alcohol that is present in water vapour emitted from the skin through a device that rests on the surface of the skin (Fairbairn & Kang, 2019). Several studies have established strong correlations between transdermal alcohol concentration (TAC) and blood alcohol concentration (BAC) (Luczak & Rosen, 2014; Sakai, Mikulich-Gilbertson, Long, & Crowley, 2006). The first wearable transdermal alcohol sensor was the Secure Continuous Remote Alcohol Monitor Continuous Alcohol Monitoring™ (SCRAM CAM) anklet (Alcohol Monitoring Systems, Inc., 2019) which is primarily used and developed for law enforcement purposes. The anklet is the most widely used and validated wearable transdermal alcohol sensor (Leffingwell et al., 2013; Sakai et al., 2006). Furthermore, several studies have tested the anklet and concluded that the SCRAM CAM™ is an objective monitoring device which improves the prediction of alcohol dependence and efficacy of alcohol intervention (Barnett, Meade, & Glynn, 2014; Barnett et al., 2017; Dougherty et al., 2015; Neville, Williams, Goodall, Murer, & Donnelly, 2013). Nevertheless, the SCRAM CAM™ anklet suffers from several limitations which hinder its application for monitoring alcohol use over longer time periods such as its large size and weight, high purchase costs combined with daily user fees and large sampling intervals (30 min.) (Barnett, 2015; Leffingwell et al., 2013). The newest generation of wearable transdermal alcohol sensors is wrist-worn. Consequently, these wearable devices are smaller, making them more acceptable for daily use as their size and weight are comparable to that of a smartwatch. The ION™ (Milo Sensors, 2019) and BACtrack Skyn™ (BACtrack Inc., 2019) are two wrist-worn transdermal alcohol sensors that are commercially available and are specially designed for consumer use making them more suited to use for research purposes than the SCRAM CAM™ anklet. Both devices overcome several weaknesses of the SCRAM CAM™ anklet: they are user–friendly (light, easy to wear), relatively inexpensive, can sample alcohol impairment at a higher frequency (i.e. every second), transmit data in real-time, and display data through a specialized smartphone app or in an online database (Fairbairn & Kang, 2019; Wang et al., 2018). The specialized smartphone apps for both devices provide real-time visualizations of the current alcohol sensor reading, skin temperature, data connection status, battery status of the device and allow selection of sampling intervals for the alcohol sensor (i.e., every 1 s, 10 s, 30 s, 1 min, or 5 min (BACtrack Inc., 2019; Milo Sensors, 2019; Wang et al., 2018). Fairbairn and Kang (2019) evaluated both devices and found a strong correlation between BAC and TAC measured by means of both wrist sensors indicating that both devices are valid tools to measure (transdermal) alcohol intoxication.

Secure Remote Alcohol Monitor (SCRAM™) (adopted from Alcohol Monitoring Systems, Inc., 2019)

ION™ alcohol tracking wearable by Milo Sensors (adopted from Milo Sensors, 2019)

BACtrack Skyn® from BACtrack (adopted from BACtrack Inc., 2019)

Figure 4: transdermal alcohol monitoring wear

Physiology based wearable sensors

The consumption of psychoactive substances causes a physiological change in the body. Therefore, several studies have focused on monitoring physical reactions to substance use such as electrocardiogram (ECG), electro dermal activity (EDA), skin temperature, heart rate, and locomotion (Angarita et al., 2015; Carreiro, Fang, et al., 2015; Carreiro, Smelson, et al., 2015; Carreiro et al., 2016; Howell, Nag, McKnight, Narsipur, & Adelegan, 2015; Natarajan et al., 2016, 2013). Researchers from the University of Massachusetts Amherst and Yale University used a wearable ECG sensor to detect cocaine use (Angarita et al., 2015; Natarajan et al., 2016, 2013). Cocaine and other stimulant drugs have an impact on the heart, which makes it possible to measure their effect through ECG signals. It should be kept in mind that regular daily activities such as stress, anxiety, and workouts also influence the heart rate which can lead to biased results (Mahmud et al., 2019). The researchers used a commercially available Zephyr Bioharness 3 chest band® (Medtronic, n.d.) to monitor cocaine impairment and this system was able to measure cocaine use with high accuracy (>0.9) and sensitivity. Furthermore, Howell et al. (2015) developed a wearable biosensor in the form of a wristwatch to detect substance use by measuring multiple physiological data such as heart rate, temperature, and skin conductance. The wristwatch was linked to a smartphone app that sends alerts to the user if he tested positive for substance impairment. Substance use also influences motor activity. Park et al. (2017) developed a smart shoe equipped with a pressure sensor that registers the changes in gait caused by alcohol consumption with an accuracy rate of 86.2.

Ocular measures

Ocular measures can also be used to detect alcohol impairment. BreathalEyes® is a smartphone app that can be used to detect alcohol impairment by scanning your eyes in order to detect and analyse the Horizontal Gaze Nystagmus (HGN) (United States Patent No. US9042615B1, 2015). By analysing HGN the app yields a result of how impaired you are based on the involuntary twitching of the eye (United States Patent No. US9042615B1, 2015). To the best of the i-DREAMS consortium's knowledge, there is no study publicly available that evaluates the accuracy rate of the BreathalEyes® app.

Touch-based wearable sensors

Trutouch technologies has created TruTouch® which uses a touch-based system to measure blood alcohol concentration by spectroscopically measuring alcohol in the user's tissue (TruTouch Technologies, 2019). In other words, this technology measures the presence of alcohol beneath the skin's surface or more specifically in the capillaries. The measurement works as follows, infrared light is shined into the user's skin and reflected back to the skin's surface where it is collected by a touchpad (TruTouch Technologies, 2019). The light that is reflected back contains information about the unique chemical properties of the skin, including the level of alcohol concentration (TruTouch Technologies, 2019). TruTouch technologies are currently investigating how this technology can be embedded in the start button or steering wheel of vehicles. If this technology proves to be reliable, it could have the potential to replace the installation of rather invasive alcohol interlocks in vehicles.

Breath-based wearable sensors

Similar to handheld breathalysers used by police officers to measure alcohol impairment during roadside testing, companies have produced different wearable devices for personal use. For example, Tokyoflash developed the Kisai Intoxicated LCD watch® with a built-in breathalyser to check blood alcohol levels (Tokyoflash Japan, 2019). The procedure works as follows, users breathe into the built-in breathalyser and the watch determines and displays the user's blood alcohol level by means of colour code (red: highly intoxicated, yellow: medium intoxicated and green: not intoxicated) (Tokyoflash Japan, 2019). The manufacturer claims that the Kisai Intoxicated LCD watch® can be used to provide an indication of alcohol impairment but does not provide completely accurate results.

Figure 5: The Kisai Intoxicated LCD watch® (adopted from Tokyoflash Japan, 2019).

4.6.3 Review of studies

The presented literature review on substance impairment is primarily based on deliverables from the DRUID (Driving Under the Influence of Drugs, Alcohol and Medicines) project. Seven years after completion, this project still provides the most extensive and precise information and insights regarding the prevalence, impairment effects on driver skills and measuring techniques regarding psychoactive substances within the European road traffic system. These findings were complemented with the findings of deliverable D4.1 from the recently completed Horizon 2020 project SafetyCube (Talbot et al., 2016). This deliverable provides a literature review on road user-related risk factors and, amongst other risk factors partly focused on the safety risks of drink and drug driving.

The relevant studies regarding the use of new emerging technologies to monitor driver substance impairment were obtained by a Google Search using the following keywords:

- Driver monitoring and wearables
- Alcohol impairment and wearables
- Drugs impairment and wearables
- Medicines impairment and wearables

4.6.4 Conclusion for measuring substance impairment in i-DREAMS

Substance impairment can be measured in various ways. Traditional methods to detect substance impairment use biological samples by means of blood tests, saliva tests, urine samples, and breathalysers. Self-report measures and behavioural tests can also be used to monitor driver impairment and are mostly used to gain the first confirmation of substance impairment. However, self-report measurements are very often characterized by inaccurate reporting rates or by participants providing socially desirable answers resulting in biased results (Mahmud et al., 2019). These methods are often used for enforcement purposes in the context of roadside police detection, which can affect the physical integrity of the individual and thus the participant compliance rate. They also do not allow for the continuous monitoring of substance impairment. Because of all these reasons, these traditional methods are difficult and less suited to carry out for research purposes within i-DREAMS.

However, recent technological advancements in the area of wearable sensor technologies create new opportunities for the continuous monitoring of substance impairment by means of biological samples (breath- and transdermal-based measures), touch-based measures, ocular-measures, and physiological measures. Wearable sensor technologies using touch-based, breath-based and ocular measures are still under development and/or have not been validated. Physiological measures are validated and provide high accuracy but wearing a chest-band for continuous monitoring purposes can be unpleasant for the participant. Therefore, commercially available wrist-worn transdermal alcohol sensors, which are comparable to a smartwatch, have more potential to be used as a tool to measure substance impairment within i-DREAMS. These devices are user-friendly, non-invasive, relatively inexpensive, sample alcohol impairment at a high frequency (i.e. every second), allow for real-time continuous monitoring, display the data through a specialized smartphone app and/or in an online database and provide objective and validated measurements of the level of alcohol impairment (Fairbairn & Kang, 2019; Wang et al., 2018).

As became apparent from the literature review, driving under the influence of drugs and medicines is still not as well understood as drink driving. Numerous studies have investigated and demonstrated the effects of psychoactive drugs and medicines on driving ability, but there appears to be no universal agreement on how best to measure the levels of impairment that psychoactive drugs cause to the driver (UK Department for Transport, 2013). This is primarily caused by the variety of substances and their diverse effects. This lack of knowledge is also reflected in the development of new wearable technologies to monitor substance impairment as the currently available technologies primarily focus on monitoring alcohol impairment. All these aspects should be taken into consideration when designing the substance impairment component of the i-DREAMS system.

Regardless of the measurement methods and their quality, practical considerations for implementations in i-DREAMS should be noted. Although impairment may be measurable in real-time with increasing reliability, the effects of the specific impairment may be expressed in impaired attention and alertness and thus, already accounted for by the corresponding real-time measurements. This should be borne in mind for the model of the safety tolerance zone.

5 Driver characteristics – Indicators and measuring methods

As argued in the introduction, not all individual-related factors that determine the driver's capacity to cope with the task demand are able to be measured in real time while driving. Factors such as personality traits, driving experience or health status are relatively stable over time and thus, they do not change suddenly. Nevertheless, they will be considered and explored here, since some of those factors will be subject to a one-time measurement for one or more reasons:

1) It has to be considered whether or not some of the driver characteristics will be introduced into the safety tolerance zone model as a constant and serve as a kind of correction factor (see also deliverable 3.1: Talbot et al., 2020). When exploring this option, it is important to exercise care. It has to be avoided to account for a contributing factor twice, since they can be already reflected in the driver state or the driver behaviour. Although this can be difficult to assess, e.g. risk-taking tendencies might be reflected already in driving behaviour or certain diseases might be expressed through the mental state.
2) Some of the factors which will be measured real-time during driving can benefit from contextualisation and validation. Interpreting the construct 'emotions' for example, measured in real-time may be facilitated by providing information on inter-individual differences in e.g. emotional regulation competence or anger proneness.
3) One of the aims of i-DREAMS is to make the collected data of the trials available for further research in form of a comprehensive database. Since we cannot anticipate future needs for control variables, background factors of the test subjects may be very valuable.
4) Some of the driver characteristics may help to provide customized interventions (real-time and post-trip, respectively).

Constructs and variables of different categories will be relevant to capture, such as competences, personality traits, health conditions, habits and socio-demographic factors. The methods to collect data accordingly are manifold. While age and profession can simply be queried, while e.g. reactivity is measured via standardised, validated and normed performance tests. The latter provide for objectivity compared to self-assessments. On the other hand, tests often require equipment, can be expensive and long. Thus, the trade-off regarding performance tests is often between validity and efficiency. Although, self-reported information is subject to desirability bias (among other problems) it is often the sensible choice due to resource restrains.

The aim of this review was to identify what will be relevant for one or more of the above-mentioned purposes and to recommend measuring tools accordingly. The evaluation and recommendation are based on the applicability in i-DREAMS in terms of:

- Language availability
- reliability and validity
- Costs and required materials
- Length (processing time for participants)
- Reasonableness

In chapters 5.1 to 5.5, stable factors which are known to be road safety relevant will be described and a range of measurement methods will be documented. Methods were not preselected in order to give a full picture of the options, independent of i-DREAMS requirements. The selection and recommendation will be provided in 5.6.

5.1 Competences

For safe driver performance and safe driver behaviour numerous of physical and mental competencies, and skills are preconditioned which are beyond the mere handling of the vehicle itself (steering, braking, switching gear etc.) and can also be referred to as 'psychological and medical fitness to drive'. While many of these competencies and skills can be tested through single or combined (psychological) tests.

In the field of differential psychology or psychological diagnostics, a test battery or a test system is a combination of different individual tests. Hence, some fitness to drive tests are applicable to multiple constructs or dimensions for measuring driving abilities. A selection of these test batteries is summarised below, as they will be mentioned throughout the chapter. The individual tests of these batteries or systems as a whole will be mentioned and described in detail where appropriate in the sub-chapters.

- The single tests in the **Vienna Test System** were developed specifically for the use in driver assessments and are thus precisely tailored to this field of application. All individual tests are validated and are also appropriate for people with limited or no computer literacy. The standardized and objective process ensures that the same conditions apply to all test persons and therefore have the same opportunities – regardless of their level of education or cultural background (Chaloupka & Risser, 1995).
- The **Perception and Attention Functions Test battery** (WAF) consists of different sub-tests for measuring alertness, vigilance and sustained attention, divided attention, focused attention, selective attention, spatial attention and neglect, smooth pursuit eye movements and visual scanning. A total of 42 subtests are available, which can be specified independently of each other or in any desired combination (Häusler & Sturm, 2009).
- The **Test of Attentional Performance-Mobility** (TAP-M)[10] is able to measure attentional aspects of the ability to drive and includes 9 sub-scales.
- The test battery itself consists of different subtests that measure the active visual field, alertness, distractibility, executive control, sustained attention, divided attention and others.
- The neuropsychological test battery **CERAD-Plus** consists of the following tests: Verbal fluidity (Animals), Boston Naming Test (15 items), Mini Mental Status Examination, Word list Learning – recalling – recognizing; Copying of Figures; Trail Making Test A and B; Phonematic fluidity (Aebi, 2002).
- The **Assessment of Driving-Related Skills** (ADReS) is a collection of individual tests and a screening tool for physicians. It uses two tests to determine visual abilities: the confrontational field testing and the Snellen chart to determine visual acuity.
- The **Corporal Plus** is used in the field of fitness to drive assessments and medicine in Germany, occupational medicine, gerontology, clinical psychology and neuropsychology. The system offers a variety of diagnostic methods for capturing visual and combined visual-auditory responsiveness, concentration, attention, spatial orientation, as well as working memory.

[10] https://www.psytest.net/index.php?page=TAP-M&hl=en_US

5.1.1 Emotional regulation

According to Gross (2002) emotional regulation includes those processes that enable us to influence what emotions we have, when we feel them, and how we experience and express them. However, emotional regulation goes beyond the reduction of negative emotions. It also includes the control of positive emotions and in addition, the intensity of emotion can either be reduced, sustained or increased as it is seen best for a respective event or situation.

Therefore, individuals with a low emotional awareness level on the one hand and a lack of emotional coping mechanisms on the other may have reduced driver safety and would be more frequently distressed (Legree et al., 2003). Consequently, a person that is able to permit or delay spontaneous reactions may be regarded as a safer driver. Below a selection of tests for measuring the ability of emotion regulation is listed:

Cognitive Emotion Regulation Questionnaire (CERQ)

CERQ is specifically designed to assess the conscious cognitive components of emotion regulation. It is a self-report questionnaire consisting of 36 items that measures nine different cognitive coping strategies.

The CERQ makes it possible to identify individual cognitive strategies and compare them to norm scores from various population groups. In addition, the questionnaire offers the opportunity to investigate relationships between the use of specific cognitive coping strategies, personality variables, psychopathology and other problems (Garnefski et al., 2001).

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Add. equipment needs	Validity
Cognitive Emotion Regulation Questionnaire (CERQ)	Questionnaire	DUT	ENG*	-	Free	-	Good factorial validity and high reliabilities, with Cronbach's α ranging between .75 and .87.
		FRE	GER				
		GRE	POR				

* translation possible

Perth Emotion Regulation Competency Inventory (PERCI)

The Perth Emotion Regulation Competency Inventory a 32 item self-report questionnaire that measures the ability to regulate both, negative and positive emotions. The individuals have to agree or disagree how much the statements apply to their self. The questionnaire itself is freely available for use (Preece et al., 2018).

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Perth Emotion Regulation Competency Inventory (PERCI)	Questionnaire	DUT	ENG*	6-10	Free	-	All subscales have good internal consistency reliability (α=0.85–0.94) and all composite scores had good internal consistency reliability (α=0.92–0.94).
		FRE	GER				
		GRE	POR				

* translation possible

Inventory of traffic-related personality traits – Revision (IVPE-R)

The IVPE-R[11] as an individual test of the Viennese Test System measures personality traits that are relevant for driving such as self-control, social responsibility, mental stability and risk avoidance. On the basis of a free scale, the individual assess to what extent a certain statement is applicable to them.

The IVPE also allows the detection of other relevant personality traits like sensation seeking or social sense of responsibility.

The test is not available in Dutch, Greek and Portuguese language.

Measurement tool	Type of measurement tool	i-DREAMS languages	Time in minutes	Costs	Additional equipment needs	Validity
Inventory of traffic-related personality traits (IVPE-R)	Personality test	DUT ENG FRE GER GRE POR	12-15	Price quote upon applica tion	-	Reliability in terms of internal con-sistency is given due to compliance with the Rasch model for each scale. Validity has in addition been demonstrated by a confirmatory factor analysis.

STAI (State Trait Anxiety Inventory)

The concept of the STAI is to measure trait anxiety as a proneness to interpret situations as menacing which then elicitings state anxiety. It is also available as a short version with ten items. Participants rate themselves on a four-point scale to indicate the frequency of certain situations (Spielberger et al., 1970).

5.1.2 Stress regulation

Drivers are regularly involved in more or less risky situations while driving. The threat that always resonates with these risks is subsequently experienced as stress (James & Nahl, 2002). Stress in turn may lead to an increased amount of near crashes and accidents. Referring to this, Westerman & Haigney (2000) pointed out that drivers, who experience high levels of stress also report frequent lapses, errors and violations. Hereafter, a selection of tests for measuring the ability of stress regulation can be found:

Determination test (DT)

The DT[12] is part of the Viennese Test System and measures the individual's stress tolerance. The task is to continuously, quickly and differently respond to rapidly changing optical or acoustical stimuli by pressing the appropriate buttons on a keyboard or by means of a foot pedal. Due to this adaptive test specification, each person can be put into a situation of an overwhelming task demand with a correspondingly high stimulation.

The test is language-free. Special standard norms for professional drivers are available.

[11] https://www.schuhfried.at/test/IVPE-R
[12] https://www.schuhfried.at/test/dt

Measurement tool	Type of measurement tool	i-DREAMS languages [13]		Time in minutes	Costs	Additional equipment needs	Validity
Determination Test (DT)	Multiple stimulus-response test	DUT	ENG	6-15	Price quote upon application	+	No information
		FRE	GER				
		GRE	POR				

Stress processing questionnaire (SVF)

The SVF[14] enables the measurement of coping or processing mechanisms in stressful situations. The questionnaire represents an inventory of methods that relate to different aspects of stress management.

The SVF is only available in German and Czech language.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Stress processing questionnaire (SVF)	Questionnaire	DUT	ENG	10-15	ca. € 1,000	-	Numerous results on construct validity as well as on differential and criteria validity are documented.
		FRE	GER				
		GRE	POR				

Differential stress Inventory (DSI)

The DSI[15] is also part of the Viennese Test System and captures both, the level of extent and causes of individual stress experiences. The method enables a differentiated measurement of stress triggers, stress manifestations, available coping strategies and risks of stress stabilization.

The test is not available in Dutch, French and Greek language.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Differential stress inventory (DSI)	Special personality test	DUT	ENG	23	Price quote upon application	-	DSI is a valid measure of the construct of daily stress.
		FRE	GER				
		GRE	POR				

TICS (Trierer Inventar zum chronischen Stress)

The TICS[16] is a standardized questionnaire with 57 items for the differentiated diagnosis of chronic stress. In answering the TICS, the individuals indicate how often they have experienced a given situation in the past three months.

However, there are no hints that the questionnaire could be relevant for traffic related research.

[13] language free task
[14] https://www.testzentrale.de/shop/stressverarbeitungsfragebogen.html
[15] https://psyexpert.de/wp-content/uploads/2019/04/DSI.pdf
[16] https://www.testzentrale.de/shop/trierer-inventar-zum-chronischen-stress.html

5.1.3 Attention regulation

While driving a vehicle the driver has to cope with numerous attention-demanding tasks. Above all, selective and sustained attention is a fundamental ability for safe driving.

The following selection shows different tests for measuring the ability of attention and attention regulation.

Test of Attentional Performance (TAP-M)

Based on the Test of Attentional Performance (TAP)[17], which was initially designed for the assessment of attentional deficits in patients with cerebral lesions, a mobility version with 9 sub-scales was developed. The test is able to measure attentional aspects of the ability to drive.

The tasks involve easily distinguishable stimuli that individuals react to by a simple motor response. The test battery itself consists of different subtests that measure the active visual field, alertness, distractibility, executive control, sustained attention, divided attention and others. Thus, it is also applicable for other dimensions, especially for visual perception.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional Equipment needs	Validity
Test of Attentional Performance (TAP-M)	Motor Response Test	DUT	ENG	~45	€ 1,000 for 1st licence	Installation on local PC	Validation of the test battery was supported by several European institutions being members of the European project "AGILE"
		FRE	GER				
		GRE	POR				

Trail Making Test (TMT-L) – Version 'Trail A' (abbreviated), 'Trail B' (long)[18]

The TMT is part of several neuropsychological test batteries and is composed of a paper-and-pencil test.

It is an internationally widely used test procedure for checking brain function performance and consists of two parts. The TMT-A predominantly checks the processing speed e.g. visual search, whereas the TMT-B tests higher cognitive performance, such as mental flexibility e.g. attention switching.

The test covers also other dimensions e.g. visual perception.

Measurement tool	Type of Measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional Equipment needs	Validity
Trail Making Test (TMT-L), Version 'Trail A' (short), 'Trail B' (long)	Performance test	DUT	ENG	1-2 (A) 1-3 (B)	-	Timer	Construct validity was assessed by confirmatory factor analysis (TMT-L; RMSEA < .06, CFI > .98)
		FRE	GER				
		GRE	POR				

[17] https://www.psytest.net/index.php?page=TAP-M&hl=en_US

[18] http://apps.usd.edu/coglab/schieber/psyc423/pdf/IowaTrailMaking.pdf

Adaptive tachistoscopic traffic perception test (ATAVT)[19]

This test – part of the Vienna Test System – measures the ability to gain a quick visual overview as part of one's attention performance. Especially in traffic, a quick and accurate detection of complex visual situations is essential. Pictures of traffic situations are shown very briefly. Before the presentation the individual hears an announcement stimulus (sound) and after each picture it should be reported what was seen on the picture.

In addition to traffic psychology, the test is also used for psychological safety assessments for professional drivers. International special standards for older drivers are available as well.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Adaptive tachisto-scopic traffic perception test (ATAVT)	Performance test	DUT	ENG	8-14	Price quote upon application	Installation on local PC	No information
		FRE	GER				
		GRE	POR				

Cognitrone (COG)[20]

The test that is also used within the Vienna Test System measures the individual's concentration performance. Due to its high practical relevance, this test is not only used in clinical neuropsychology but also for fitness to drive assessments.

The individual compares a geometric figure with four other geometric figures and indicates whether it is identical to one of the other four figures.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Cognitrone (COG)	Test for concentration performance	DUT	ENG	8-17	Price quote upon application		Validity of the test battery could be documented.
		FRE	GER				
		GRE	POR				

Perception and attention functions battery (WAF)[21]

The test captures sub-functions of attention such as alertness, vigilance and constant attention, focused attention, divided attention, selected attention or eye movements and visual scanning. A total of 42 subtests are available, which can be specified independently of each other or in any desired combination.

For many of the used subtests, long and short forms are available.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Perception and attention functions battery (WAF)	Special performance test	DUT	ENG	5-30	Price quote upon application	To conduct the auditive and cross-modal subtests, a standardized USB headset is required.	No information
		FRE	GER				
		GRE	POR				

19 https://www.schuhfried.com/test/atavt
20 https://www.schuhfried.com/test/cog
21 https://www.schuhfried.at/test/WAF

Corporal Plus[22]

The test is used in the field of fitness to drive assessments and medicine in Germany, occupational medicine, gerontology, clinical psychology and neuropsychology.

The system offers a variety of diagnostic methods for measuring visual and combined visual-auditory responsiveness, concentration, attention, spatial orientation, as well as working memory.

The test can also be used for the dimension “Reactivity”.

Measurement tool	Type of measurement tool	i-DREAMS languages	Time in minutes	Costs	Additional equipment needs	Validity
Corporal Plus		DUT ENG FRE GER GRE POR	2-4 per subtest	€ 5,400	Is included in purchase	The internal consistency for attention and spatial orientation is .96 and .99 (Spearman/Brown)

d2-R

The d2-R[23] is the electronic form of one of the most widely-used measures of attention – particularly visual attention – throughout Europe. It is not only used within clinical and educational settings, but also within the transport sectors.

With d2-R the ability of attention and concentration can be tested quickly, reliably and with validation in individual or group settings. Additionally, the test integrated the rapidity and accuracy of distinguishing similar visual stimuli. Among other fields of application, d2-R is suitable for the use in driver assessments.

Measurement tool	Type of measurement tool	i-DREAMS languages	Time in minutes	Costs	Additional equipment needs	Validity
d2-R		DUT ENG FRE GER GRE POR	9-15	ca. € 2.600	-	Consistently values of at least .77 for concentration, rapidity and accuracy indicate a high validity

Test of Variables of Attention (TOVA)

TOVA[24] is a culture- and language-free computerised and objective measure of attention and inhibitory control. It calculates response time variability (consistency), response time (speed), commissions (impulsivity), and omissions (focus and vigilance). The test helps to assist in the assessment and evaluation of treatment for attention deficits, including attention-deficit/hyperactivity disorder (ADHD). However, as TOVA is not used in the field of fit to drive assessments it may not be applicable for calculating a driver tolerance zone.

[22] https://www.testzentrale.de/shop/psychometrisches-testsystem.html
[23] https://www.hogrefe.co.uk/shop/d2-test-of-attention-revised.html
[24] https://www.tovatest.com/

Measurement tool	Type of measurement tool	i-DREAMS languages	Time in minutes	Costs	Additional equipment needs	Validity
Test of Variables of Attention (TOVA)	Computerised test	DUT ENG FRE GER GRE POR	10-15	Ca. € 3,500	Included in purchase - use on multiple computers possible	

5.1.4 Readiness to take risks

Among other things, risky behaviour can include driving while impaired, driving too fast for the conditions, tailgating, unsafe passing or lane changing. For instance, study results showed that drivers who were involved in traffic accidents or crashes the year before took more risks when driving (Iverson, 2004). However, risky driving is notably affected by sensation seeking as it correlates with a various number of unsafe driving practices and mistakes in driving (Linkov, 2019). Most studies that focused on sensation seeking and risky behaviour showed positive relationships between sensation seeking and risky driving, with correlations between 0.30 and 0.40 (Jonah, 1997). Below a selection of tests for measuring the readiness to take risks in traffic is listed:

Vienna Risk Taking Test (WRBTV)[25]

This test, as part of the Vienna Test System, assesses the willingness to take risks by examining 24 different types of traffic situations: speeding and overtaking situations, decision situations at intersections and traffic situations in bad or good weather conditions. Due to its everyday relevance, this objective personality test is successfully used in driver assessments.

Measurement tool	Type of measurement tool	i-DREAMS languages	Time in minutes	Costs	Additional equipment needs	Validity
Vienna Risk Taking Test (WRBTV)	Objective personality test	DUT ENG FRE GER GRE POR	18	Price quote upon application	-	The significant correlations with the sensation-seeking scales, the adventurousness scale, and self-control confirm the construct validity.

Sensations-Seeking Scale (SSS-V)

A subscale of the SSS-V is dedicated to risk taking. See 5.2 for further details on the SSS-V.

5.1.5 Self-appraisal

Subjective driving skills are frequently assessed with self-reports and various studies have shown that a lot of drivers overestimate their own skills. Sundström (2008) pointed out that subjective driving skills should be assessed through a comparison with the actual driving skills in order to obtain indicators of reliability and validity.

[25] https://www.schuhfried.at/test/WRBTV

5.1.6 Hazard perception

Perception of hazards and coping test (GECO)

GECO[26] is part of the Vienna Test System and is used for measuring the perception of hazards, the knowledge of specific dangerous situations in traffic as well as handling these situations. For this purpose, video-supported dynamic traffic situations are presented from the perspective of e.g. a bicyclist, a motorcyclist or a car driver. The respondents' task is to recognize an immediate hazardous situation. Despite the fact that this performance test is not available in all of the i-DREAMS languages, it is not suitable for British participants due unfamiliarity of right side driving.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Perception of hazards and coping test (GECO)	Special ability tests	DUT	ENG	Ca. 40	Price quote upon applicati on	-	
		FRE	GER				
		GRE	POR				

5.1.7 Reactivity

Reaction test (RT)

The RT[27] is a single test of the Vienna Test System and assesses the ability to react under simple stimulus constellations (simple and choice reactions). Individuals have to react as quickly as possible to optical or acoustic signals. Special norms are available for older drivers and professional drivers.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Reaction test (RT)	Special ability test	DUT	ENG	5-10	Price quote upon applicatio n	+	
		FRE	GER				
		GRE	POR				

Corporal Plus

See above "Attention regulation" for further details.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Corporal Plus		DUT	ENG	2-4 per subtest	€ 5,400	Is included in purchase	The internal consistency for attention and spatial orientation is .96 and .99 (Spearman/Brown)
		FRE	GER				
		GRE	POR				

5.1.8 Eye-hand coordination

The ability to coordinate eye and hand when making small movements is important when driving. This ability can also be measured with help of the Two-Hand Coordination test.

26 https://www.schuhfried.com/test/geco
27 https://www.schuhfried.com/test/rt

Two-Hand Coordination (2HAND)

2HAND[28] assesses the two-dimensional visuomotor coordination between eye and hand as well as the coordination between the left and right hand. The subject has to move a red dot along a given track using either two control knobs or two joysticks.

The test is included in the Vienna Test System and the task is language-free.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Two-Hand Coordination (2HAND)	Special ability test	DUT	ENG	8-15	Price quote upon application	-	
		FRE	GER				
		GRE	POR				

5.1.9 Visual perception and visual orientation

Visual Pursuit Test (LVT)

The LVT[29] measures visual orientation performance and visual perception and is used among others in the field of fitness to drive assessments. The test presents a number of random and disorderly lines and the individual has to visually identify the end of a particular line as quickly as possible. LVT is language-free and special norms are available for professional drivers.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Visual Pursuit Test (LVT)	Special ability test	DUT	ENG	5-25	Price quote upon application	-	Coefficient of internal consistency is r=.96; n=785 (Kubinger & Ortner, 2010)
		FRE	GER				
		GRE	POR				

Perception and attention functions battery (WAF)

The test battery includes subscales for eye movements and visual scanning.

The subject's task is to search a 6x6 matrix of similar visual stimuli and decide whether a previously defined stimulus is present or not.

See above "Attention regulation" for further details on WAF

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Perception and attention functions battery (WAF)	Special performance test	DUT	ENG	5-30	Price quote upon application	A standardized USB head-set is required.	No information
		FRE	GER				
		GRE	POR				

Test of Attentional Performance (TAP-M)

One of the nine subscales measures the individual's visual field.

See above "Attention regulation" for further details.

[28] https://www.schuhfried.com/test/2hand
[29] https://www.schuhfried.com/test/lvt

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional Equipment needs	Validity
Test of Attentional Performance (TAP-M)	Motor response test	DUT	ENG	Ca. 45	1,000 EUR for 1st licence	Installation on local PC	Validation of the test battery was supported by several European institutions being members of the European project "AGILE"
		FRE	GER				
		GRE	POR				

Line Orientation Test (LAT)

The Line Orientation Test[30] captures the visual orientation ability, which is an essential component of spatial-perceptual functions. The test is used when visual orientation ability is relevant for fitness to drive assessments.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Line Orientation Test (LAT)	Special ability test	DUT	ENG	4-5		-	No information
		FRE	GER				
		GRE	POR				

Motor Free Visual Perceptual Test Version 4 (MVPT-4)

The MVPT-4[31] is used to assess the visual-perceptual ability via a series of visual-perceptual tasks that do not require a motor response from the individual.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Motor Free Visual Perceptual Test Version 4 (MVPT-4)	Special ability test	DUT	ENG	20-25	195$ for 25 forms	-	No information
		FRE	GER				
		GRE	POR				

Corporal Plus

See above "Attention regulation" for further details.

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Corporal Plus		DUT	ENG	2-4 per subtest	€ 5,400	Is included in purchase	The internal consistency for attention and spatial orientation is .96 and .99 (Spearman/Brown)
		FRE	GER				
		GRE	POR				

[30] https://www.schuhfried.com/test/lat

[31] https://www.wpspublish.com/mvpt-4-motor-free-visual-perception-test-4

5.2 Personality traits

A careful and thoughtful driving behaviour is crucial for traffic safety. Personality factors are known to have an impact on road safety. Although in many studies the isolated effect of e.g. sensation seeking is rather small, it is also beyond dispute (Goldenbeld & van Schagen, 2016; Sümer et al., 2003). The following scales and questionnaires are frequently used to measure safety relevant personality traits.

Sensation Seeking Scale (SSS-V)

The SSS (form V) is a normed questionnaire (Zuckerman, 2007). As sensation seeking is a multidimensional construct, the SSS includes four different subscales, which measure thrill and adventure seeking (TAS), disinhibition (Dis), experience seeking (ES) and boredom susceptibility (BS).

Each subscale contains 10 items. The SSS total score of all 40 items is used to measure sensation seeking as an overall score. Regarding the internal consistency of the subscales, the Cronbach's alpha for the English version were $\alpha = 0.81$ and $\alpha = 0.82$ for the TAS in males and females; $\alpha = 0.78$ and $\alpha = 0.77$ for the DIS; $\alpha = 0.65$ and $\alpha = 0.59$ for the BS; and $\alpha = 0.65$ and $\alpha = 0.67$ for the ES (Gianfranchi et al., 2017).

A German version of the SSS-V is available. Further language versions were not identified. Since the SSS-V is text-based a translation into other i-DREAMS languages would be possible. However, validity scores apply only to the English version.

Brief Sensation Seeking Scale (BSSS)

The BSSS is based on the Sensation Seeking Scale and was developed by Hoyle et al. (2002). The BSSS consists only of 8 questions and still carries reasonable reliability and validity (Hoyle et al., 2002). Limitations:

- There are different answer formats of the short German and English version
- There is no evidence that BSSS is used in fitness to drive assessments.

Deffenbacher Driving Anger Scale (DAS)

The DAS is a 14-item questionnaire and the most widely used instrument to measure trait driving anger. The items include brief descriptions of potentially anger-provoking driving situations (e.g., sticking in a traffic jam), and respondents rate the degree to which each situation would anger them using a five-point Likert scale (Deffenbacher et al, 2016; Deffenbacher et al., 1994).

Propensity for Angry Driving Scale (PADS)

The 19-items questionnaire describes driving scenarios that are likely to provoke anger and respondents indicate how they would react to each situation by selecting one of four options ranging from mild to extreme. The PADS has a very good internal consistency (as=.88 and .89) and a four-week test-retest reliability (r=.91). The scale significantly (>.05) predicted moving tickets, minor accidents, aggressive driving, risky driving, and maladaptive driving anger expression, above and beyond gender, miles driven per week, and trait anger (Dahlen & Ragan, 2004).

State-Trait-Anger Expression Inventory (STAXI-2)

The STAXI-2[32] is a 57-item inventory which measures the experience, expression, and control of anger. It consists of the following subscales: state anger, trait anger, anger expression - out, anger expression - in, anger control - out, anger control - in, and anger expression index. However, there is no indication that the STAXI-2 is used in fitness to drive assessments.

UPPS-P Impulsive Behavior Scale

The UPPS-P is a 59 item self-assessment scale with the five subscales urgency, premeditation, perseverance, sensation seeking, and positive urgency, used for adolescents and adults (12 years and older). While the scale is not directly measuring impulsivity as a trait, it assesses personality aspects which lead to impulsive behaviour. Acts and incidents are rated on a four-point scale for their frequency during the past six months (Whiteside et al., 2005).

5.3 Habitual driving behaviour

A multitude of questionnaires exist which try to capture self-reported past or habitual behaviours of drivers. One of the most frequently used questionnaires is, for example, the Manchester Driver Behaviour Questionnaire (DBQ). On the other hand, many project teams and research groups have introduced their own items depending on the specific research question. One of the recent large-scale surveys in the domain of road safety is the ESRA initiative (Meesman et al., 2019), which used a combination of various items from established questionnaires with slight variations.

With regards to the behaviours **speeding, tailgating, fatigued driving, impaired and distracted driving,** the DBQ as well as the ESRA questionnaire used relevant items. Although simply using self-constructed or adapted items is feasible and legitimate, using established items to query those behaviours from i-DREAMS participants allows for comparability with other data sets.

Manchester Driver Behaviour Questionnaire (DBQ)

The DBQ is a self-report questionnaire as a measure of deviant driving behaviours. The questionnaire exists in numerous different versions using various combinations of items. The original version consists of 50 items, assessed through a six-point Likert scale ranging from 0=never to 5=nearly all the time. The short version consists of 24 items investigating general factors, aggressive violations, ordinary violations, slips and errors (de Winter & Dodou, 2010).

For past and habitual **aggressive and other risky driving behaviours**, various dedicated questionnaires are available:

- ADBQ (Aggressive Driver Behaviour Questionnaire)
- DDDI (Dula Dangerous Driving Index, dimension 'aggressive driving', ‘risky driving’)
- DBQ (Manchester Driver Behaviour Questionnaire, subscale ‘aggressive violations’)
- DAX (Driving Anger Expression Inventory)
- AVIS (Aggressive driving behaviour)

[32] https://www.testzentrale.de/shop/state-trait-anger-expression-inventory-2tm.html

5.4 Health status

The general health status of a driver is of utmost importance as a prerequisite to safely participate in traffic and steer a vehicle. This importance is also reflected in the mandatory health check before talking the driving test in many countries.

There are many standard tests available for screening neurological symptoms. Some of them, however, are designed to allow for early diagnosis of dementia and therefore are normed for the elderly.

5.4.1 Neurologic assessment

Montreal Cognitive Assessment (MoCA)

The MoCA[33] test is a proven cognitive screening tool for illnesses, including Alzheimer's disease; Parkinson's disease; VCI/Stroke; ALS; Sleep behaviour disorder and others. The test has been validated for 55-85-year olds. There is no indication that MoCA is used in fitness to drive assessments.

CERAD-Plus

The neuropsychological test battery **CERAD-Plus** consists of the following tests: Verbal fluidity (Animals), Boston Naming Test (15 Items), Mini Mental Status Examination, Word list Learning – recalling – recognizing; Copying of Figures; Trail Making Test A and B; Phonematic fluidity (Aebi, 2002).

Mini Mental Status Examination (MMSE)

The Mini Mental Status Test was published by Folstein et al. (1975) and designed to provide a routine clinical screening tool for detecting cognitive deficits. However, the test is not appropriate to detect mild cognitive impairment. The test is used for persons older than 65 years.

Boston naming test

This test measures confrontational word retrieval in individuals with aphasia or other language disturbance caused by stroke, Alzheimer's disease, or other dementing disorders (Kreutzer et al., 2018).

Corporal Plus (see "Attention regulation" for further details)

DemTect

This test focuses on the main symptom of Alzheimer's dementia, i.e. memory and decelerated memory and consists of five task demands: word list task, number transcoding task, word fluency task, digit span reverse, and delayed recall of the word list. DemTec is suited to detect minor cognitive losses (Kalbe et al., 2004).

Measurement tool	Type of measurement tool	i-DREAMS languages	Time in minutes	Costs	Additional equipment needs	Validity
DemTect		possible to translate	8-10	Free	-	No information

33 https://www.mocatest.org/

Short Blessed Test (SBT)

The SBT[34] is a brief performance-based screening instrument to identify elderly persons with cognitive dysfunction (quicker than MMSE). The SBT provides good diagnostic test characteristics and overlaps with MMSE results.

Measurement tool	Type of measurement tool	i-DREAMS languages	Time in minutes	Costs	Additional equipment needs	Validity
Short Blessed Test (SBT)		possible to translate	5-10	Free	-	Good

Test for an early detection of dementia (TFDD)

The test for an early detection of dementia consists of two parts. The first part serves as an early diagnosis of dementia. In the second part, an assessment of depression is conducted. The person concerned should assess oneself, followed by an assessment of a medical practitioner or a close person (Ihl et al., 2000). The TFDD is assessed as not applicable for i-DREAMS.

ADAS-cog

The ADAS is a scale for assessing the progression of dementia symptoms. It covers cognitive performance (orientation, memory, naming objects, following instructions), but also the behaviour during the interview and psychopathological symptoms (Mohn et al., 1983; Rosen et al., 1984). As interviews and behaviour observations are part of ADAS, the test seems not practicable for i-DREAMS.

Gross Impairments Screening Battery (GRIMPS)

The GRIMPS[35] is a tool for screening gross impairments in physical and perception functions that are essential for safe driving. It is a collection of individual tests targeted for persons aged 72+. The battery consists of various tests for the two domains 1) physical measures (Rapid-pace walk, foot-tap test, head-neck rotation and arm reach) and 2) perceptual-cognitive measures (Motor-free Visual Perception Test, Trail Making Test Part A and B, Cued/delayed Recall, Scan Test and Visual Acuity).

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
Gross Impairments Screening Battery (GRIMPS)	Test battery	DUT	ENG	15	$40/kit	Various	No information
		FRE	GER				
		GRE	POR				

Assessment of Driving-Related Skills (ADReS)

The ADReS is a collection of individual tests and a screening tool for physicians. It uses two tests to determine visual abilities: the confrontational field testing and the Snellen chart to determine visual acuity. To assess motor ability, the test uses three measures: (1) the Rapid

34 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3080244
35 https://one.nhtsa.gov/people/injury/olddrive/safe/01c02.htmhttps://dict.leo.org/franz%C3%B6sisch-deutsch/

pace walk, (2) range of motion testing, and (3) the manual muscle test. To screen for cognitive deficits that might affect driving ability, the test includes the Clock Drawing Test and the Trailmaking Test, Part B. However, a study from the American Medical Association (AMA), with support from the National Highway Traffic Safety Administration (McCarthy & Mann, 2006) showed that the ADReS may not be an efficient predictor in identifying older drivers who may or may not be at an increased risk for unsafe driving.

5.4.2 Musculoskeletal system

Rapid-Pace Walk (lower limb mobility), Questionnaire/self-report

For measuring the lower limb mobility, a three-meter path is marked on the floor. The individual is asked to walk the path, turn around, and come back to the starting point as quickly as possible. The test is scored by the total number of seconds required to complete the exam.

The cut-off scores, as reported in Staplin, et al. (2003), are more than 7.5 seconds to complete the walk indicated they were 2.5 times more likely than age-matched controls to be involved in an at-fault crash. Those with completion times greater than nine seconds had a three-fold increased risk of being in an at-fault car accident.

Alternatively, the participants can be asked to indicate any known musculoskeletal conditions or pain.

5.4.3 Cardio-vascular diseases

While there are plenty of medical procedures and tests to assess the cardio-vascular system, such as blood tests, ECG, MRI and CT, stress tests etc., a medically sound diagnosis is not required for the i-DREAMS participants. An open question on known cardio-vascular conditions or pain will be sufficient for this purpose.

5.4.4 Indication of fatigue and sleepiness

Epworth Sleepiness Scale (ESS)

The ESS is a quick test and designed to check the daytime sleepiness to see if the individual is getting enough sleep at night (Johns, 1991). Individuals are asked to score the likelihood of falling asleep in certain situations and an overall score as an indication of excessive daytime sleepiness is provided.

Measurement tool	Type of measurement tool	i-DREAMS languages	Time in minutes	Costs	Additional equipment needs	Validity
Epworth Sleepiness Scale	questionnaire	possible to translate	5	free	-	The questionnaire had a high level of internal consistency as measured by Cronbach's alpha (0.88). Reliability (r = 0.82).

Berlin Questionnaire

The Berlin questionnaire is a self-administered 10-items-questionnaire that was developed to identify subjects with obstructive sleep apnea (OSA) (Tan et al., 2017).

Measurement tool	Type of measurement tool	i-DREAMS languages		Time in minutes	Costs	Additional equipment needs	Validity
	questionnaire	DUT	ENG	5	free	-	

Berlin Questionnaire		FRE	GER				Scoring test-retest reliability: 0.74 – 0.98 (Cohen's kappa) (2-4). Scoring Internal consistency: 0.68 - 0.98 (Cronbach's alpha) (2-5)
		GRE	POR				

Stanford Sleepiness Scale (SSS)

The SSS is a subjective measure of sleepiness, frequently used for both research and clinical purposes (Connor et al., 2002). The questionnaire measures sleepiness at a specific moment in time, which can be relevant in simulator studies.

Karolinska Sleepiness Scale (KSS)

The KSS (Åkerstedt & Gillberg, 1990), is a 9-point, one dimensional scale ranging from ‘extremely alert’ to ‘very sleepy’, ‘great effort to stay awake’ and ‘fighting sleep’. It KSS has been validated against EEG variables (Kaida et al., 2006; Sagaspe et al., 2008) and is considered a reliable tool for evaluating sleepiness, both in a laboratory environment and in field studies (Åkerstedt, Anund, Axelsson & Kecklund, 2014).

Chalder Fatigue Scale (CFS)

The Chalder Fatigue Scale is designed to measure the severity of fatigue in adults. The 14-item instrument is indicated for use in both clinical and research settings. Symptoms examined by the scale can be divided into two categories: physical and mental fatigue (Chalder et al., 1993). The scale has been used in studies of various diseases and is explicitly not recommended for fitness to drive assessment.

In addition to standardized tests/questionnaires, items on overall sleep quality, working patterns and known sleep apnoea can be useful.

5.4.5 Vision impairment

Vision assessments are conducted, for example, with high and low contrast acuity charts or the so-called ‘Snellen E Chart’. The chart is placed on a wall at a 20 feet distance. Also, the participant’s driver licence may contain information on visual impairment. According to the ICD-10, normal vision is given when the Corrected Distance Visual Acuity (CDVA) ≥ 3/10.

5.4.6 Hearing impairment

An example of a hearing test is the Pure tone audiometry test. However, asking about known conditions or problems with hearing may be a simpler method.

5.5 Socio-demographic factors and driving experience

Obtaining socio-demographic variables is fairly straight-forward. Age, gender, nationality, issue date of driver license and alike are most efficiently asked in open questions. For level of education, socio-economic status and occupation, the provision of categories, from which the participants can choose from, is sensible.

However, ‘culture’ or ‘cultural identity’ are multifaceted constructs with various approaches for operationalisation. Cultural values are reported to be associated with road fatalities. An established and time saving survey to measure cultural values is the Short Schwartz’s Value

Survey (SSVS), with the value dimensions power, achievement, hedonism, stimulation, self-direction, universalism, benevolence, tradition, conformity and security (Lindeman & Verkasalo, 2005). Nationality may be a proxy of 'culture', however, does not account for sub-cultural phenomena.

Driving experience is a crucial factor for safety. It is reflected in the ability to detect hazards (Shinar & Oppenheim, 2011). It may also be accounted for by merely collecting the participants' average kilometrage and years of active driving.

5.6 Conclusion & recommendations

One outcome of the i-DREAMS project will be a research database with rich information of simulator and on-road drives of hundreds of participants. Since the database aims to facilitate future research, it can be argued that the more known about the test subjects, the better. However, this is a question of time, resources and also reasonableness towards participants volunteering to support i-DREAMS' research. Thus, a selection of variables for a one-time measurement before or during the trials has to be made. Table 12 displays all variables recommended for one-time measurement (second column), a suggestion how to measure (third column) them and the potential use of the variables.

Table 12: Driver characteristic variables recommended to collect from i-DREAMS participants, suggested measurement method and potential purpose of use in the project highlighted in green.

Category	Construct	Recommended measurement method	Potentially include in STZ	Validation of inter-individual differences in real-time measure, control variables	Potential for customized intervention
Competences	Emotional regulation	PERCI			
	Stress regulation	DSI			
	Attention regulation	Trail making test (Trail A)			
	Risk-taking	- SSS-V (sub-scale) - DBQ items			
	Hazard perception	GECO			
	Reactivity	RT			
	Visual perception, orientation	LVT			
Personality	Sensation seeking	BSSS			
	Anger proneness	DAS			
Past and habitual driving behaviour	Speeding	- DBQ subscale 'ordinary violations' - ESRA2			
	Tailgating	- DBQ item 23 - ESRA2			
	Fatigued driving	- ESRA2			
	Distracted driving	- ESRA2			
	Aggressive driving	- ADBQ subscale 'conflict behavior'			

Category	Construct	Recommended measurement method	Potentially include in STZ	Validation of inter-individual differences in real-time measure, control variables	Potential for customized intervention
		- DDDI, subscale 'aggressive driving'			
Health, diseases	Neurological	Question on any known conditions			
	Musculoskeletal	Question on any known conditions			
	Cardio-vascular	Question on any known conditions			
	Fatigue, sleepiness	- Epworth Sleepiness Scale - Berlin Questionnaire (Sleep apnoea, shift work, professional driving)			
	Vision impairment	- Snellen chart - Question on any known conditions			
	Hearing impairment	Question on any known conditions			
Socio-demographics	Age, gender, nationality	Open question			
	Level of education, socio-economic status, occupation	Closed question (provide response options)			
	Cultural identity	SSVS			
	Experience	Average kilometrage, issue date of driver licence, professional driving			

6 Driver behaviour – Indicators, methods and technology for real-time monitoring

6.1 Selection of the variables, technologies and metrics to review

The aim of this chapter is to have a comprehensive look at which variables are related to driver behaviour, which can be measured by on-board devices, which technologies provide these measurements, and which measured indicators/parameters provide useful information.

The starting point for the review is the Appendix 1.a of the proposal, where some of the variables, metrics and technologies to be addressed are highlighted. Table 13 summarizes this information with regard to operator behaviour. Specifically, for each ADAS the type of vehicle, road infrastructure, traffic and environmental conditions have been reported.

Table 13: Overview of variables, metrics and data collection tools as reported in Appendix 1-a of i-DREAMS proposal

	Variable	Metrics	Car	Bus	Truck	Train
Operator behaviour	Longitudinal movement	% exceedance speed limit	OBD/SP	OBD/SP	OBD/SP	SIM/OTMR
		ACC/DEC	OBD/SP	OBD/SP	OBD/SP	SIM/OTMR
		Longitudinal g-force	SP	SP	SP	NA
		Headway distance	MBE	MBE	MBE	NA
	Lateral movement	Lateral position	OBD/SP	OBD/SP	OBD/SP	NA
		Lateral g-force	SP	SP	SP	NA
		Edge line crossing	MBE	MBE	MBE	NA
where OBD means On Board Diagnostics, SP indicates Smart Phones, MBE stands for MoBilEye, SIM is the train simulator, OTMR indicates the On-Train Monitoring Recorder, while NA stands for Not Applicable.						

6.2 Reviewing method

A screen-based reviewing method has been worked out to analyse the state-of-the-art about data collection tools and their measurements.
four online databases were checked: Google Scholar, Science Direct, Scopus and Web of Science. The following keywords and combinations of them (Table 14) were searched:

Table 14: Keywords and their combination used during the screen-search

Measurable parameters longitudinal driving behaviour
automotive telematics
road speeding measurements
harsh braking measures
distance control system
adaptive cruise control
forward collision warning
objective measures OR in-vehicle device OR in-vehicle data OR CarChip OR global positioning system OR GPS OR on-board

diagnostic system OR OBDII OR intelligent speed adaptation design OR ISA device OR electronic device
(objective measures OR in-vehicle device OR in-vehicle data OR CarChip OR global positioning system OR GPS OR on-board diagnostic system OR OBDII OR intelligent speed adaptation design OR ISA device OR electronic device) AND (driving behaviour OR driving behavior OR driving exposure OR speed OR driving distance OR on-road behavio*)

A first selection of the papers was conducted by checking the title and then abstract. After this, the selected papers were read and 31 of them were taken into account for further evaluations.

The considered papers belong mainly to three different fields: safety, sustainable driving and insurance.

The majority of the papers focus on car drivers, with only a few of them also considering the issues related to light duty vehicles.

6.3 Advanced Driver Assistance Systems: a brief overview.

In accordance to (Yannis & Antoniou, 2000) (Wiethoff, Oei, Penttinen, Anttila, & Marchau, 2002) it is possible to recognize 2 different intervention levels and 3 distinct phases of the accident process. The two different intervention levels are the tactical and the operational level. The tactical level consists of executing manoeuvres, for example car-following, overtaking and intersection approaching (Wiethoff et al., 2002), whereas the operational level is related to the task of keeping the car on the road by selecting the suitable speed, steering etc. (Wiethoff et al., 2002). The three stages in which the accident process can be divided are: pre-crash, crash and post-crash (Yannis & Antoniou, 2000). Focusing on driving behaviour, the ADAS which will be reviewed belongs to the tactical level support functions.

Specifically, the ADAS which aid in longitudinal and lateral controlling are: Speed Control (ISA), Advanced Cruise Control (ACC) and Road Departure/Lane Departure Collision Avoidance.

Speed Control or Intelligent Speed Adaptation (ISA) ranges from recommending speed to reducing it, and is usually integrated with traffic control systems. As stated by Jesty et al. (1999, quoted from Wiethoff et al., 2002) an ISA system should provide information about various aspects of the road network, display the current speed limit and offer the ability to keep the vehicle below the speed limit.

Normally, this kind of system can be implemented on passenger vehicles as well as heavy vehicles and buses, and should work on each type of road. The technologies needed by an ISA system are sensors to measure vehicle’s speed and sensors to detect speed limits.

Advanced Cruise Control (ACC) works over 40 km/hr and aims at maintaining a safe separation between the vehicles by sensing the presence and relative speed of moving vehicles and adjusting the travelling speed (Yannis & Antoniou, 2000).

Also, this system can be implemented in both passenger and heavy vehicles and can work on all road types. Typically used tools are radar, lidar and video cameras.

Road departure/Lane departure collision avoidance’s goal is to assist the driver by detecting and tracking the lane or road edge and by determining the speed which the vehicle can be safely kept on the road (Yannis & Antoniou, 2000). Used devices are often radar, lidar, lasers, infrared and video sensors or combinations of the same (Tewolde, 2012; Thuy & León, 2010).

6.4 Data collection tools

During the review of the existing and available technologies to control driver behaviour, 15 different kinds of sensors and systems have been selected.

The technologies that are applied – not only in academic field, but also in the commercial one – are: cameras, smartphones with their embedded sensors (i.e. gyroscopes, accelerometers and magnetometers), the OBD-II system, GPS, radar, laser, lidar, steering angle sensors, thermal radiation sensors, infrared sensors, brake/gas pedal sensors, yaw rate sensors, digital tachograph, potentiometer and inertial sensors.

In the following section each tool will be briefly described, highlighting, where available, information relating to what the tool measures, its advantages and limitations, and the context where it has been applied.

Cameras represent a technology which is widely applied in Intelligent Transportation Systems (ITS). They are used to monitor the forward scene (Dagan, Mano, Stein, & Shashua, 2004; Stein, Mano, & Shashua, 2003), as well as to detect lane boundaries (Goldbeck & Huertgen, 1999).

In (Dagan et al., 2004) and (Stein et al., 2003) a forward-facing camera is used to create a combination of Forward Collision Warning, Lane Detection Warning and Headway Monitoring systems, and it provides range, relative speed and lane position data. Also, the inputs obtained by the camera are used to calculate Time-to-Collision (TTC), which is considered to efficiently indicate the changing from a normal driving situation to a dangerous one.

A CMOS video camera is used in (Goldbeck & Huertgen, 1999) to detect and track lane boundaries, by connecting it to the vehicle's CAN bus.

Also MobilEye (Dagan et al., 2004)("Mobileye Collision Avoidance System | Mobileye for Fleets," n.d.)(Ellison, Greaves, & Daniels, 2012) makes use of cameras for its IT system, especially using a forward facing camera to detect the headway to the lead vehicle.

Cameras can be installed in the majority of vehicles. Normally, cameras work on paved roads, where road markings are clearly visible ("Mobileye Collision Avoidance System | Mobileye for Fleets," n.d.) and bad weather conditions can affect the system's capabilities ("Mobileye Collision Avoidance System | Mobileye for Fleets," n.d.). In lane detection systems presented in (Goldbeck & Huertgen, 1999) authors assert that using CMOS cameras their system can also work during adverse weather conditions, i.e. diffuse daylight on dry asphalt, but also rain and back-lighting situations on wet pavement. Also, it works both on highways and country roads, even though one boundary leaves the scene (Goldbeck & Huertgen, 1999).

Using cameras has several advantages, including, the substantial amount of information obtained, the low cost and operation power, the absence of a sweep time and their non-invasive features (Wijesoma, Kodagoda, & Balasuriya, 2004). Limitations include poor visibility issues, such as shadows, bad lighting and weather. Also complex driving environments and missing lane markings make feature extraction challenging (Wijesoma et al., 2004) and overexposure can cause loss of information. Furthermore, the need of an exact and complex calibration has to be taken into account (Thuy & León, 2010).

Smartphones, with their embedded sensors, i.e. gyroscopes, accelerometers and magnetometers, are tools often used by insurers to control driving behaviour. In (Wahlstrom, Skog, & Handel, 2017) Wi-Fi, Bluetooth and GNSS smartphone sensors are used for positioning, so that where GNSS signals are unavailable, the others can be applied.

Skog et al., 2014; Wahlstrom et al., 2017; Wahlström, Skog, & Händel, 2015 deal with smartphone technologies for vehicle telematics and highlight the opportunity of measuring position, speed, course and timestamps, horizontal acceleration and vehicular angular velocity.

Regarding smartphones, context information is scarce. Also, since they are portable devices, they are more related to the person who carries them, than to the car. This implies that these devices are not directly linked to the car structure, well-fitting many vehicle types. Moreover, no indications about particular road infrastructures are reported in literature

Smartphone solutions are increasing in vehicle telematics because they are scalable, upgradable and cheap. Also, they can provide instantaneous driver feedbacks and have many embedded sensors. Issues that have to be considered are the low quality of the sensors, which are not primarily selected for vehicular measurements. Moreover, smartphones are not fixed, leading to issues as regarding relative orientation, driver/passenger recognition and GNSS coverage.

OBD-II is an in-vehicle sensor system, which is nowadays compulsory on most vehicles, designed to record data about the driving patterns and engine performance. In Rolim et al. (2014) it has been used with CarChipPro – a data logger – to collect vehicle dynamics. The data gathered include the number of trips, travel time, distance, speed, time in each speed band, brakes and accelerations, as well as mass air flow, engine speed, acceleration, fuel cut off and VSP distribution. In Wahlstrom et al. (2017) OBD-II is mentioned because data provided by this system can be sent to a smartphone, allowing an integrated network to be created. The study also adds to the already mentioned parameters throttle position and other engine performance magnitudes. In Cai et al. (2010) the OBD system is integrated with GPS and a wireless communication component to create a system warning the driver of dangerous situations. OBD measurements are also used in the calculation of fuel consumption and emissions by monitoring driving performance (D'Orey & Ferreira, 2014).OBD-II has been tested both in city contexts and on highway segments (Cai et al., 2010) and has been applied to both cars and light duty vehicles. In (Rolim et al., 2014) it has been integrated in vehicles driven by participants on their daily usual routes.

Among the important advantages of OBD-II there is the non-subjectivity of measurements to multipath errors and the use of fixed sensors (Wahlstrom et al., 2017).

Global Positioning System (GPS) is also a widespread utilized technology for ITS, both alone and in combination with other sensors. Ellison et al. (2012) applied GPS on participant cars driving on their normal routes, with the aim of measuring speeds, speed limits, location, date and time. (Greaves & Ellison, 2011) uses GPS devices on motorbikes and to investigate the willingness of motorists to speeding. In (Cai et al., 2010) GPS devices are used in combination with OBD system to test ABS activation alert, conducting experiments on both urban and highway roads. (Thuy & León, 2010) introduces a combined system of Differential Global Positioning System (DGPS), lidar and an IMU to develop a lane detection system, while (Ryu, Rossetter, & Gerdes, 2002) introduces a system based on GPS antennas to determine vehicle sideslip angle, longitudinal velocity and attitude. In (ElBatt, Goel, Holland, Krishnan, & Parikh, 2006) a system based on GPS and wireless communication devices is proposed to warn against forward collision.

From literature it emerges that GPS is widely used on various kinds of vehicles, from cars to motorcycles, and are mainly tested on daily routes.

Among the concerns of GPS devices installed on cars there is the need for power, and issues with privacy and signal jamming ("Advantages and Disadvantages of GPS and Telematics Systems for Vehicles - National Motorists Association," n.d.).

Radar, laser and lidar technologies are mainly used for lane detection and since they are quite similar, they can be discussed in the same frame. In (Thuy & León, 2010) a 1D lidar is used in combination with an IMU and DGPS to detect the lane, while in (Wijesoma et al., 2004) the authors present a radar system to detect curb position, together with a wheel encoder and a fiber-optic gyroscope to achieve vehicle speed and yaw angle respectively. In (Regan, Gustav Tingvall, Healy, & Williams, 2000) a commercially available radar device is used to transmit and receive signals in order to determine speed and distance of other vehicles/objects from the one, where it is installed. The 1-D lidar mounted on the vehicle's front bumper and sloped towards the street, introduced in (Thuy & León, 2010) is used to develop a lane detection and tracking system, while in (Wijesoma et al., 2004) a laser device is installed to detect road curbs. Furthermore, among radar and laser devices, distance sensors are quoted in the review carried out in (Lin et al., 2014). They are part of the system developed by UTDrive and can be grouped by long distance (30-100 m) – and short distance (0-30 m) sensors (Fleming, 2008). Long range distance sensors are normally radars or near-infrared lasers and are used in ACC and FCW systems (Fleming, 2008).

Radars, lasers and lidars have been applied in various contexts, from university campuses to other usual routes, and under different weather conditions (Regan et al., 2000; Wijesoma et al., 2004).

Radar's advantages are the high-quality images of the road scene over long distances (100m) and under various weather and light conditions. Ladars operate over moderate distances (80 m) compared to radars, but cost less, have an easy packaging, and their operating power is lower. Limitations of these sensors are the low resolution and slow scanning speeds and their tendency to be affected by extreme weather conditions.

Infrared sensors are grouped together with other already mentioned technologies (e.g. video and laser) in (Tewolde, 2012) as devices used in lane departure warning systems.

Thermal radiation sensors are quoted in (Xiao & Gao, 2010) as possible sensors to be integrated in active cruise control systems. In (Fleming, 2008) they are mentioned for pedestrian detection. Indeed, they detect nonvisible long-IR wavelengths emitted by warm body objects at great distances (Fleming, 2008).

Two kinds **of steering wheel sensors** exist: absolute and relative, which differ in the way they measure the steering angle. Absolute sensors estimate it, while the relative sensors "learn" the position of the steering wheel (Tseng, Ashrafi, Madau, Brown, & Recker, 1999). In (Fleming, 2008), a dual magnet steering wheel angle sensor is quoted and applications for vehicle stability control, parking assist and road navigation are mentioned. In (Wijesoma et al., 2004) wheel sensors are used to obtain vehicle speed and, in combination to a laser device and a fiber-optic gyroscope, able to provide yaw angle measurements, they form a system to detect road boundaries.

Brake/gas pedal sensors are used to achieve pedal position and brake cylinder pressure to understand how hard the pedal was pressed. In (Feng et al., 2017) they are integrated with transmission output speed sensors, which are sensors linked to CAN bus, to derive the jerk. Also, in (Lin et al., 2014) speed sensors are used and linked to GPS device and accelerometers.

Yaw rate sensors are devices belonging to curve sensors (Xiao & Gao, 2010). They assist the driver in recognizing the highway course and are especially useful along curves.

A **digital tachograph** is a device which imports driving records from the OBD system and is compulsory on commercial vehicles driving in some countries, e.g. Korea (Lee & Jang, 2017). In Lee & Jang (2017) the authors use this tool to obtain speed, acceleration and yaw rate of taxis driving in metropolitan cities.

In (Espinosa et al., 2011) a **potentiometer** is used to control the movement of the pedals (throttle, brake, clutch) in order to have insights in driver behaviour. This sensor is integrated in a system of several devices used to monitor the triad vehicle-driver-environment with the aim of linking it to vehicular emissions.

Inertial sensors are units made up of gyroscopes and accelerometers, which are commonly linked to other tools in a sensing network. In (Espinosa et al., 2011) they are used together with OBD, potentiometers, cameras and GPS to obtain data about vehicle-driver behaviour and environment influencing emissions.

In Table 15, pros and cons of the six most applied technologies are recalled.

Table 15 Advantages and Disadvantages of the most applied technologies

Technology, sensors	Advantages	Disadvantages
Cameras	High information content, low cost and operation power, absence of a sweep time, passive non-invasive sensor	Shadows, complex driving environments, missing lane markings, low signal-to-noise ratio of images under poor lightening, visibility, bad weather making feature extraction challenging, loss of information in case of overexposure, need of exact and complex calibration
Smartphones	Scalable, upgradable and cheap; instantaneous driver feedbacks; a lot of embedded sensors	Low quality sensors, not fixed position, i.e. relative orientation issues, driver/passenger recognition, GNSS coverage
OBD-II	Compulsory on vehicles; several engine and dynamics data; non-subjected to multipath errors; use only of fixed sensors	Variations of the system according to different vehicle types, presence of a single OBD-II port in most vehicles, higher costs than smartphones, compatibility problems with vehicles released before 1996.
Global Positioning System (GPS)	Widely available and applied.	Need of power; privacy issues; signal jamming
Radar	High-quality images pf the road scene over long distances (100m), in snow, haze, dust, rain, not susceptible to ambient light	Difficult to detect small and/or static objects, difficulties in closed environments (tunnels, etc…), interference with other radars, time for warnings is rather high.
Laser, lidar	Operate over moderate distances (80 m), lower cost, easy packaging, lower operating power, signal cutter and size considerations	Low resolution, slow scanning speeds, not usable under extreme weather conditions

6.5 Measurements to define driver behaviour (other)

In relation to the already described devices, the measurements they provide have also been reviewed. The aim of this section is to understand which factors representing driving behaviour can be detected and can give useful data to manage safety. Therefore, only measurements linked to this kind of behaviour are summarized. The same technologies can also provide other data typologies, whose review is out of the scope of this work.

Generally, around 30 different measurements regarding safety have been reviewed, and can be summarized as: speed, trajectory, acceleration, steering angle, yaw rate, headway, lane position, speed limits and travel information.

As noted in SWOV (2000), the most appropriate measurement methods are vehicle control (longitudinal and lateral), time headway, time to collision, speed and lane keeping.

In the following section, the measurements are classified on the basis of the technologies to measure them.

OBD-II can provide the number of trips, travel time, distance, speed, time in each speed band, time spent over the speed limit (Kaye, Lewis, & Freeman, 2018), use of brakes and acceleration (Rolim et al., 2014).

Forward-facing cameras can measure data such as range, relative speed and lane position. They can also detect lane markings and road edges, as well as measuring the distance of the vehicle from the boundaries (Dagan et al., 2004). Systems like MobilEye, which uses a forward-facing camera, also calculate TTC in order to trigger the warnings.

Smartphones can provide measures as three-dimensional position, planar speed, course and timestamps, longitudinal and lateral acceleration and angular velocity (Skog et al., 2014; Wahlstrom et al., 2017; Wahlström et al., 2015).

GPS devices obtain driving speed, speed limits, sideslip angle, attitude, roll, yaw, distance, location, data and time (Ellison et al., 2012; Ryu et al., 2002).

Lidar and radar sensors provide position and distance information (Thuy & León, 2010).

Steering angle/wheel sensors provide steering wheel angle, speeds and lateral acceleration (Hac & Simpson, 2000). Infrared sensors generally measure distance, whereas brake/gas pedal sensors provide pedal position and brake cylinder pressure. Potentiometers provide information relating to the position of throttle, brake and clutch pedals, while yaw rate sensors provide yaw rate measurements. The use of a digital tachograph is possible to obtain speed, acceleration and yaw rate and inertial sensors can determine acceleration, velocity, displacement, angular rates, and rotation angles.

In Table 16 the measurements provided by the various sensors are summarised.

Table 16 Measurements provided by the listed sensors

TECH-NOLOGY	**Cameras**	**Smartphones**	**OBD-II**	**GPS**	**Radar**	**Ladar**	**Steering wheel sensors**	**Infrared sensors**	**brake/gas pedal sensors**	**yaw rate sensors**	**digital tachograph**	**potentiometer**	**inertial sensor**
Number of trips			•	•									
Travel time			•	•									
Travel distance			•	•									
Position		•		•	•	•							
Headway	•				•	•							
Speed	•	•	•	•			•				•		•
Speed limit	•			•									
Time over speed limit	•		•	•									
Brakes and acceleration		•	•						•		•	•	•
lane position	•				•	•							
Road markings/edges	•				•	•							
Angular velocity		•											•

Sideslip angle				•									
Roll, yaw				•					•	•	•		
Data & time	•	•	•	•									
steering wheel angle							•						
Lateral acceleration		•					•						
Distance from obstacles	•				•	•		•					

After reviewing the available sensors, it is possible to classify them in relation to the kind of movement they can provide data about, i.e. longitudinal or lateral movement, and the metrics they belong to, i.e. speed regulation, acceleration/deceleration.

This classification is briefly summarised in Table 17 and Table 18.

Table 17 Classification of the sensors on the basis of the initial selected variables and metrics – longitudinal movement

Longitudinal movement	
Speed regulation	GPS, radar, cameras, On Board Diagnostics (OBD-II), smartphones, digital tachograph, steering wheel sensors
Acceleration/deceleration	OBD-II, smartphones, cameras, steering wheel sensors, brake/gas pedal sensors, digital tachograph, potentiometer, GPS
Longitudinal g-force	OBD-II, smartphones
Headway distance	cameras, radar, OBD-II, GPS

Table 18 Classification of the sensors on the basis of the initial selected variables and metrics – lateral movement

Lateral movement	
Lateral position	smartphones, lidar, radar, cameras.
Lateral g-force	OBD-II
Edge line crossing	cameras, GPS

6.6 Technologies available in the consortium and provided measurements

Inside the Consortium developing i-DREAMS project, cardioID and OSeven are the partners which provide the sensors to detect driving behaviour. Specifically, CardioID provides MobilEye technology while OSeven makes available its smartphone solutions.

MobilEye solution is a forward-facing camera, which alerts drivers when an imminent rear-end collision may happen (FCW), helps to keep a safe following distance (HMW), warns drivers about unintentional lane departures (LDW), notifies the driver if a vulnerable road user is in the danger zone (PCW), and it provides indications about the detected speed limit signs.

OSeven solutions are smartphone-based technologies, which can provide trip data such as duration and distance of the trip, speeding, mobile use while driving, harsh brakes, accelerations and cornerings, driving in risky hours and transportation mode detection. On the basis of the obtained data a driving behaviour related scoring model is also developed.

6.7 Driver behaviour indicators and thresholds

The aim of this section is to give a first insight into indicators characterising driver behaviour and the thresholds specifying the passage from normal to dangerous situations.

Longitudinal movement, acceleration, time headway and TTC are the main utilized indicators for defining driver behaviour and thresholds. Also, variations of TTC are commonly used, such as TET, TIT and TTZ. Finally, PET and MTC have to also be considered.

According to Eboli et al. (2016), acceleration values close to ±6 m/s^2 are safe for speed near to zero. On the other hand, at speed close to 100 km/h, threshold values are about ±2 m/s^2. In (Stipancic et al., 2018), acceleration thresholds of ± 2 m/s^2 and buffers of 200m for intersections were chosen to develop GPS-based safety indicators.

Vlahogianni & Barmpounakis (2017) using an OBD-II device, proposed the following values for thresholds: harsh braking 0.21 g, harsh acceleration 0.18 g and harsh cornering (left/right) ±0.18 g. (Skog et al., 2014) uses a detection threshold for harsh braking of -2 m/s^2. In (Rolim et al., 2014) the following ranges and descriptions are proposed (Table 19).

Table 19 events and thresholds as proposed in (Rolim et al., 2014).

Events	Description
Hard brake	0.34G<Brake force<0.51G
Extreme brake	Brake force>0.51G
Hard acceleration	0.31G<Acceleration Force<0.45G
Extreme acceleration	Acceleration Force>0.45G

Moreover, deceleration rate to avoid crash (DRAC) is recognized as a safety performance indicator, as it considers the role of differential speeds and decelerations in risk avoidance. DRAC is the differential speed between a following vehicle and corresponding lead vehicle divided by their closing time. A threshold of 3.4 m/s^2 is considered as a cut-off value (Shi et al., 2018).

Another widely used indicator is Time Headway (H), which is the elapsed time between the front of the lead vehicle passing a point on the roadway and the front of the following vehicle passing the same point (Mahmud et al., 2017). In the US, it has been found that it is impossible to follow a vehicle safely with headway of <2 s. In Germany, the recommended minimum distance is "half the speedometer", which means, a car travelling at 80 km/h should keep a distance of at least 40 m. This rule translates to recommended time headway of 1.8 s. The values between 1.0–1.5 seconds are the most often cited in literature. Yannis et al. (2004) found through a driving simulator experiment that the average distance headway on rural road sections and on two-lane motorway segments is 51 m without the use of any ACC, and time headway stands in the range 1.4-2.8 s. In (Michael, Leeming, & Dwyer, 2000) a study to infer the average values of headways in urban areas has been developed. Starting from the consideration that the 2 s rule pointed out by various driver training programs is normal too large, and that headways of 4s indicates isolated vehicles, the authors measured by video footage the headway between following cars in three situations: normal driving situation, without any intervention, and two situations where the participants were shown different signs.

They found that in normal driving on urban roads, average headway values are set in the range 1,4 – 2,2 s. In (Ayres, Li, Schleuning, & Young, 2001) speed and time headway data are collected by means of loop detectors in both free flow and rush hour conditions. During free-flow situations a variable time-headway with a lower bound set at 1s was noticed, while during rush hours the values of the time-headway set in the range 1-2 s and that this range keeps constant with vehicle speeds from 20 to 60 mph (32 – 97 km/h). However, many road administrations in European countries recommend a safe headway of 2 s. (MobilEye, n.d.-b) and (MobilEye, n.d.-a) report the thresholds kept by MobilEye technology. Specifically, the headway monitoring system, which works at speeds higher than 30 km/h, gives notice when the time distance from the front vehicle equals 2.5 s, and returns a warning when it is 0.6 s. The forward collision system provides an alert up to 2.7 s before a possible collision with the front vehicle, while the lane departure system is active above 65 km/h. For the Speed Limit Indicator Option, the user can manually change the thresholds of the alerts, letting the notice be issued when the speed exceeds the speed limit respectively by 5,10,15, 20, 25 or 30 km/h.

Temporal and spatial proximity can be used to evaluate high-risk driving behaviours, such as excessive speeding, driving too close to the preceding vehicle etc. Set of commonly used indicators in studies are referred as proximal indicators. These indicators define closeness of other vehicles or road users in relation to possible point of collision. The main advantage of proximal indicators is that they occur frequently.

Among time-based indicators, Time to Collision (TTC) calculates the time remaining before the collision, if the involved road users continue with their respective speeds and trajectories (Mahmud et al., 2017). TTC is a continuous variable that can only be calculated while the road users remain on the collision course. TTCs higher than 5 seconds are not very feasible for traffic safety research, whereas 1.5 seconds may be suitable as a cut-off value beyond which driving becomes unsafe (Niezgoda et al., 2012). Yannis et al. (2004) sum up the thresholds reported by SIMONE and SISTM, underlining that values < 3s are considered in SIMONE as uncomfortable situations, and TTC < 1.5 s as dangerous situations, while the micro-simulation package SISTM reports all TTC < 10s and TTC > 1s. Also (Laureshyn, Goede, Saunier, & Fyhri, 2017) cites thresholds for TTC, indicating dangerous situations to be the ones when TTC values are < 1.5 s. In (Graham & Hirst, 1994) a TTC value of 3 s has been suggested as suitable values, warning about a TC=4-5 s, which seems to provide too many alarms if applied to Collision Warning Systems. In (Hogema & Janssen, 1996), a minimum TTC value of 3.5s has been reported for non-supported drivers and 2.6 s for supported ones. This last value is also the one regarded as a safety concern. Finally, (Van der Horst & Hogema, 1994) confirms a TTC of 2.5 s and a minimum TTC value of 1.1 s. In (Lamble, Laakso, & Summala, 1999) a study about the influence of phone dialling and cognitive tasks in comparison to normal driving is developed. As measures to evaluate the different performances TTC and Brake Reaction Time (BRT) have been considered. In addition to the differences among control task and phone/cognitive tasks, he authors found that TTC stands in the range [11;11.5]s, when drivers are not distracted, while BRT belongs to the range [3.5;4]s. These performances increase in both distraction tasks, i.e. a growth of 0.62 s and 0.95 s happens in TTC for dialling and cognitive tasks respectively, and an increase equal to 0.48 s and 0.50 s has been measured for BRT. Based on TTC, Time Exposed TTC (TET) and Time Integrated TTC (TIT) are further used to measure risk duration and risk integration, respectively. TET expresses the total time of a vehicle exposed in risk situations (the length of time a TTC-event remains below a designated TTC-threshold). TIT is the integral part of the TTC-profile during the time below the threshold (Mahmud et al., 2017).

Another variation of TTC measure is Time-to-Zebra (TTZ), which aims to estimate traffic safety at pedestrian's crossings and analyse if drivers regard pedestrians not on the crossing as a potential danger risk. TTZ value is defined as the distance to the zebra crossing divided by the speed. It can be calculated as the time left for the car to the zebra crossing at the moment the pedestrian arrives at the curb (Niezgoda et al., 2012). In particular, for TTZ ≤ 3 s, the driver is approaching the pedestrian crossing with high speed values and adopts the most abrupt speed reductions. This behaviour highlights a certain driver's aggressiveness. For TTZ values belonging to the range (3;5 s), the driver adopts lower speed and less abrupt speed reductions than those shown for TTZ≤ 3 s (Bella & Silvestri, 2015).

Time-to-Accident (TA) is a special value of the TTC, based on evasive action undertaken by any of the road user to avoid collision. TA is the time that remains to an accident from the moment that one of the road users starts an evasive action, if they had continued with unchanged speed and directions. A TA value of 1.5 s is used to distinguish serious conflict and slight conflict (Niezgoda et al., 2012).

Another commonly used value is Post-Encroachment-Time (PET). It is a calculation of the time difference between the passages of two road users with a common area of potential collision (common spatial point) that is below a given cut-off value (Niezgoda et al., 2012). PET is calculated as the time between the moment when the first road user leaves the path of the second and the moment when the second reaches the path of the first (i.e. the PET indicates the extent to which they miss each other). Typically, it is assumed that the threshold values 1.0 s or 1.5 s are considered critical (Mahmud et al., 2017).

Margin to Collision (MTC) is a distance based proximal indicator that represents the possibility of collision in a case where the preceding vehicle and the following vehicle decelerate abruptly at the same time (the abrupt deceleration is assumed to be 0.7 g). MTC is the ratio of the summation of the inter-vehicular distance and the stopping distance of the preceding vehicle divided by the stopping distance of the following vehicle. MTC of <1 indicates a high likelihood of collision in a case where the preceding vehicle decelerates abruptly, even if the following vehicle at the same time also decelerates abruptly (Mahmud et al., 2017).

Among traffic safety proximal indicators, there are several indicators concerning lateral behaviour of a driver. Lateral behaviour, also called lane keeping, describes driving performance, but it is very difficult to indicate where limits between safet and unsafe behaviours are. Two of the most common lane keeping indicators are: standard deviation of lane position (SDLP) and Time-To-Line-Crossing (TLC). SDLP reflects the degree of vehicular control a driver exerts in any particular driving situation (Martens et al., 2011). TLC is defined as the time it takes to reach the lane marking, assuming fixed steering angle and a constant speed. It is accepted, that TLC < 1s implies an increased safety risk. TLC indicates that a lane exceed is likely to occur within a short time frame and therefore detects a possible risk before the lane exceed actually occurs (Niezgoda et al., 2012).

Table 20 summarizes the basic driver behaviour indicators and thresholds.

Table 20 : Literature synthesis for driver behaviour indicators and thresholds

Type of indicator	Study	Indicators	Thresholds	Suggested STZ phase
Acceleration /deceleration /cornering indicators	Vlahogianni & Barmpounakis (2017)	Harsh Brakings	2.1 m/s^2	Avoidable crash
		Harsh Accelerations	1.8 m/s^2	
		Harsh Cornerings	1.8 m/s^2	
	Eboli et al. (2016)	Acceleration at speed 100km/h	± 2 m/s^2	Normal/dangerous driving
		Acceleration at speed near zero	± 6 m/s^2	Normal driving
	Stipancic et al. (2018)	Acceleration	± 2 m/s^2	Dangerous driving
	Shi et al. (2018)	Deceleration Rate to Avoid the Crash (DRAC)	3.4 m/s^2	Dangerous driving/avoidable crash
Time based indicators	Niezgoda et al. (2012)	Time to Collision (TTC)	1.5 s	Avoidable crash
		Time to Accident (TA)	1.5 s	
		Time-To-Line-Crossing (TLC)	<1 s	
	Mahmud et al. (2017)	Time Headway (H)	<2 s	Normal driving
		Post-Encroachment-Time (PET)	1.0-1.5 s	Avoidable crash
	Bella & Silvestri (2015)	Time-to-Zebra (TTZ)	1-4 s	Normal/dangerous driving
Distance based indicators	Mahmud et al. (2017)	Margin to Collision (MTC)	<1	Avoidable crash
Lateral behaviour indicators	Martens et al. (2011)	Standard deviation of lateral position (SDLP)	>0.25 cm	

Finally, some studies are reported, which tried to give an insight into the behavioural pattern of normal driving.

In (André, 2004) driving conditions, represented by running and average speed (km/h) and average positive acceleration (m/s^2), are obtained by analysing and clustering a set of data from France, UK, Germany and Greece. Congested urban conditions, free-flow urban situations, secondary roads, main roads and motorways have been considered. Table 21 shows the maximum values reported by the authors for each category.

Table 21 Classes of driving conditions as introduced in (André, 2004).

Classes of driving conditions	Running speed (km/h)	Average speed (km/h)	Average positive acceleration (m/s2)
Congested urban	25.9	15.9	0.87
Free-flow urban	35.6	32.3	0.81

Secondary roads	65.0	64.0	0.75
Main roads	86.1	85.7	0.67
Motorways	123.8	123.7	0.53

Research relating to driving cycles provide information about driving patterns in normal conditions.

In (Amirjamshidi & Roorda, 2015) driving cycles for commercial vehicles in the Toronto area, including arterial, collector, local roads and two freeways, are worked out. The parameters considered for the assessment of the cycles are average speed, average running speed, average acceleration and deceleration, time proportion of driving modes in idling, accelerating, decelerating, cruising and creeping, average micro-trip duration, average percentage of acceleration-deceleration changes, root mean square acceleration, root mean square of positive kinetic energy overweight. In Table 22 running speed, average velocity and acceleration/deceleration are summarized.

Table 22 Assessment measures for driving cycles developed in different locations.

Assessment measure	**Simulated Toronto LDT driving cycle**	**Beijing**	**Shanghai**	**Chongqing**	**Tianjin**	**Chengdu**	**Hong Kong1 driving cycle**	**US HWFET**	**Artemis freeway**	**WLTC Class 2 high phase**	**WLTC Class 3 extra high phases**
V (km/h)	52.7	25	38	36	29	50	38.3	77.1	97.0	54.1	71.4
Vr (km/h)	54.0	27	41	37	35	51	41.8	77.7	98.3	57.8	74.9
Acc (m/s2)	0.293	0.43	0.48	0.44	0.34	0.39	0.398	0.288	0.408	0.238	0.372
Dec (m/s2)	-0.544	-0.42	-0.51	-0.48	-0.45	-0.43	-0.414	-0.383	-0.496	-0.323	-0.381

The authors of (Ho, Wong, & Chang, 2014) develop driving cycles for passenger cars driving in Singapore and compared them to the New European Driving Cycle (NEDC). The measures they used to compare the cycles are maximum and average speed, and the percentage of driving modes – idling, acceleration, deceleration and cruising (Table 23).

Table 23 Maximum and average speed as reported by (Ho, Wong, & Chang, 2014).

Driving cycle	Maximum speed [km/h]	Average speed [km/h]
SDC – expressway	88.0	49.7
SDC – arterial road	77.0	33.7
NEDC – expressway	120.0	69.4
NEDC – arterial road	50.0	27.2

In (Li, Xiong, Wang, & Lu, 2015) 3 scenarios determined by TTC values on three different road typologies (urban access roads, urban distributor roads and freeways) in Bejing were studied: a) stable car following; b) unstable car following and c) car-approaching. Speed, distance headway (DHW), time headway (THW), the inverse of TTC (TTCi), DHW (accelerator release and brake activation) and TTC (accelerator release and brake activation) are considered for the evaluation (Table 24).

Table 24 Control parameters and results found by (Li, Xiong, Wang, & Lu, 2015).

Control parameter	Speed [km/h]	DHW [m]	DHW accelerator release [m]	DHW brake activation [m]	THW [s]	TTCi [s^-1]	TTC accelerator release [s]	TTC brake activation [s]
A – urban access roads	47.14+-3.82	24.08+-4.48	-	-	1.83+-0.32	0.012+-0.030	-	-
B – urban access roads	45.68+-3.95	22.60+-3.92	-	-	1.78+-0.29	0.039+-0.011	-	-
C – urban access roads	-	-	25.29+-3.93	22.24+-3.78	-	-	19.05+-3.15	14.20+-3.80
A – urban distributor	68.98+-4.30	33.46+-6.50	-	-	1.72+-0.27	0.024+-0.035	-	-
B – urban distributor	68.18+-6.20	31.67+-6.17	-	-	1.64+-0.24	0.085+-0.012	-	-
C – urban distributor	-	-	33.50+-6.80	27.37+-8.02	-	-	21.96+-3.46	14.68+-3.35
A – freeways	94.65+-8.25	47.86+-9.75	-	-	1.80+-0.36	0.025+-0.033	-	-
B – freeways	91.49+-13.16	45.19+-9.41	-	-	1.77+-0.32	0.093+-0.087	-	-
C - freeways	-	-	44.60+-7.65	34.16+-9.16	-	-	20.04+-2.59	14.64+-5.71

6.8 Conclusions

In total 15 tools have been found to be used in literature. From the most frequently used to the least frequently used there are: cameras, smartphones, OBD-II, GPS, radar, lidar, laser, steering angle sensors, distance sensors, brake and gas pedal sensors, speed sensors, yaw rate sensors, thermal radiation sensors, infrared sensors, digital tachograph, potentiometer, inertial sensors.

More than 30 different direct measurements have been reviewed. Of these, the 14 most reported are: speed, trajectory, acceleration, latitude and longitude, jerk, acceleration/brake pedal status, steering angle, yaw rate, time/distance headway, lane position, speed limits, time over speed limit and travel information.

As indirectly calculated parameters, Time-To-Collision, its inverse and Tres have been also cited.

Considering both the reviewed technologies and the ones owned by the consortium, it can be concluded that all the available and most promising measurements are covered by the technologies provided by CardioID and Oseven, providing a reliable insight in driving behaviour characteristics.

7 Conclusion and considerations for the next steps

The aim of the work documented in this report was to review and assess state-of-the-art approaches and methods to monitor the driver's mental state and contextual factors of the driving environment that impact task demand. In addition, a selection of driver trait factors including measurement methods were summarized and driver behaviour indicators were reviewed.

The vast majority of reviewed literature and information concerns car driving. An assessment was conducted to see what extent the conclusions are transferable to other modes. Apart from the driving behaviour indicators and corresponding threshold ranges, no indication was found that contradict the assumption that the identified methods and indicators can be transferred from the context car to the other i-DREAMS modes: trucks, buses, trains and trams. However, the particular situation of professional driving should be borne in mind at all times.

One of the main conclusions that can be drawn is that two physiological/behavioural measurement methods should be used for the continuous driver monitoring. This insight applies to measuring all of the single constructs: task demand attention and distraction, fatigue and sleepiness as well as emotions and other related constructs. By using two measures, the drawbacks of a single measurement methods can be compensated for.

The benefit of using CardioWheel is that a real-time measure of ECG can be obtained, with equipment available in the consortium. However, heart rate and heart rate variability is sensitive to inter-individual differences and confounding factors. An additional dashboard mounted eye tracking system is beneficial for measuring task demand, sleepiness and fatigue. A thermal or standard camera for facial feature tracking would support emotions monitoring as well as sleepiness and distraction detection. Another consideration is using a wrist band with EDA sensors, which also supports emotion detection. However, the impact on the naturalistic driving character has to be considered when asking the participants to wear a device whenever they drive. When using cameras facing the participant, GDPR is to be considered carefully.

Due to the circumstance that each measuring method described is not the one and only standard in research, a thorough testing in the simulator stage is indispensable.

Considerations for the single constructs are summarized hereafter.

Task demand

Self-reported task demand, driving performance measures and physiological measures are three main approaches to measuring task demand. The most frequently used method in reviewed literature is the use of physiological and behavioural indicators, where the number and duration of eye fixations as well as ECG measures are indicated to be the most reliable ones. Although ECG can be captured by the consortium's cardio wheel (sensors on the steering wheel), a supplementary eye tracking system might prove beneficial. With regard to driving performance, lateral position deviation and reduced speed are important indicators of increased task demand. Both can be detected by means of mobile phones. However, physiological and behavioural measures[36] are to be prioritized over driving performance measures to measure task demand.

[36] Physiological measures refer to the activity of the autonomous nervous system, for example the heartbeat. This activity cannot (or hardly) be controlled by an individual whereas behavioural measures refer to the movement of body parts that can actively be controlled, such as the eyes and facial expression (which, however, are not always controlled consciously).

Attention and distraction

The best-studied forms of distraction are visual but also cognitive distraction. Eye tracking systems and cameras are most frequently used to measure attention and distraction, with head position, viewing and scanning patterns and PERCLOS (percentage of time that eyelid covers ≥80% of the pupil) being the most reliable indicators. Driver behaviour indicators are focused on lateral and longitudinal control measurements. With regards to the i-DREAMS system, a dashboard eye tracking system or camera might prove beneficial for measuring attention and distraction in real-time. To measure distraction due to mobile phone use, smartphone sensors which detect the movement of the phone can also be beneficial and are non-intrusive.

Fatigue and sleepiness

The literature on measuring fatigue and sleepiness indicates that eye tracking is the most commonly used measure of sleepiness with blink rate and PERCLOS shown to be the most robust indicators in terms of ocular measures. Heart rate and heart rate variability also shows potential and can be developed as a minimally invasive technique. Utilising multiple measures and indicators could help to improve the reliability of sleepiness detection.

Emotions and stress

Research designs of reviewed studies on measuring emotions and related constructs are very heterogeneous with a broad variety of underlying theoretical assumptions regarding the operationalization of 'emotions', due to a lack of a standard definition. However, anger, frustration, aggression, stress as well as fear and anxiety appeared to be the most studied emotional categories, composed of combined arousal and valence levels. EDA and heart-based measures are most frequently used to measure emotions. A combination of two physiological measures is advised. Complementing the ECG measure of the Cardio Wheel with a wrist worn EDA sensor or a (thermal) camera for facial feature tracking might prove beneficial.

Substance impairment

Conventional methods to detect substance impairment – blood tests, saliva tests, urine samples and breathalysers – are reliable, especially to infer impairment due to alcohol, but not applicable for i-DREAMS since they do not allow for continuous monitoring. Wearable sensor technologies using touch-based, breath-based and ocular measures are still under development or have not been validated, respectively. Wrist-worn transdermal alcohol sensors have more potential to be used within i-DREAMS. Impairment by drugs and medicine is much less understood than alcohol impairment and thus, there is no common understanding of real-time measurement. Regardless of the measurement methods and their quality, practical considerations for implementations in i-DREAMS should be noted. Although impairment may be measurable in real-time with increasing reliability, the effects of the specific impairment may be expressed in impaired attention and alertness and thus, already accounted for by the corresponding real-time measurements. This should be borne in mind for the model of the safety tolerance zone.

Driver characteristics

Collecting further information about the i-DREAMS participants serves various goals in the project: populating the i-DREAMS research data base, customizing interventions, accounting for covariates and possibly introducing stable factors into the Safety Tolerance Zone model as correction factors. The driver characteristic variables will be subject to a one-time measurement, likely before starting the trials. The most efficient methods to collect this data is surveying. However, it is recommended to contemplate using a few additional performance tests such as measuring attentional regulation capabilities. Table 12 provides an overview of variables suggested to account for.

Driver behaviour indicators

The most commonly used systems to measure driver behaviour are cameras, smartphones, OBD-II, GPS, radar, lidar, laser, steering angle sensors, distance sensors, brake and gas pedal sensors, speed sensors and yaw rate sensors with the indicators speed, trajectory, acceleration, latitude and longitude, jerk, acceleration/brake pedal status, steering angle, yaw rate, time/distance headway, lane position and time over speed limit and travel information. An important indirect parameter is Time-To-Collision and its inverse. It can be concluded that the most promising measurements are covered by the equipment already available to the consortium.

7.1 Recommendations for measurement

Revisiting the starting position and the initial project outline, some modifications had to be made to arrive with a set of relevant factors measurable in real time or one-off. Figure 6 provides the adapted version of the operator related measures which ideally should be considered in a conceptual model of the safety tolerance zone.

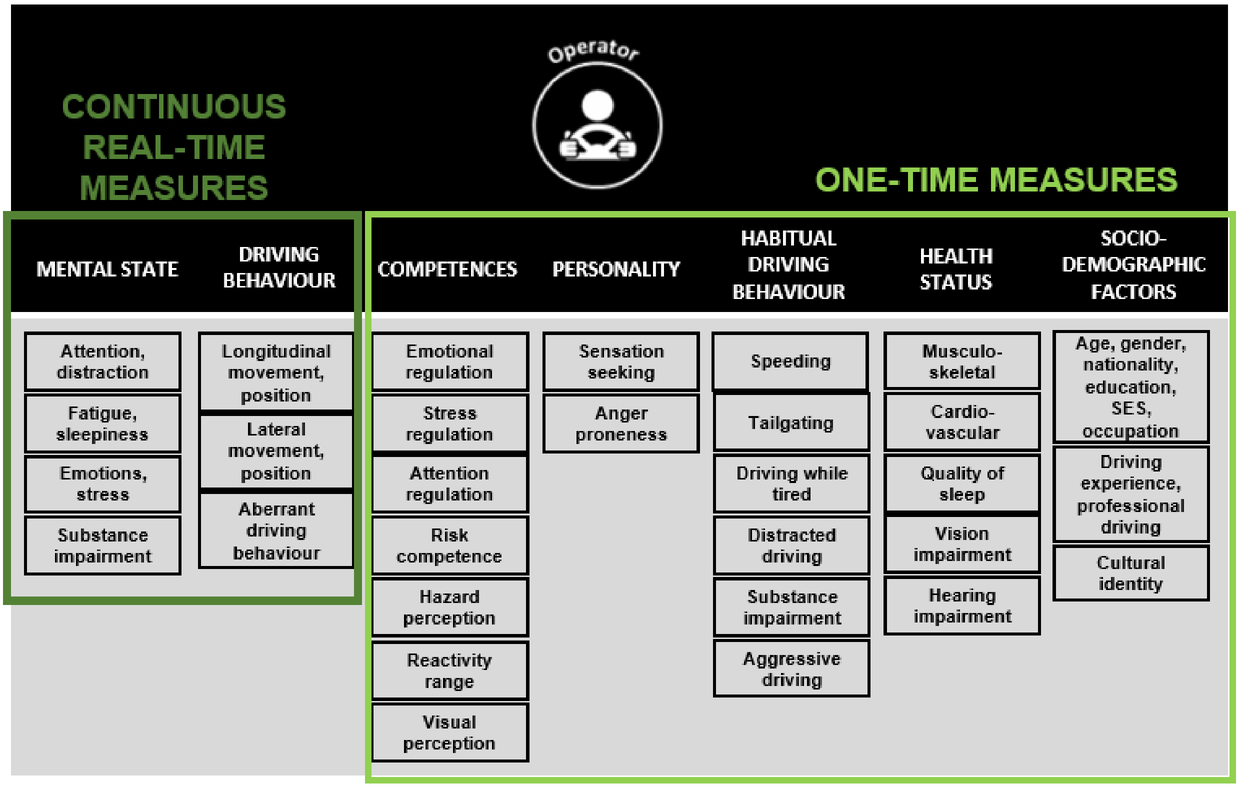

Figure 6: Adapted figure of factors ideally to be considered in the STZ after review of scientific literature.

In general, the most important conclusions and recommendations for i-DREAMS include:

- Most of the evidence is available for car drivers. The transferability of some of the findings to trucks, busses, trams and trains may partly be determined in an iterative process and by actual trial and error
- 'Mental state', 'emotions', 'distraction' etc. are theoretical constructs that ask for deciding on one of the plethora of definitions and theoretical concepts.
- Using at least two approaches for driver state monitoring will be beneficial for assuring validity and reliability
- Majority of driver mental state variables could be measured with cameras, eye tracking, and heart rate sensors either embedded in the steering wheel or incorporated into It should be considered that the use of devices that have to be put on or activated by the participant before driving may compromise the naturalistic driving character of the trials.
- The potential to consider the drivers' traits and characteristics in the calculation of the safety tolerance zone should be explored.
- Thoroughly testing indicators and measures at the simulator stage is indispensable

The table below summarises the operator states, and the recommended measures, technology and thresholds for use when monitoring driver task complexity and coping capacity, based on the systematic review of scientific literature.

Table 25: Summary of factors to be considered for modelling the driver state and corresponding indicators

Operator state	**Optimal measure**	**Ideal technology**	**Influence on coping capacity/ task demand**	**Safety critical threshold**	**Frequency of measure (real time or one-off)**
Attention and distraction	- PERCLOS - PERLOOK - Glance duration - Head movement - driver behaviour (lateral and longitudinal measures, reaction time, gap acceptance)	- Eye tracker (glasses / system) - Driver facing camera - Forward facing camera and collision avoidance system (Mobileye)	Increased PERCLOS, PERLOOK, glance duration, head movements = increased distraction and reduced coping capacity.	- PERCLOS and PERLOOK > 35% - Glace duration of 2 seconds - Head turns > 5 seconds	Real time
Alertness (fatigue / sleepiness)	- Blink rate - PERCLOS - Heart rate variability (HRV)	- Eye tracker (glasses / system) - Driver facing camera - Heart rate sensors embedded in steering wheel (CardioWheel) - Wearable heart rate monitor	Slowed blink rate, increased PERCLOS = increased sleepiness and reduced coping capacity. HRV data mixed findings	Various thresholds reported	Real time
Emotion, stress	- ECG (heart rate) - EDA	- ECG sensors (CardioWheel) - EDA wearable device	Increased heart rate and EDA = increased emotional response and	Unsure	Real time

		- Driver facing camera - Eye tracker (glasses / system)	reduced coping capacity		
Substance impairment	- Blood and Urine samples - Tissue readings -Breathalysers - EDA	- Wearable sensors (TruTouch)	Increased reading of impairment = reduced coping capacity	Unsure	Real time and one-off
Driving behaviour	- Speed - Braking - Lateral and longitudinal movement - Trajectory - Acceleration - Time to collision	- Forward facing camera and collision avoidance system (Mobileye) - Smart phones - Various driving sensors	Increased variables = reduced coping capacity	Various thresholds reported	Real time, post trip

8 References

Abdel-Aty, M., Ekram, A. A., Huang, H., & Choi, K. (2011). A study on crashes related to visibility obstruction due to fog and smoke. *Accident Analysis & Prevention*, *43*(5), 1730-1737.

Abtahi, F., Anund, A., Fors, C., Seoane, F., & Lindecrantz, K. (2017). Association of drivers' sleepiness with heart rate variability: A pilot study with drivers on real roads. In EMBEC & NBC 2017, 65, 149-152. Springer, Singapore.

Advantages and Disadvantages of GPS and Telematics Systems for Vehicles - National Motorists Association. (n.d.). Retrieved December 2, 2019, from https://www.motorists.org/blog/advantages-and-disadvantages-of-gps-and-telematics-systems-for-vehicles.

Aebi, C. (2002). Validierung der neuropsychologischen Testbatterie CERAD-NP: eine Multi-Center Studie (Doctoral dissertation, University_of_Basel).

Aeschbach, D., Matthews, J. R., Postolache, T. T., Jackson, M. A., Gieson, H. A., & Wehr, T. A. (1997). Dynamics of the human EEG during prolonged wakefulness: evidence for frequency- specific circadian and homeostatic influences. Neuroscience Letters, 239, 121–124.

Agrawal, U., Giripunje, S., & Bajaja, P. (2013). Emotion and gesture recognition with soft computing tool for driver assistance system in human centered transportation. In 2013 IEEE International Conference on Systems, Man, and Cybernetics (pp. 4612-4616).

Ahlström, C., Anund, A., Fors, C., & Åkerstedt, T. (2018). The effect of daylight versus darkness on driver sleepiness: a driving simulator study. Journal of Sleep Research, 27(3), e12642.

Ahlström, C., Anund, A., Fors, C., & Åkerstedt, T. (2018). Effects of the road environment on the development of driver sleepiness in young male drivers. Accident Analysis & Prevention, 112, 127-134.

Ahlström, C., Anund, A., & Kjellman, E. H. (2018) Stress, fatigue and inattention amongst city bus drivers–an explorative study on real roads within the ADAS&ME project.

Ahlström, C., Jansson, S., & Anund, A. (2017). Local changes in the wake electroencephalogram precedes lane departures. Journal of Sleep Research, 26(6), 816-819.

Ahn, S., Nguyen, T., Jang, H., Kim, J. G., & Jun, S. C. (2016). Exploring neuro-physiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data. Frontiers in Human Neuroscience, 10, 219.

Aidman, E., Johnson, K., Paech, G. M., Della Vedova, C., Pajcin, M., Grant, C., ... & Banks, S. (2018). Caffeine reduces the impact of drowsiness on driving errors. Transportation Research Part F: Traffic Psychology and Behaviour, 54, 236-247.

Aigner-Breuss, E., Kaiser, S., Usami, D. S., Reed, S., & Weijermars, W. (2017). Inventory of road user related risk factors and safety measures. Deliverable 4.4 of the H2020 project SafetyCube.

Åkerstedt, T., Anund, A., Axelsson, J., & Kecklund, G. (2014). Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. Journal of Sleep Research, 23(3), 242-254.

Åkerstedt, T., Connor, J., Gray, A., & Kecklund, G. (2008). Predicting road crashes from a mathematical model of alertness regulation—The Sleep/Wake Predictor. Accident Analysis & Prevention, 40(4), 1480-1485.

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. International Journal of Neuroscience, 18(10), 890–894.

Åkerstedt, T., Hallvig, D., Anund, A., Fors, C., Schwarz, J., & Kecklund, G. (2013). Having to stop driving at night because of dangerous sleepiness–awareness, physiology and behaviour. Journal of Sleep Research, 22(4), 380-388.

Åkerstedt, T., Ingre, M., Kecklund, G., Anund, A., Sandberg, D., Wahde, M., Philip, P., & Kronberg, P. (2010). Reaction of sleepiness indicators to partial sleep deprivation, time of day and time on task in a driving simulator—the DROWSI project. Journal of Sleep Research, 19, 298–309.

Al Machot, F. A., Mosa, A. H., Dabbour, K., Fasih, A., Schwarzlmuller, C., Ali, M., & Kyamakya, K. (2011). A novel real-time emotion detection system from audio streams based on Bayesian Quadratic Discriminate Classifier for ADAS. In Proceedings of the Joint INDS'11 & ISTET'11 (pp. 1-5). IEEE.

Alcohol Monitoring Systems, Inc. (2019). SCRAM Continuous Alcohol Monitoring | SCRAM Systems. Retrieved 26 November 2019, from https://www.scramsystems.com/products/scram-continuous-alcohol-monitoring [04.12.2019].

Ali, M., Al Machot, F., Mosa, A.H., & Kyamakya, K. (2016). CNN based subject-independent driver emotion recognition system involving physiological signals for ADAS. In Advanced Microsystems for Automotive Applications 2016 (pp. 125-138). Springer, Cham.

Ali, S. F., & Hassan, M. T. (2018). Feature Based Techniques for a Driver's Distraction Detection using Supervised Learning Algorithms based on Fixed Monocular Video Camera. TIIS, 12(8), 3820-3841.

Ali, S. S., Wilson, M. P., Castillo, E. M., Witucki, P., Simmons, T. T., & Vilke, G. M. (2013). Common hand sanitizer may distort readings of breathalyzer tests in the absence of acute intoxication. *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine*, *20*(2), 212–215. https://doi.org/10.1111/acem.12073

Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. Physiological measurement, 28(3), R1.

Alvaro, P. K., Jackson, M. L., Berlowitz, D. J., Swann, P., & Howard, M. E. (2016). Prolonged eyelid closure episodes during sleep deprivation in professional drivers. Journal of Clinical Sleep Medicine, 12(08), 1099-1103.

Amirjamshidi, G., & Roorda, M. J. (2015). Development of simulated driving cycles for light, medium, and heavy duty trucks: Case of the Toronto Waterfront Area. *Transportation Research Part D: Transport and Environment*, *34*, 255–266.

Anderson, C., Chang, A. M., Sullivan, J.P., Ronda, J. M., & Czeisler, C.A. (2013). Assessment of drowsiness based on ocular parameters detected by infrared reflectance oculography. Journal of Clinical Sleep Medicine, 9(9), 907–920.

Anderson, C., & Horne, J. A. (2013). Driving drowsy also worsens driver distraction. Sleep Medicine, 14(5), 466-468.

André, M. (2004). The ARTEMIS European driving cycles for measuring car pollutant emissions. *Science of the Total Environment*, *334–335*, 73–84.

Angarita, G. A., Natarajan, A., Gaiser, E. C., Parate, A., Marlin, B., Gueorguieva, R., … Malison, R. T. (2015). A remote wireless sensor network/electrocardiographic approach to discriminating cocaine use. Drug and Alcohol Dependence, C(146), e209. ttps://doi.org/10.1016/j.drugalcdep.2014.09.037

Anund, A., Ahlström, C., Fors, C., & Åkerstedt, T. (2018). Are professional drivers less sleepy than non-professional drivers?. Scandinavian Journal of Work, Environment and Health, 44(1), 88- 95.

Anund, A., Fors, C., & Ahlström, C. (2017). The severity of driver fatigue in terms of line crossing: a pilot study comparing day-and night time driving in simulator. European Transport Research Review, 9(2), 31.

Anund, A., Fors, C., Ihlström, J., & Kecklund, G. (2018). An on-road study of sleepiness in split shifts among city bus drivers. Accident Analysis & Prevention, 114, 71-76.

Anund, A., Kecklund, G., Vadeby, A., Hjälmdahl, M., & Åkerstedt, T. (2008). The alerting effect of hitting a rumble strip—A simulator study with sleepy drivers. Accident Analysis & Prevention, 40(6), 1970-1976.

Ariansyah, D., Caruso, G., Ruscio, D., & Bordegoni, M. (2018). Analysis of Autonomic Indexes on Drivers' Workload to Assess the Effect of Visual ADAS on User Experience and Driving Performance in Different Driving Conditions. Journal of Computing and Information Science in Engineering, 18(3), 031007.

Arnedt, J. T., Wilde, G. J., Munt, P. W., & MacLean, A. W. (2001). How do prolonged wakefulness and alcohol compare in the decrements they produce on a simulated driving task? Accident; Analysis and Prevention, 33(3), 337–344. https://doi.org/10.1016/s0001-4575(00)00047-6

Asbridge, M., & Ogilvie, R. (2015). *Feasibility Study of Roadside Oral Fluid Drug Testing*. Retrieved from http://www.madd.ca/media/docs/feasibility-roadside-oral-fluid-drug-testing.pdf [04.12.2019].

Auflick, J. L. (2015). Resurrecting driver workload metrics: a multivariate approach. Procedia Manufacturing, 3, 3160-3167.

Ayres, T. J., Li, L., Schleuning, D., & Young, D. (2001). Preferred time-headway of highway drivers. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 826–829.

BACtrack Inc. (2019). BACtrack Skyn™—The World's 1st Wearable Alcohol Monitor. Retrieved 26 November 2019, from BACtrack Skyn website: https://skyn.bactrack.com [04.12.2019].

Bailenson, J. N., Pontikakis, E. D., Mauss, I. B., Gross, J. J., Jabon, M. E., Hutcherson, C. A., ... & John, O. (2008). Real-time classification of evoked emotions using facial feature tracking and physiological responses. International journal of human-computer studies, 66(5), 303-317.

Bakhit, P. R., Osman, O. A., Guo, B., & Ishak, S. (2019). A distraction index for quantification of driver eye glance behavior: A study using SHRP2 NEST database. Safety Science, 119, 106-111.

Balasubramanian, V., & Bhardwaj, R. (2018). Can cECG be an unobtrusive surrogate to determine cognitive state of driver?. Transportation Research Part F: Traffic Psychology and Behaviour, 58, 797-806.

Balkin, T. J., Horrey, W. J., Graeber, R. C., Czeisler, C. A., & Dinges, D. F. (2011). The challenges and opportunities of technological approaches to fatigue management. Accident Analysis & Prevention, 43(2), 565-572.

Balters, S., Bernstein, M., & Paredes, P. E. (2019). On-road stress analysis for in-car interventions during the commute. In Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (p. LBW2110). ACM.

Balters, S., & Steinert, M. (2015). Capturing emotion reactivity through physiology measurement as a foundation for affective engineering in engineering design science and engineering practices. Journal of Intelligent Manufacturing, 28(7), 1585-1607.

Bandodkar, A. J., Jeerapan, I., & Wang, J. (2016). Wearable Chemical Sensors: Present Challenges and Future Prospects. *ACS Sensors*, *1*(5), 464–482.

Barnard, M. P., & Chapman, P. (2016). Are anxiety and fear separable emotions in driving? A laboratory study of behavioural and physiological responses to different driving environments. Accident Analysis and Prevention, 86, 99-107.

Barnett, N. P. (2015). Alcohol sensors and their potential for improving clinical care. *Addiction (Abingdon, England)*, *110*(1), 1–3.

Barnett, N. P., Celio, M. A., Tidey, J. W., Murphy, J. G., Colby, S. M., & Swift, R. M. (2017). A preliminary randomized controlled trial of contingency management for alcohol use reduction using a transdermal alcohol sensor. *Addiction (Abingdon, England)*, *112*(6), 1025–1035.

Barnett, N. P., Meade, E. B., & Glynn, T. R. (2014). Predictors of detection of alcohol use episodes using a transdermal alcohol sensor. *Experimental and Clinical Psychopharmacology*, *22*(1), 86–96.

Bartley, S. H., & Chute, E. (1947). Fatigue and impairment in man. New York, NY: McGraw-Hill Book Company

Barua, S., Ahmed, M. U., Ahlström, C., & Begum, S. (2019). Automatic driver sleepiness detection using EEG, EOG and contextual information. Expert Systems with Applications, 115, 121-135.

Bella, F., & Silvestri, M. (2015). Effects of safety measures on driver's speed behavior at pedestrian crossings. Accident Analysis & Prevention, 83, 111-124.

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., & Montanari, R. (2011). Driver workload and eye blink duration. Transportation research part F: traffic psychology and behaviour, 14(3), 199-208.

Bergel, R., Rattaire, H., Aron, M., Doucet, D., & Violette, E. (2010, July). Adeed risk in case of rain: some recent results for France.

Berry, R. B., & Wagner, M. H. (2014). Sleep medicine pearls E-Book. Elsevier Health Sciences.

Bioulac, S., Franchi, J. A. M., Arnaud, M., Sagaspe, P., Moore, N., Salvo, F., & Philip, P. (2017). Risk of motor vehicle accidents related to sleepiness at the wheel: a systematic review and metaanalysis. Sleep, 40(10).

Blower, D., Campbell, K. L., & Green, P. E. (1993). Accident rates for heavy truck-tractors in Michigan. Accident Analysis & Prevention, 25(3), 307-321.

Bongiorno, N., Bosurgi, G., Pellegrino, O., & Sollazzo, G. (2017). How is the driver's workload influenced by the road environment?. Procedia engineering, 187, 5-13.

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neuroscience & Biobehavioral Reviews, 44, 58-75.

Bosch, E., Drewitz, U., Ihme, K., & Jipp, M. (2019, February 6-7). The role of emotion recognition in future mobility visions. AAET Conference, Braunschweig, Germany.

Botta, M., Cancelliere, R., Ghignone, L., Tango, F., Gallinari, P., & Luison, C. (2019). Real-time detection of driver distraction: random projections for pseudo-inversion-based neural training. Knowledge and Information Systems, 60(3), 1549-1564.

Brijs, T., Karlis, D., & Wets, G. (2008). Studying the effect of weather conditions on daily crash counts using a discrete time-series model. Accident Analysis & Prevention, 40(3), 1180-1190.

Brookhuis, K. A., & De Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. Accident Analysis & Prevention, 42(3), 898-903.

Buendia, R., Forcolin, F., Karlsson, J., Arne Sjöqvist, B., Anund, A., & Candefjord, S. (2019). Deriving heart rate variability indices from cardiac monitoring—An indicator of driver sleepiness. Traffic Injury Prevention, 20(3), 249-254.

Bunn, T. L., Slavova, S., Struttmann, T. W., & Browning, S. R. (2005). Sleepiness/fatigue and distraction/inattention as factors for fatal versus nonfatal commercial motor vehicle driver injuries. Accident Analysis & Prevention, 37(5), 862-869.

Burgess, H. J., Trinder, J., Kim, Y., & Luke, D. (1997). Sleep and circadian influences on cardiac autonomic nervous system activity. American Journal of Physiology-Heart and Circulatory Physiology, 273(4), H1761-H1768.

Butler, G. (1993). Definitions of stress. Occasional paper series RCGP, 61, 1-5.

Cai, S., Becherif, M., Bakhouya, M., & Wack, M. (2010). *A Context-aware Embedded System for Intelligent Vehicles and Smart Roads*. Berlin: UPC workshop in conjunction with ICPS 2010.

Campbell, A. S., Kim, J., & Wang, J. (2018). Wearable electrochemical alcohol biosensors. *Current Opinion in Electrochemistry*, *10*, 126–135.

Caponecchia, C., & Williamson, A. (2018). Drowsiness and driving performance on commuter trips. Journal of Safety Research, 66, 179-186.

Carreiro, S., Fang, H., Zhang, J., Wittbold, K., Weng, S., Mullins, R., … Boyer, E. W. (2015). iMStrong: Deployment of a Biosensor System to Detect Cocaine Use. *Journal of Medical Systems*, *39*(12), 186.

Carreiro, S., Smelson, D., Ranney, M., Horvath, K. J., Picard, R. W., Boudreaux, E. D., … Boyer, E. W. (2015). Real-Time Mobile Detection of Drug Use with Wearable Biosensors: A Pilot Study. *Journal of Medical Toxicology*, *11*(1), 73–79.

Carreiro, S., Wittbold, K., Indic, P., Fang, H., Zhang, J., & Boyer, E. W. (2016). Wearable Biosensors to Detect Physiologic Change During Opioid Use. *Journal of Medical Toxicology*, *12*(3), 255–262.

Carskadon, M. A. (1986). Guidelines for the multiple sleep latency test (MSLT): a standard measure of sleepiness. Sleep, 9(4), 519-524.

Chaiard, J., Deeluea, J., Suksatit, B., & Songkham, W. (2019). Factors associated with sleep quality of Thai intercity bus drivers. Industrial Health, 2018-0168.

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., & Wallace, E. P. (1993). Development of a fatigue scale. Journal of Psychometric Research, 37(2), 147–153.

Chaloupka, C. & Risser, R. (1995). Don't wait for accidents - possibilities to assess risk in traffic by applying the 'Wiener Fahrprobe'. Safety Science, 19(2–3), 137-147.

Chan, M., & Singhal, A. (2015). Emotion matters: implications for distracted driving. Safety Science 72, 302-309.

Chen, H., Liu, P., Lu, J. J., & Behzadi, B. (2009). Evaluating the safety impacts of the number and arrangement of lanes on freeway exit ramps. Accident Analysis & Prevention, 41(3), 543-551.

Chengye, P., & Ranjitkar, P. (2013). Modelling motorway accidents using negative binomial regression. Journal of the Eastern Asia Society for Transportation Studies, 10, 1946-1963.

Choi, M., Koo, G., Seo, M., & Kim, S. W. (2017). Wearable device-based system to monitor a driver's stress, fatigue, and drowsiness. IEEE Transactions on Instrumentation and Measurement, 67(3), 634-645.

Clifford, G. D., & Tarassenko, L. (2005). Quantifying errors in spectral estimates of HRV due to beat replacement and resampling. IEEE Transactions on Biomedical Engineering, 52(4), 630-638.

Connor, J., Norton, R., Ameratunga, S., Robinson, E., Civil, I., Dunn, R., ... & Jackson, R. (2002). Driver sleepiness and risk of serious injury to car occupants: population-based case control study. British Medical Journal, 324(7346), 1125.

Cori, J. M., Anderson, C., Soleimanloo, S. S., Jackson, M. L., & Howard, M. E. (2019). Narrative review: Do spontaneous eye blink parameters provide a useful assessment of state drowsiness?. Sleep Medicine Reviews, 45, 95-104.

Costa, J. O., Maria, A. P., Pereira, P. A., Freitas, E. F., & Soares, F. E. (2018). Portuguese two-lane highways: modelling crash frequencies for different temporal and spatial aggregation of crash data. Transport, 33(1), 92-103.

Costa, M., Oliveira, D., Pinto, S., & Tavares, A. (2019). Detecting Driver's Fatigue, Distraction and Activity Using a Non-Intrusive Ai-Based Monitoring System. Journal of Artificial Intelligence and Soft Computing Research, 9(4), 247-266.

Cunningham, M. L., & Regan, M. A. (2017). Are Happy Drivers Better Drivers? The Impact of Emotion, Life Stress and Mental Health Issues on Driving Performance and Safety. In Australasian Road Safety Conference, 2017, Perth, Western Australia, Australia.

Cunningham, M. L., & Regan, M. (2018). Driver distraction and inattention. Safe mobility: Challenges, methodology and solutions, 57-82.

Cyders, M. A., & Smith, G. T. (2008). Emotion-based dispositions to rash action: positive and negative urgency. Psychological bulletin, 134(6), 807.

D'Orey, P. M., & Ferreira, M. (2014). ITS for sustainable mobility: A survey on applications and impact assessment tools. *IEEE Transactions on Intelligent Transportation Systems*, *15*(2), 477–493.

Dagan, E., Mano, O., Stein, G. P., & Shashua, A. (2004). Forward collision warning with a single camera. *IEEE Intelligent Vehicles Symposium, Proceedings*, 37–42.

Dahlen, E. R., & Ragan, K. M. (2004). Validation of the propensity for angry driving scale. Journal of safety research, 35(5), 557-563.

Damasio, A. R. (2001). Fundamental feelings. Nature, 413(6858), 781.

Darzi, A., Gaweesh, S., Ahmed, M., & Novak, D. (2018). Identifying the causes of drivers' hazardous states using driver characteristics, vehicle kinematics and physiological measurements. Frontiers in Neuroscience, 12, 568.

Dawson, D., & Reid, K. (1997). Fatigue, alcohol and performance impairment. Nature, 388(6639), 235.

Deffenbacher, J. L., Oetting, E. R., & Lynch, R. S. (1994). Development of a driving anger scale. Psychological Reports, 74, 83-91.

Deffenbacher, J. L., Stephens, A. N., & Sullmanc, M. J. M. (2016). Driving anger as a psychological construct: Twenty years of research using the Driving Anger Scale, Transportation Research Part F: Traffic Psychology and Behaviour, 42(2), 236-247.

Dehais F., Duprès A, Blum S., Drougard N., Scannella S., Roy N.R., & Lotte F. (2019). Monitoring Pilot's Mental Workload Using ERPs and Spectral Power with a Six-Dry-Electrode EEG System in Real Flight Conditions. Sensors 2019, 19, 1324; doi:10.3390/s19061324.

Dement, W. C., & Carskadon, M. A. (1982). Current perspectives on daytime sleepiness: the issues. Sleep: Journal of Sleep Research & Sleep Medicine, 5(2), 56-66.

De Winter, J. C. F., & Dodou, D. (2010). The Driver Behaviour Questionnaire as a predictor of accidents: A meta-analysis. *Journal of safety research*, *41*(6), 463-470.

Diaz-Piedra, C., Gomez-Milan, E., & Di Stasi, L. L. (2019). Nasal skin temperature reveals changes in arousal levels due to time on task: An experimental thermal infrared imaging study. Applied Ergonomics, 81, 102870.

Di Milia, L., Smolensky, M. H., Costa, G., Howarth, H. D., Ohayon, M. M., & Philip, P. (2011). Demographic factors, fatigue, and driving accidents: An examination of the published literature. Accident Analysis & Prevention, 43(2), 516-532

Dinges, D. F., & Grace, R. (1998). PERCLOS: A Valid Psychophysiological Measure of Alertness As Assessed by Psychomotor Vigilance (Report No. FHWA-MCRT-98-006). National Highway Transportation Safety Administration, Washington, D.C.

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. Proceedings of the National Academy of Sciences, 113(10), 2636–2641.

Dobbins, C., & Fairclough, S. (2018). Detecting negative emotions during real-life driving via dynamically labelled physiological data. In: 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). IEEE, 2018. S. 830-835.

Doran, S. M., Van Dongen, H. P. A., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: evidence of state instability. Archives Italiennes de Biologie, 139(3), 253–267.

Dougherty, D. M., Karns, T. E., Mullen, J., Liang, Y., Lake, S. L., Roache, J. D., & Hill-Kapturczak, N. (2015). Transdermal alcohol concentration data collected during a contingency management program to reduce at-risk drinking. *Drug and Alcohol Dependence*, *148*, 77–84.

Dumitru, A. I., Girbacia, T., Boboc, R. G., Postelnicu, C. C., & Mogan, G. L. (2018). Effects of smartphone based advanced driver assistance system on distracted driving behavior: A simulator study. Computers in Human Behavior, 83, 1-7.

Dunn, N., & Williamson, A. (2011). Monotony in the rail industry: The role of task demand in mitigating monotony-related effects on performance. In Ergonomics Australia–HFESA 2011 Conference Edition (Vol. 11, pp. 1-5).

Eboli, L., Mazzulla, G., & Pungillo, G. (2016). Combining speed and acceleration to define car users' safe or unsafe driving behaviour. Transportation research part C: emerging technologies, 68, 113-125.

Eichhorn, A., & Pilgerstorfer, M. (2017). Emotions – Anger, Aggression, European Road Safety Decision Support System, developed by the H2020 project SafetyCube. Retrieved from www.roadsafety-dss.eu on 27/10/2019.

Ekman, P. (1992). An argument for basic emotions. Cognition & emotion, 6(3-4), 169-200.

ElBatt, T., Goel, S. K., Holland, G., Krishnan, H., & Parikh, J. (2006). Cooperative collision warning using dedicated short range wireless communications. *VANET - Proceedings of the Third ACM International Workshop on Vehicular Ad Hoc Networks*, *2006*, 1–9.

Ellis, H. C., & Moore, B. A. (1999). Mood and memory. In T. Dalgleish & M. J. Power (Eds.), Handbook of Cognition and Emotion (pp. 193–210). Wiley: Chichester.

Ellison, A. B., Greaves, S. P. &, & Daniels, R. (2012). Profiling Drivers' Risky Behaviour Towards All Road Users. *Australasian College of Road Safety Conference*. Sydney

Elvik, R. (2013). Risk of road accident associated with the use of drugs: A systematic review and meta-analysis of evidence from epidemiological studies. *Accident Analysis and Prevention*, *60*, 254–267.

Elvik, R., Fridstrøm, L., Kaminska, J., & Meyer, S. F. (2013). Effects on accidents of changes in the use of studded tyres in major cities in Norway: A long-term investigation. Accident Analysis & Prevention, 54, 15-25.

Espinosa, F., Jiménez, J. A., Santiso, E., Gardel, A., Pérez, D., Casanova, J., & Santos, C. (2011). Design and implementation of a portable electronic system for vehicle-driver-route activity measurement. *Measurement: Journal of the International Measurement Confederation*, *44*(2), 326–337.

ETSC. (2017). *Preventing Drug Driving in Europe: Policy measures for national and EU action* (p. 24). Brussels, Belgium: ETSC/European Transport Safety Council.

European Commission (2019). Lack of driving experience. Retrieved from: https://ec.europa.eu/transport/road_safety/specialist/knowledge/young/contributing_factors_to_crash_risk/lack_of_driving_experience_en

European Monitoring Centre for Drugs and Drug Addiction. (2012). Driving Under the Influence of Drugs, Alcohol and Medicines in Europe—Findings from the DRUID project (p. 57). Luxembourg: Publications Office of the European Union.

Fairbairn, C. E., & Kang, D. (2019). Temporal Dynamics of Transdermal Alcohol Concentration Measured via New-Generation Wrist-Worn Biosensor. Alcoholism: Clinical and Experimental Research, 43(10), 2060–2069.

Fairclough, S. H., Venables, L., & Tattersall, A. (2005). The influence of task demand and learning on the psychophysiological response. International Journal of Psychophysiology, 56(2), 171-184.

Fan, X., Bi, l., & Chen, Z. (2010). Using EEG to detct drivers' emotion with Bayesian networks. In 2010 International Conference on Machine Learning and Cybernetics (Vol. 3, pp. 1177-1181). IEEE.

Feng, F., Bao, S., Sayer, J. R., Flannagan, C., Manser, M., & Wunderlich, R. (2017). Can vehicle longitudinal jerk be used to identify aggressive drivers? An examination using naturalistic driving data. *Accident Analysis and Prevention*, *104*, 125–136.

Filtness, A. J., Anund, A., Fors, C., Ahlström, C., Åkerstedt, T., & Kecklund, G. (2014). Sleep-related eye symptoms and their potential for identifying driver sleepiness. Journal of Sleep Research, 23(5), 568-575.

Fleming, W. J. (2008). New automotive sensors - A review. *IEEE Sensors Journal*, *8*(11), 1900–1921.

Folstein, M. F., Folstein, S. E., &. McHugh, P. R. (1975). “MINI-MENTAL-STATE” A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatry Research, 12, 189-198.

Forcolin, F., Buendia, R., Candefjord, S., Karlsson, J., Sjöqvist, B. A., & Anund, A. (2018). Comparison of outlier heartbeat identification and spectral transformation strategies for deriving heart rate variability indices for drivers at different stages of sleepiness. Traffic Injury Prevention, 19(sup1), S112-S119.

Fors, C., Ahlström, C., & Anund, A. (2018). A comparison of driver sleepiness in the simulator and on the real road. Journal of Transportation Safety & Security, 10(1-2), 72-87.

Forsman, P. M., Vila, B. J., Short, R. A., Mott, C. G., & Van Dongen, H. P. A. (2013). Efficient driver drowsiness detection at moderate levels of drowsiness. Accident Analysis & Prevention, 50, 341-350.

Foy, H. J., & Chapman, P. (2018). Mental workload is reflected in driver behaviour, physiology, eye movements and prefrontal cortex activation. Applied ergonomics, 73, 90-99.

Fridstrøm, L., & Ingebrigtsen, S. (1991). An aggregate accident model based on pooled, regional time-series data∗. Accident Analysis & Prevention, 23(5), 363-378.

Ftouni, S., Rahman, S. A., Crowley, K. E., Anderson, C., Rajaratnam, S. M. W., & Lockley, S. W. (2013). Temporal dynamics of ocular indicators of sleepiness across sleep restriction. Journal of Biological Rhythms, 28(6), 412–424.

Fuller, R. (2011). Driver control theory: From task difficulty homeostasis to risk allostasis. In Handbook of traffic psychology (pp. 13-26). Academic Press.

Gaca, S., & Kiec, M. (2013). Risk of accidents during darkness on roads with different technical standards. In 16th International Conference Road Safety on Four Continents. Beijing, China (RS4C 2013). 15-17 May 2013. Statens väg-och transportforskningsinstitut.

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

Garbarino, S., Lino, N., Beelke, M., Carli, F. D., & Ferrillo, F. (2001). The contributing role of sleepiness in highway vehicle accidents. Sleep, 24(2), 201-206.

Garnefski, N., Kraaij, V., & Spinhoven, Ph. (2001). Negative life events, cognitive emotion regulation and emotional problems. Personality and Individual Differences, 30, 1311-1327.

Gateau, T., Ayaz, H., & Dehais, F. (2018) In silico vs. Over the Clouds: On-the-Fly Mental State Estimation of Aircraft Pilots, Using a Functional Near Infrared Spectroscopy Based Passive-BCI. Front. Hum. Neurosci. 12:187. doi: 10.3389/fnhum.2018.00187

Gelau, C., Gasser, T. M., Seeck, A. (2012). Fahrerassistenz und Verkehrssicherheit. In H. Winner, S. Hakuli & G. Wolf, Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort. Wiesbaden: Vieweg + Teubner.

Gianfranchi, E., Tagliabue, M., Spoto, A., & Vidotto, G. (2017). Sensation seeking, non-contextual decision making, and driving abilities as measured through a moped simulator. Frontiers in psychology, 8, 2126.

Girardi, D., Lanubile, F., & Novielli, N. (2017). Emotion detection using noninvasive low cost sensors. In 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII) (pp. 125-130). IEEE.

Goldbeck, J., & Huertgen, B. (1999). Lane detection and tracking by video sensors. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 74–79.

Goldenbeld, C., van Schagen, I. (2016), Personal Factors - Sensation Seeking, European Road Safety Decision Support System, developed by the H2020 project SafetyCube. Retrieved 19.12.2019 from www.roadsafety-dss.eu

Golob, T. F., Recker, W., & Pavlis, Y. (2008). Probabilistic models of freeway safety performance using traffic flow data as predictors. Safety science, 46(9), 1306-1333.

Golz, M., Sommer, D., Trutschel, U., Sirois, B., & Edwards, D. (2010). Evaluation of fatigue monitoring technologies. Somnologie-Schlafforschung und Schlafmedizin, 14(3), 187-199.

Gonzalez-Sanchez, J., Baydogan, M., Chavez-Echeagaray, M. E., Atkinson, R. K., & Burleson, W. (2017). Affect measurement: A roadmap through approaches, technologies, and data analysis. In M. Jeon (Ed.), Emotions and Affect in Human Factors and Human-Computer Interaction (pp. 255-288). San Diego: Academic Press.

Gotardi, G., Schor, P., Van Der Kamp, J., Navarro, M., Orth, D., Savelsbergh, G., Polastri, P. F. et al. (2018). The influence of anxiety on visual entropy of experienced drivers. In Proceedings of the 3rd Workshop on Eye Tracking and Visualization (p. 4). ACM.

Graham, R., & Hirst, S. (1994). The effect of a collision avoidance system on drivers' braking responses. Moving towards deployment. Proceedings of the 1994 annual meeting of IVHS America.

Greaves, S. P., & Ellison, A. B. (2011). Personality, risk aversion and speeding: An empirical investigation. *Accident Analysis and Prevention*, *43*(5), 1828–1836.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. Psychophysiology, 39, 281–291.

Guo, F., Wang, X., & Abdel-Aty, M. A. (2010). Modeling signalized intersection safety with corridor-level spatial correlations. Accident Analysis & Prevention, 42(1), 84-92.

Guo, Y., Wang, X., Xu, Q., Liu, F., Liu, Y., & Xia, Y. (2019). Change-point analysis of eye movement characteristics for female drivers in anxiety. International journal of environmental research and public health, 16(7), 1236.

Hac, A., & Simpson, M. D. (2000). Estimation of vehicle side slip angle and yaw rate. SAE Technical Papers.

Haleem, K., & Abdel-Aty, M. (2010). Examining traffic crash injury severity at unsignalized intersections. Journal of safety research, 41(4), 347-357.

Halim, Z., & Rehan, M. (2020). On identification of driving-induced stress using electroencephalogram signals: A framework based on wearable safety-critical scheme and machine learning. Information Fusion, 53, 66-79.

Hallvig, D., Anund, A., Fors, C., Kecklund, G., Karlsson, J. G., Wahde, M., & Åkerstedt, T. (2013). Sleepy driving on the real road and in the simulator - A comparison. Accident Analysis & Prevention, 50, 44-50.

Hallvig, D., Anund, A., Fors, C., Kecklund, G., & Åkerstedt, T. (2014). Real driving at night – predicting lane departures from physiological and subjective. Biological Psychology, 101, 18-23.

Hammoud, R. I., Smith, M. R., Dufour, R., Bakowski, D., & Witt, G. (2008). Driver distraction monitoring and adaptive safety warning systems (No. 2008-01-2694). SAE Technical Paper.

Haouij, E. (2018). Biosignals for driver's stress level assessment: functional variable selection and fractal characterization. (Doctoral dissertation, Paris Saclay).

Hari, C. V., & Sankaran, P. (2017, August). Embedding vehicle driver face poses on manifolds. In 2017 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES) (pp. 1-5). IEEE.

Hatakka, M., Keskinen, E., Gregersen, N.P., Glad, A., &Hernetkoski, K. (2002). From control of the vehicle to personal self-control: broadening the perspectives of driver education. Transportation Research Part F: Psychology and Behavior, 201-215.

Häusler, J., & Sturm, W. (2009). Construct validation of a new test battery for perceptual and attention functions. Zeitschrift für Neuropsychologie, 20, 327-339.

He, J., Choi, W., Yang, Y., Lu, J., Wu, X., & Peng, K. (2017). Detection of driver drowsiness using wearable devices: A feasibility study of the proximity sensor. Applied Ergonomics, 65, 473-480.

Heikenfeld, J., Jajack, A., Rogers, J., Gutruf, P., Tian, L., Pan, T., … Kim, J. (2018). Wearable sensors: Modalities, challenges, and prospects. *Lab on a Chip*, *18*(2), 217–248

Hels, T., Bernhoft, I., Lyckegaard, A., Houwing, S., Hagenzieker, M., Legrand, S.-A., … Verstraete, A. (2011). *Risk of injury by diving with alcohol and other drugs.* DRUID (Driving under the Influence of Drugs, Alcohol and Medicines). 6th Framework programme. Deliverable 2.3.5.

Ho, S. H., Wong, Y. D., & Chang, V. W. C. (2014). Developing Singapore Driving Cycle for passenger cars to estimate fuel consumption and vehicular emissions. *Atmospheric Environment*, *97*, 353–362.

Hogema, J. H., & Janssen, W. H. (1996). Effects of intelligent cruise control on driving behaviour: a simulator study. Intelligent Transportation: Realizing the Future. Abstracts of the Third World Congress on Intelligent Transport Systems. TNO.

Holmes, A. L., Burgess, H. J., & Dawson, D. (2002). Effects of sleep pressure on endogenous cardiac autonomic activity and body temperature. Journal of Applied Physiology, 92(6), 2578-2584.

Horne, J. A., & Reyner, L. A. (1995). Sleep related vehicle accidents. British Medical Journal, 310(6979), 565-567.

Houwing, S., Hagenzieker, M., Mathijssen, R., Bernhoft, I., Hels, T., Janstrup, K., … Verstraete, A. (2011). *Prevalence of alcohol and other psychoactive substances in drivers in general traffic Part I: General results.* DRUID (Driving under the Influence of Drugs, Alcohol and Medicines). 6th Framework programme. Deliverable 2.2.3 Part I.

Howell, J., Nag, A., McKnight, M., Narsipur, S., & Adelegan, O. (2015). A low-power wearable substance monitoring device. *2015 IEEE Virtual Conference on Applications of Commercial Sensors (VCACS)*, 1–9.

Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Pugzles Lorch, E., & Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. Personality and Individual Differences, 32(3), 401-414.

Huang, H., Chen, H., & Lin, S. (2019, June). MagTrack: Enabling Safe Driving Monitoring with Wearable Magnetics. In Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services (pp. 326-339).

Ihl, R., Grass-Kapanke, B., Lahrem, P., Brinkmeyer, J., Fischer, S., Gaab, N., & Kaupmannsennecke, C. (2000). Entwicklung und Validierung eines Tests zur Früherkennung der Demenz mit Depressionsabgrenzung (TFDD). Fortschritte der Neurologie Psychiatrie, 68, 413-422.

Ingre, M., Åkerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2006). Subjective sleepiness, simulated driving performance and blink duration; examining individual differences. Journal of Sleep Research, 15, 47-53.

Ingre, M., Kecklund, G., Åkerstedt, T., & Kecklund, L. (2004). Variation in sleepiness during early morning shifts: a mixed model approach to an experimental field study of train drivers. Chronobiology International, 21(6), 973-990.

Isalberti, C., Van der Linden, T., Legrand, S.-A., Verstraete, A., Bernhoft, I., Hels, T., … Mathijssen, R. (2011). *Prevalence of alcohol and other psychoactive substances in injured and killed drivers.* DRUID (Driving under the Influence of Drugs, Alcohol and Medicines). 6th Framework programme. Deliverable 2.2.5.

Iversen, H. (2004). Risk-taking attitudes and risky driving behaviour. Transportation Research Part F: Traffic Psychology and Behaviour, 7(3), 135-150.

Jackson, M. L., Kennedy, G. A., Clarke, C., Gullo, M., Swann, P., Downey, L. A., Hayley, A. C., Pierce, R. J., & Howard, M. E. (2016). The utility of automated measures of ocular metrics for detecting driver drowsiness during extended wakefulness. Accident Analysis & Prevention, 87, 127–133.

James, L., & Nahl, D. (2002). Taxonomy of driving behavior: Affective, cognitive and sensorimotor. Driving Lessons: Exploring systems that make traffic safer, 21-49.

Jeon, M. (2017). Emotions and affect in human factors and human-computer interaction: taxonomy, theories, approaches, and methods. In M. Jeon (Ed.), Emotions and affect in human factors and human-computer interaction (pp. 3-26). San Diego: Academic Press.

Jeon, M., Walker, B. N., & Gable, T. M. (2015). The effects of social interactions with in-vehicle agents on a driver's anger level, driving performance, situation awareness, and perceived workload. Applied ergonomics, 50, 185-199.

Jia, Y., & Tyler, C. W. (2019). Measurement of saccadic eye movements by electrooculography for simultaneous EEG recording. Behavioral Research Methods, 51(5), 2139-2151.

Johansson, Ö., Wanvik, P. O., & Elvik, R. (2009). A new method for assessing the risk of accident associated with darkness. Accident Analysis & Prevention, 41(4), 809-815.

Johns, M. W. (1991). “A new method for measuring daytime sleepiness: the Epworth sleepiness scale.” Sleep, 14(6), 540–545.

Jonah, B. A. (1997). Sensation seeking and risky driving: a review and synthesis of the literature. Accident Analysis & Prevention, 5, 651-65.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. Aviation Space and Environmental Medicine, 67(6), 507–512.

Kahneman, D. (2011). Thinking, fast and slow. Macmillan.

Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. Clinical Neurophysiology, 117(7), 1574-1581.

Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A. P., Brand, M. & Bullock, R. (2004). DemTect: a new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. International Journal of Geriatric Psychiatry, 19, 136–143.

Kanaan, D., Ayas, S., Donmez, B., Risteska, M., & Chakraborty, J. (2019). Using Naturalistic Vehicle-Based Data to Predict Distraction and Environmental Demand. International Journal of Mobile Human Computer Interaction (IJMHCI), 11(3), 59-70.

Katsis, C. D., Katertsidis, N., Ganiatsas, G., & Fotiadis, D. I. (2008). Toward emotion recognition in car-racing drivers: a biosignal processing approach. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 38(3), 502-512.

Kaye, S. A., Lewis, I., & Freeman, J. (2018). Comparison of self-report and objective measures of driving behavior and road safety: A systematic review. *Journal of Safety Research*, *65*, 141–151.

Kecklund, G., & Åkerstedt, T. (1993). Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. Ergonomics, 36(9), 1007-1017.

Khan, M. Q., & Lee, S. (2019). A comprehensive survey of driving monitoring and assistance systems. Sensors, 19(11), 2574.

Kolli, A., Fasih, A., Al Machot, F., & Kyamakya, K. (2011). Non-intrusive car driver's emotion recognition using thermal camera. In Proceedings of the Joint INDS'11 & ISTET'11 (pp. 1-5). IEEE.

Koohestani et al. (2019) 'Drivers awareness evaluation using physiological measurement in a driving simulator', Proceedings of the IEEE International Conference on Industrial Technology. IEEE, 2019-Febru, pp. 859–864.

Kreutzer, J. S., DeLuca, J., & Caplan, B. (Eds.) (2018). Encyclopedia of Clinical Neuropsychology. SpringerLink (Online service).

Kubinger, K. D., & Ortner, T. M. (2010). Psychologische Diagnostik in Fallbeispielen. Hogrefe Verlag.

Lafont, A., Rogé, J., Ndiaye, D., & Boucheix, J.-M. (2018). Driver's emotional state and detection of vulnerable road users: Towards a better understanding of how emotions affect drivers' perception using cardiac and ocular metrics. Transportation Research Part F 55, 141–152.

Lagarde, E., Chastang, J. F., Gueguen, A., Coeuret-Pellicer, M., Chiron, M., & Lafont, S. (2004). Emotional stress and traffic accidents: the impact of separation and divorce. Epidemiology, 15(6), 762-766.

Lamble, D., Laakso, M., & Summala, H. (1999). Detection thresholds in car following situations and peripheral vision: Implications for positioning of visually demanding in-car displays. *Ergonomics*, *42*(6), 807–815.

Landowska, A. (2014). Emotion monitoring - verification of physiological characteristics measurement procedures. Metrology and measurement systems, 21(4), 719-732.

Laureshyn, A., Goede, M. de, Saunier, N., & Fyhri, A. (2017). Cross-comparison of three surrogate safety methods to diagnose cyclist safety problems at intersections in Norway. *Accident Analysis and Prevention*, *105*, 11–20.

Lee, J., & Jang, K. (2017). A framework for evaluating aggressive driving behaviors based on in-vehicle driving records. *Transportation Research Part F: Traffic Psychology and Behaviour*.

Lee, K. W., Yoon, H. S., Song, J. M., & Park, K. R. (2018). Convolutional neural network-based classification of driver's emotion during aggressive and smooth driving using multi-modal camera sensors. Sensors, 18(4), 957.

Lees, T., Chalmers, T., Burton, D., Zilberg, E., Penzel, T., Lal, S., & Lal, S. (2018). Electroencephalography as a predictor of self-report fatigue/sleepiness during monotonous driving in train drivers. Physiological Measurement, 39(10), 105012.

Leffingwell, T. R., Cooney, N. J., Murphy, J. G., Luczak, S., Rosen, G., Dougherty, D. M., & Barnett, N. P. (2013). Continuous objective monitoring of alcohol use: Twenty-first century measurement using transdermal sensors. *Alcoholism, Clinical and Experimental Research*, *37*(1), 16–22.

Legrand, S.-A., Gjerde, H., Isalberti, C., Van der Linden, T., Lillsunde, P., Dias, M. J., … Verstraete, A. G. (2014). Prevalence of alcohol, illicit drugs and psychoactive medicines in killed drivers in four European countries. *International Journal of Injury Control and Safety Promotion*, *21*(1), 17–28.

Legree, P. J., Heffner, T. S., Psotka, J., Martin, D. E., & Medsker, G. J. (2003). Traffic crash involvement: Experiential driving knowledge and stressful contextual antecedents. Journal of applied psychology, 88(1), 15.

Lehrer P., Karavidas M., Lu S-E, Vaschillo E., Vaschillo B., & Cheng A. (2010). Cardiac data increase association between self-report and both expert ratings of task load and task performance in flight simulator tasks: An exploratory study. International Journal of Psychophysiology 76 (2010) 80–87.

Li, K., Xiong, C., Wang, J., & Lu, M. (2015). Longitudinal driving behaviour on different roadway categories: an instrumented-vehicle experiment, data collection and case study in China. *IET Intelligent Transport Systems*, *9*(5), 555–563.

Li, X., & Seignez, E. (2018). Driver inattention monitoring system based on multimodal fusion with visual cues to improve driving safety. Transactions of the Institute of Measurement and Control, 40(3), 885-895.

Liang, Y., Horrey, W. J., Howard, M. E., Lee, M. L., Anderson, C., Shreeve, M. S., ... & Czeisler, C. A. (2019). Prediction of drowsiness events in night shift workers during morning driving. Accident Analysis & Prevention, 126, 105-114.

Lim, J., & Dinges, D. (2008). Sleep deprivation and vigilant attention. Annals of the New York Academy of Sciences, 1129(1), 305.

Lin, N., Zong, C., Tomizuka, M., Song, P., Zhang, Z., & Li, G. (2014). An overview on study of identification of driver behavior characteristics for automotive control. *Mathematical Problems in Engineering*, Vol. 2014.

Lindeman, M., & Verkasalo, M. (2005). Measuring values with the Short Schwartz's Value Survey. Journal of Personality Assessment, 85(2),170-178.

Linkov, V., Zaoral, A., Řezáč, P., &. Pai, C-W. (2019). Personality and professional drivers' driving behavior. Transportation Research Part F: Traffic Psychology and Behaviour, 60, 105-110.

Lippman, N. E. A. L., Stein, K. M., & Lerman, B. B. (1994). Comparison of methods for removal of ectopy in measurement of heart rate variability. American Journal of Physiology-Heart and Circulatory Physiology, 267(1), H411-H418.

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

Lohani, M., Payne, B. R., & Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. Frontiers in Human Neuroscience, 13.

Lotz, A., Ihme, K., Charnoz, A., Maroudis, P., Dmitriev, I., & Wendemuth, A. (2018). Recognizing behavioral factors while driving: a real-world multimodal corpus to monitor the driver's affective state. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018).

Lounis C., Peysakhovich V., & Causse M. (2020) Flight Eye Tracking Assistant (FETA): Proof of Concept. In: Stanton N. (eds) Advances in Human Factors of Transportation. AHFE 2019. Advances in Intelligent Systems and Computing, vol 964. Springer, Cham.

Luczak, S. E., & Rosen, I. G. (2014). Estimating BrAC from transdermal alcohol concentration data using the BrAC estimator software program. *Alcoholism, Clinical and Experimental Research*, *38*(8), 2243–2252.

Luig J., & Sontacchi a. (2014) A speech database for stress monitoring in the cockpit. Proc IMechE Part G: J Aerospace Engineering 2014, Vol 228(2) 284–296. DOI: 10.1177/0954410012467944

Ma, J., Gu, J., Jia, H., Yao, Z., & Chang, R. (2018). The relationship between drivers' cognitive fatigue and speed variability during monotonous daytime driving. Frontiers in Psychology, 9, 459.

Mabry, J. E., Glenn, T. L., & Hickman, J. S. (2019). Commercial Motor Vehicle Operator Fatigue Detection Technology Catalog and Review.

Macías, R., García, M. A., Ramos, J., Bragos, R., & Fernández, M. (2013). Ventilation and heart rate monitoring in drivers using a contactless electrical bioimpedance system. In Journal of Physics: Conference Series, 434(1), 012047.

Maderthaner, R. (2008). Psychologie. Vienna: Facultas WUV.

Mahmud, M. S., Fang, H., Carreiro, S., Wang, H., & Boyer, E. W. (2019). Wearables technology for drug abuse detection: A survey of recent advancement. *Smart Health*, *13*, 100062.

Mahmud, S. S., Ferreira, L., Hoque, M. S., & Tavassoli, A. (2017). Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs. IATSS research, 41(4), 153-163.

Majumder, S., Verma, A. K., Wang, C., Mohamud, A., Archer, L., Tavakolian, K., & Wilson N. (2019). Using photoplethysmography based features as indicators of drowsiness: preliminary results. Proceedings of the 2019 Design of Medical Devices Conference DMD2019 April 15, 16-18, 2019, Minneapolis, MN, USA.

Marillier, M., & Verstraete, A. G. (2019). Driving under the influence of drugs. Wiley Interdisciplinary Reviews: Forensic Science, 1(3), e1326.

Marquart, G., Cabrall, C., & de Winter, J. (2015). Review of eye-related measures of drivers' mental workload. Procedia Manufacturing, 3, 2854-2861.

Martens, M., & Brouwer, R. (2003). Linking behavioural indicators to safety: What is safe and what is not. In 3rd International conference on Road Safety and Simulation Proceedings, Indianapolis, USA.

Martensen, H., Focant, N., & Diependaele, K. (2016). Let's talk about the weather–Interpretation of short term changes in road accident outcomes. Transportation research procedia, 14, 96-104.

May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. Transportation Research Part F: Traffic Psychology and Behaviour, 12(3), 218-224.

McCarthy, D. P., & Mann, W. C. (2006). Sensitivity and Specificity of the American Medical Association's Assessment of Driving Related Skills (ADReS). Topics in Geriatric Rehabilitation 22(2).

McDonald, A. D., Ferris, T. K., & Wiener, T. A. (2019). Classification of driver distraction: A comprehensive analysis of feature generation, machine learning, and input measures. Human factors, 0018720819856454.

Meesmann, U., Torfs, K. & Van den Berghe, W. (2019) ESRA2 methodology. ESRA2 report Nr. 1. ESRA project (E-Survey of Road users' Attitudes). Brussels, Belgium: Vias institute.

Medtronic. (n.d.). Components | Zephyr™ Performance Systems. Retrieved 27 November 2019, from https://www.zephyranywhere.com/system/components

Michael, P. G., Leeming, F. C., & Dwyer, W. O. (2000). Headway on urban streets: Observational data and an intervention to decrease tailgating. *Transportation Research Part F: Traffic Psychology and Behaviour*, *3*(2), 55–64.

Milo Sensors. (2019). ION™. Retrieved 27 November 2019, from ION™ website: https://www.ionwearable.com

Mitler, M. M., Carskadon, M. A., Czeisier, C. A., Dement, W. C., Dinges, D. F., & Graeber, R. C. (1988). Catastrophes, sleep, and public policy: consensus report. Sleep, 11(1), 100-109.

MobilEye. (n.d.-a). Mobileye 630/60 Mobileye 6-Series User Manual.

MobilEye. (n.d.-b). Mobileye C2-Series Quick Refence Guide.

Mobileye Collision Avoidance System | Mobileye for Fleets. (n.d.). Retrieved December 2, 2019, from https://www.mobileye.com/uk/fleets/products/mobileye-6-collision-avoidance-system/

Mohs R. C., Rosen W. G., & Davis K. L. (1983). The Alzheimer's disease assessment scale: an instrument for assessing treatment efficacy. Psychopharmacology Bulletin, 19 (3), 448-50.

Natarajan, A., Angarita, G., Gaiser, E., Malison, R., Ganesan, D., & Marlin, B. M. (2016). Domain Adaptation Methods for Improving Lab-to-field Generalization of Cocaine Detection using Wearable ECG. *Proceedings of* the ACM International Conference on Ubiquitous Computing. UbiComp (Conference), 2016, 875–885.

Natarajan, A., Parate, A., Gaiser, E., Angarita, G., Malison, R., Marlin, B., & Ganesan, D. (2013). Detecting Cocaine Use with Wearable Electrocardiogram Sensors. *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 123–132.

Nesbit, S.M., Conger, J.C., & Conger, A.J. (2007). A quantitative review of the relationship between anger and aggressive driving. Aggression and Violent Behavior, 12, 56–176.

Neville, F. G., Williams, D. J., Goodall, C. A., Murer, J. S., & Donnelly, P. D. (2013). An Experimental Trial Exploring the Impact of Continuous Transdermal Alcohol Monitoring upon Alcohol Consumption in a Cohort of Male Students. *PLOS ONE*, *8*(6), e67386.

Niezgoda, M., Kamiński, T., & Kruszewski, M. (2012). Measuring driver behaviour-indicators for traffic safety. Journal of KONES, 19, 503-511.

O'Donnel, R. D., & Eggemeier, F. T. (1986). Workload Assessment Methodology In Handbook Of Perception And Human Performance. Vol II.

Olszewski, P., Szagała, P., Wolański, M., & Zielińska, A. (2015). Pedestrian fatality risk in accidents at unsignalized zebra crosswalks in Poland. Accident Analysis & Prevention, 84, 83-91.

Ooi, J. S. K., Ahmad, S. A., Chong, Y. Z., Ali, S. H. M., Ai, G., & Wagatsuma, H. (2016). Driver emotion recognition framework based on electrodermal activity measurements during simulated driving conditions. In 2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES) (pp. 365-369). IEEE.

Otmani, S., Pebayle, T., Roge, J., & Muzet, A. (2005). Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. Physiology & Behaviour, 84, 715–724.

Öz, B., Özkan, T., & Lajunen, T. (2010). Professional and non-professional drivers' stress reactions and risky driving. Transportation Research Part F: Traffic Psychology and Behaviour, 13(1), 32- 40.

Papantoniou, P., Papadimitriou, E., & Yannis, G. (2017). Review of driving performance parameters critical for distracted driving research. Transportation research procedia, 25, 1796-1805.

Park, E., Lee, S. I., Nam, H. S., Garst, J. H., Huang, A., Campion, A., … Sarrafzadeh, M. (2017). Unobtrusive and Continuous Monitoring of Alcohol-impaired Gait Using Smart Shoes. Methods of Information in Medicine, 56(01), 74–82.

Paschalidis, E., Choudhury, C. F., & Hess, S. (2018). Modelling the effects of stress on gap-acceptance decisions combining data from driving simulator and physiological. Transportation Research Part F, 59, 418–435.

Patten, C. J., Kircher, A., Östlund, J., Nilsson, L., & Svenson, O. (2006). Driver experience and cognitive workload in different traffic environments. Accident Analysis & Prevention, 38(5), 887-894.

Paxion, J., Galy, E., & Berthelon, C. (2014). Mental workload and driving. Frontiers in psychology, 5, 1344.

Peißl S., Wickens C.D., & Baruah R. (2018) Eye-Tracking Measures in Aviation: A Selective Literature Review, The International Journal of Aerospace Psychology, 28:3-4, 98-112, DOI: 10.1080/24721840.2018.1514978

Philip, P., & Åkerstedt, T. (2006). Transport and industrial safety, how are they affected by sleepiness and sleep restriction? Sleep Medicine Reviews, 10(5), 347-356.

Philip, P., Sagaspe, P., Moore, N., Taillard, J., Charles, A., Guilleminault, C., & Bioulac, B. (2005). Fatigue, sleep restriction and driving performance. Accident Analysis & Prevention, 37(3), 473-478.

Philip, P., Chaufton, C., Orriols, L., Lagarde, E., Amoros, E., Laumon, B., ... & Sagaspe, P. (2014). Complaints of poor sleep and risk of traffic accidents: a population-based case-control study. PloS One, 9(12), e114102.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and psychopathology, 17(3), 715-734.

Preece, D., Becerra, R., Robinson, K., & Dandy J. (2018). Perth emotion regulation competency inventory (PERCI): Copy of questionnaire and scoring instructions. Retrieved from https://www.researchgate.net/publication/326547664

Pulugurtha, S. S., & Nujjetty, A. P. (2011). Crash Estimation Models for Intersections. In Transportation and Development Institute Congress 2011: Integrated Transportation and Development for a Better Tomorrow (pp. 1306-1314).

Rabe, E., & Gerlach, H. (2005). Praktische Phlebologie. Berlin: Springer.

Rangel, T., Vassallo, J. M., & Herraiz, I. (2013). The influence of economic incentives linked to road safety indicators on accidents: The case of toll concessions in Spain. Accident Analysis & Prevention, 59, 529-536.

Rechtschaffen, A., & Kales, A. A. (1968). Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. UCLA Brain Information Service, Los Angeles, CA.

Rebolledo-Mendez, G., Reyes, A., Paszkowicz, S., Domingo, M. C., & Skrypchuk, L. (2014). Developing a body sensor network to detect emotions during driving. IEEE transactions on intelligent transportation systems, 15(4), 1850-1854.

Regan, M., Gustav Tingvall, C., Healy, D., & Williams, L. (2000). *Trial and evaluation of integrated in-car ITS technologies: Report on an Australian research program*. Retrieved from https://www.researchgate.net/publication/228433270

Reimann, M., & Bechara, A. (2010). The somatic marker framework as a neurological theory of decision-making: Review, conceptual comparisons, and future neuroeconomics research. Journal of Economic Psychology, 31(5), 767-776.

Rodrigues, J. G. P., Kaiseler, M., Aguiar, A., Silva Cunha, J. P., & Barros, J. (2015). A mobile sensing approach to stress detection and memory activation for public bus drivers. IEEE Transactions on Intelligent Transportation Systems, 16(6), 3294-3303.

Rolim, C. C., Baptista, P. C., Duarte, G. O., & Farias, T. L. (2014). Impacts of On-board Devices and Training on Light Duty Vehicle Driving Behavior. *Procedia - Social and Behavioral Sciences*, *111*, 711–720.

Rosen W. G., Mohs R. C., Davis K. L. (1984). A new rating scale for Alzheimer`s disease. American Journal of Psychiatry, 11, 1356-1364.

Russo, F., Busiello, M., Biancardo, S. A., & Dell'Acqua, G. (2014). Assessing Transferability of Highway Safety Manual Crash Prediction Models to Data from Italy. Transportation Research Record, 2433(1), 129-135.

Ryu, J., Rossetter, E. J., & Gerdes, J. C. (2002). Vehicle Sideslip and Roll Parameter Estimation using GPS. - Google Učenjak. *AVEC 2002 6th Int. Symposium on Advanced Vehicle Control*. Retrieved from https://scholar.google.com/scholar?hl=sl&as_sdt=0%2C5&q=Vehicle+Sideslip+and+Roll+Parameter+Estimation+using+GPS.&btnG=

Sabir, M. (2011). Weather and travel behaviour (Doctoral dissertation, Vrije Universiteit).

Sagaspe, P., Taillard, J., Åkerstedt, T., Bayon, V., Espié, S., Chaumet, G., ... & Philip, P. (2008). Extended driving impairs nocturnal driving performances. PloS One, 3(10), e3493.

Sakai, J. T., Mikulich-Gilbertson, S. K., Long, R. J., & Crowley, T. J. (2006). Validity of transdermal alcohol monitoring: Fixed and self-regulated dosing. Alcoholism, Clinical and Experimental Research, 30(1), 26–33.

Sangani, S., Lamontagne, A., & Fung, J. (2015). Cortical mechanisms underlying sensorimotor enhancement promoted by walking with haptic inputs in a virtual environment. Progress in brain research, 218, 313-330.

Sauvet, F., et al. (2014). In-Flight Automatic Detection of Vigilance States Using a Single EEG Channel. IEEE Transactions on Biomedical Engineering, Volume: 61, Issue: 12 , Dec. 2014.

Schmidt, J., Laarousi, R., Stolzmann, W., & Karrer-Gauß, K. (2018). Eye blink detection for different driver states in conditionally automated driving and manual driving using EOG and a driver camera. Behavior Research Methods, 50(3), 1088-1101.

Schmidt, R. F., Lang, F. & Heckmann, M. (2017). Physiologie des Menschen. Berlin: Springer.

Schulze, H., Schumacher, M., Urmeew, R., & Auerbach, K. (2012). *Final Report: Work performed, main results and recommendations.* DRUID (Driving under the Influence of Drugs, Alcohol and Medicines). 6th Framework programme. Deliverable 0.1.8.

Schwarz, J. F., Geisler, P., Hajak, G., Zulley, J., Rupprecht, R., Wetter, T. C., & Popp, R. F. (2016). The effect of partial sleep deprivation on computer-based measures of fitness to drive. Sleep and Breathing, 20(1), 285-292.

Schwarze, A., Ehrenpfordt, I., & Eggert, F. (2014). Workload of younger and elderly drivers in different infrastructural situations. Transportation research part F: traffic psychology and behaviour, 26, 102-115.

Seppelt, B. D., Seaman, S., Lee, J., Angell, L. S., Mehler, B., & Reimer, B. (2017). Glass half-full: On-road glance metrics differentiate crashes from near-crashes in the 100-car data. Accident Analysis & Prevention, 107, 48-62.

Sezgin, E., & Lin, S. (2019). Technology-Based Interventions, Assessments, and Solutions for Safe Driving Training for Adolescents: Rapid Review. JMIR mHealth and uHealth, 7(1), e11942.

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. Front Public Health, 5, 258.

Shi, Q., Abdel-Aty, M., & Lee, J. (2016). A Bayesian ridge regression analysis of congestion's impact on urban expressway safety. Accident Analysis & Prevention, 88, 124-137.

Shi, X., Wong, Y. D., Li, M. Z. F., & Chai, C. (2018). Key risk indicators for accident assessment conditioned on pre-crash vehicle trajectory. Accident Analysis & Prevention, 117, 346-356.

Shiferaw, B. A., Downey, L. A., Westlake, J., Stevens, B., Rajaratnam, S. M., Berlowitz, D. J., & Howard, M. E. (2018). Stationary gaze entropy predicts lane departure events in sleep-deprived drivers. Scientific Reports, 8(1), 2220.

Shinar, D. (2006). *Drug Effects and Their Significance for Traffic Safety*. 52–64. Retrieved from https://trid.trb.org/view/783169

Shinar, D. (1998). Aggressive driving: the contribution of the drivers and the situation. Transportation Research Part F: Traffic Psychology and Behaviour, 1(2), 137–160.

Shinar, D., & Oppenheim, I. (2011). Review of models of driver behaviour and development of a unified driver behaviour model for driving in safety critical situations. In Human Modelling in Assisted Transportation (pp. 215-223). Springer, Milano.

da Silva, F. P. (2014). Mental workload, task demand and driving performance: What relation. Procedia-Social and Behavioral Sciences, 162, 310-319.

Simpson, G. (1987). Accuracy and precision of breath-alcohol measurements for a random subject in the postabsorptive state. *Clinical Chemistry*, *33*(2 Pt 1), 261–268.

Skog, I., Wahlstroem, J., Bonawiede, F., Welch, R., Ohlsson, J., Ohlsson, M., & Händel, P. (2014). Insurance telematics: opportunities and challenges with the smartphone solution. *IEEE Intelligent Transportaion Systems Magazine*. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-143329

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. Risk Analysis: An International Journal, *24*(2), 311-322.

Smith, S. S., Horswill, M. S., Chambers, B., & Wetton, M. (2009). Hazard perception in novice and experienced drivers: The effects of sleepiness. Accident Analysis & Prevention, 41(4), 729- 733.

Socha L., Hanakova L., Socha V., Lali A., Rozenberg R., & Hana K. (2016). Telemetry System Utilization for Stress Monitoring of Pilots During Training. Magazine of Aviation Development 4(20):33-37, 2016.

Sowmya, D., Suneetha, I., & Pushpalatha, N. (2014). Driver behavior monitoring through sensors and tracking the accident using wireless technology. International Journal of Computer Applications, 102(2), 21-27.

Sparrow, A. R., LaJambe, C. M., & Van Dongen, H. P. (2019). Drowsiness measures for commercial motor vehicle operations. Accident Analysis & Prevention, 126, 146-159.

Sparrow, A. R., Mollicone, D. J., Kan, K., Bartels, R., Satterfield, B. C., Riedy, S. M., Unice, A., & Van Dongen, H. P. A. (2016). Naturalistic field study of the restart break in US commercial motor vehicle drivers: truck driving, sleep and fatigue. Accident Analysis & Prevention, 93, 55–64.

Spielberger, C. D., Gorsuch, R. L. & Lushene R. E. (1970). STAI Manual for the State-Trait Anxiety Inventory. Palo Alto: Consulting Psychologists Press.

Stapel, J., Mullakkal-Babu, F. A., & Happee, R. (2019). Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving. Transportation research part F: traffic psychology and behaviour, 60, 590-605.

Staplin, L., Lococo, K. H., Gish, K. W., & Decina, L. E. (2003). Model Driver Screening and Evaluation Program. Volume 2: Maryland Pilot Older Driver Study. Washington, D.C.: National Highway and Traffic Safety Administration. Retrieved 19.12.2019 from: http://www.nhtsa.dot.gov/people/injury/olddrive/modeldriver/

Stedman's Medical Dictionary. (2005). *Drugs* (28th edition). Retrieved from https://stedmansonline.com/public/Learnmore.aspx?resourceID=Medical [04.12.2019].

Stein, G. P., Mano, O., & Shashua, A. (2003). Vision-based ACC with a single camera: Bounds on range and range rate accuracy. *IEEE Intelligent Vehicles Symposium, Proceedings*, 120–125.

Stephan & Newstead (2014). Characteristics of the Road and Surrounding Environment in Metropolitan Shopping Strips: Association with the Frequency and Severity of Single-Vehicle Crashes. Traffic Injury Prevention, 15, 74–80.

Stephens, A. N., Trawley, S. L., Madigan, R., & Groeger, J. A. (2013). Drivers display anger congruent attention to potential traffic hazards: Anger and visual attention. Applied Cognitive Psychology, 27(2), 178–189.

Stipancic, J., Miranda-Moreno, L., & Saunier, N. (2018). Vehicle manoeuvers as surrogate safety measures: Extracting data from the gps-enabled smartphones of regular drivers. Accident Analysis & Prevention, 115, 160-169.

Stockburger, M., & Möckel, M. (2016). EKG-Ableitungen. In N. Suttorp, M. Möckel, B. Siegmund & M. Dietel (Eds.), Harrisons Innere Medizin (19. edition). Berlin: ABW Wissenschaftsverlag.

Stojmenova, K., & Sodnik, J. (2015). Methods for assessment of cognitive workload in driving tasks. In ICIST 2015 5th International Conference on Information Society and Technology (pp. 229-234).

Stojmenova, K., & Sodnik, J. (2018). Detection-Response Task—Uses and Limitations. Sensors, 18(2), 594.

Strano-Rossi, S., Castrignanò, E., Anzillotti, L., Serpelloni, G., Mollica, R., Tagliaro, F., … Chiarotti, M. (2012). Evaluation of four oral fluid devices (DDS®, Drugtest 5000®, Drugwipe 5+® and RapidSTAT®) for on-site monitoring drugged driving in comparison with UHPLC-MS/MS analysis. Forensic Science International, 221(1–3), 70–76.

Stuiver, A., Brookhuis, K. A., de Waard, D., & Mulder, B. (2014). Short-term cardiovascular measures for driver support: increasing sensitivity for detecting changes in mental workload. International Journal of Psychophysiology, 92(1), 35-41.

Summala, H. (1988). Risk control is not risk adjustment: the zero-risk theory of driver behavior and its implications. Ergonomics, 31(4), 491-506.

Sümer, N. (2003). Personality and behavioral predictors of traffic accidents: testing a contextual mediated model. Accident Analysis & Prevention, 35(6), 949-964.

Sundström. A. (2008). Self-assessment of driving skill – A review from a measurement perspective. Transportation Research Part F: Traffic Psychology and Behaviour, 11(1) 1-9.

SWOV. (2015). *The use of drugs and medicines behind the wheel.* [SWOV Fact sheet]. The Hague, The Netherlands: SWOV.

SWOV (2000). Action for advanced Driver assistance and Vehicle control systems Implementation, Standardisation, Optimum use of the Road network and Safety Funding: European (5th RTD Framework Programme) Background & policy context.

Talbot, R., Aigner-Breuss, E., Kaiser, S., Alfonsi, R., Braun, E., Eichhorn, A., … Ziakopoulos, A. (2016). Identification of Road User Related Risk Factors. Deliverable 4.1 of the H2020 project SafetyCube.

Talbot, R., & Filtness. A. (2017). Fatigue – Not Enough Sleep/Driving While Tired, European Road Safety Decision Support System, developed by the H2020 project SafetyCube. Retrieved from www.roadsafety-dss.eu on 01/11/2019

Talbot, R., Hancox, G., Pilkington-Cheney, Filtness, A., Brijs, K., Brijs, T., Ross, V., Katrakazas, C., Al Haddad, C., Taveria, R., Kaiser, S., & Mateus, T. (2020). Safety tolerance zone calculation and interventions for driver-vehicle-environment interactions under challenging conditions Deliverable 3.1 of the Horizon 2020 project i-DREAMS.

Tan, A., Yin, J. D., Tan, L. W., van Dam, R. M., Cheung, Y. Y., & Lee, C. H. (2017). Using the Berlin questionnaire to predict obstructive sleep apnea in the general population. J Clin Sleep Med., 13(3), 427–432.

Tewolde, G. S. (2012). Sensor and network technology for intelligent transportation systems. *IEEE International Conference on Electro Information Technology.*

Thatcher S., & Kilingaru K. (2012). Intelligent Monitoring of Flight Crew Situation Awareness. Advanced Materials Research, Vols. 433-440, 6693-6701.

Thayer, J. F., Ahs, F., Fredrikson, M., Sollers, J. J., & Wager, T.D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. Neuroscience & Behavioural Reviews, 36, 747–756.

Thuy, M., & León, F. P. (2010). Lane detection and tracking based on lidar data. *Metrology and Measurement Systems*, *17*(3), 1.

Tokyoflash Japan. (2019). 'Kisai Intoxicated LCD Watch,' Tokyoflash Japan. Retrieved 26 November 2019, from https://www.tokyoflash.com/en/watches

Torsvall, L., & Åkerstedt, T. (1987). Sleepiness on the job: continuously measured EEG changes in train drivers. Electroencephalography and Clinical Neurophysiology, 66, 502–511.

TruTouch Technologies. (2019). TruTouch. Retrieved 26 November 2019, from TruTouch Technologies website: http://tttinc.com

Tseng, H. E., Ashrafi, B., Madau, D., Brown, T. A., & Recker, D. (1999). The development of vehicle stability control at Ford. *IEEE/ASME Transactions on Mechatronics*, *4*(3), 223–234.

Tso, C., Currie, G. M., Gilmore, D., & Kiat, H. (2015). Electrocardiography: A technologist's guide to interpretation. Journal of nuclear medicine technology, 43(4), 247-252.

UK Department for Transport. (2013). *Driving Under the Influence of Drugs: Report from the Expert Panel on Drug Driving, Wolff, K., Brimblecombe, R., Forfar, J.C., Forrest, A.R., Gilvarry, E., Johnston, A., Morgan, J., Osselton, M.D., Read, L., Taylor, D.* Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/167971/drug-driving-expert-panel-report.pdf

Vaa, T. (2014). From Gibson and Crooks to Damasio: The role of psychology in the development of driver behaviour models. Transportation research part F: traffic psychology and behaviour, 25, 112-119.

Valent, F., Schiava, F., Savonitto, C., Gallo, T., Brusaferro, S., & Barbone, F. (2002). Risk factors for fatal road traffic accidents in Udine, Italy. Accident Analysis & Prevention, 34(1), 71-84.

Van Der Horst, R., & Hogema, J. (1994). Time-to-Collision and collision avoidance systems master. Managing Speeds of Traffic on European Roads View project Observation study conflicts on bicycle lanes View project TIME-TO-COLLISION AND COLLISION AVOIDANCE SYSTEMS. 6th ICTCT Workshop Salzburg Proceedings. Retrieved from https://www.researchgate.net/publication/237807114

Van Dongen, H. P. A., & Belenky, G. (2012). Model-based fatigue risk management. In Matthews, G., Desmond, P. A., Neubauer, C., & Hancock, P. A. (Eds.), The Handbook of Operator Fatigue. Ashgate, Farnham, U.K, 487–506.

Van Dongen, H. P. A., Maislin, G., Mullington, J. M., & Dinges, D. F. (2003). The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep, 2, 117–126.

Veisten, K., Houwing, S., Mathijssen, R., & Akhtar, J. (2011). *Cost-benefit analysis of drug driving enforcement by the police.* DRUID (Driving under the Influence of Drugs, Alcohol and Medicines). 6th Framework programme. Deliverable 3.3.1.

Verster, J. C., Taillard, J., Sagaspe, P., Olivier, B., & Philip, P. (2011). Prolonged nocturnal driving can be as dangerous as severe alcohol-impaired driving. Journal of Sleep Research, 20(4), 585- 588.

Verstraete, A. G. (2004). Detection times of drugs of abuse in blood, urine, and oral fluid. *Therapeutic Drug Monitoring*, *26*(2), 200–205.

Vicente, J., Laguna, P., Bartra, A., & Bailón, R. (2016). Drowsiness detection using heart rate variability. Medical & Biological Engineering & Computing, 54, 927–937.

Villarejo, M. V., Zapirain, B. G., & Zorrilla, A. M. (2012). A stress sensor based on galvanic skin response (GSR) controlled by ZigBee. Sensors, 12, 6075-6101.

Vlahogianni, E. I., & Barmpounakis, E. N. (2017). Driving analytics using smartphones: Algorithms, comparisons and challenges. Transportation Research Part C: Emerging Technologies, 79, 196-206.

de Waard, D., & Brookhuis, K. A. (1991). Assessing driver status: a demonstration experiment on the road. Accident analysis & prevention, 23(4), 297-307.

de Waard, D., & de Groningen, R. (1996). The measurement of drivers' mental workload. Netherlands: Groningen University, Traffic Research Center.

de Waard, D., Kruizinga, A., & Brookhuis, K. A. (2008). The consequences of an increase in heavy goods vehicles for passenger car drivers' mental workload and behaviour: a simulator study. Accident Analysis & Prevention, 40(2), 818-828.

Wahlstrom, J., Skog, I., & Handel, P. (2017, October 1). Smartphone-Based Vehicle Telematics: A Ten-Year Anniversary. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 18, pp. 2802–2825.

Wahlström, J., Skog, I., & Händel, P. (2015). Driving behavior analysis for smartphone-based insurance telematics. *WPA 2015 - Proceedings of the 2nd Workshop on Physical Analytics*, 19–24.

Wang, C., Sun, Q., Fu, R., Li, Z., & Zhang, Q. (2018). Lane change warning threshold based on driver perception characteristics. Accident Analysis & Prevention, 117, 164-174.

Wang, C., Quddus, M., & Ison, S. (2013). A spatio-temporal analysis of the impact of congestion on traffic safety on major roads in the UK. Transportmetrica A: Transport Science, 9(2), 124-148.

Wang, H., Dragomir, A., Abbasi, N. I., Li, J., Thakor, N. V., & Bezerianos, A. (2018). A novel real-time driving fatigue detection system based on wireless dry EEG. Cognitive Neurodynamics, 12(4), 365-376.

Wang, Y., Fridberg, D. J., Leeman, R. F., Cook, R. L., & Porges, E. C. (2018). Wrist-worn alcohol biosensors: Strengths, limitations, and future directions. *Alcohol*.

Wartzek, T., Eilebrecht, B., Lem, J., Lindner, H. J., Leonhardt, S., & Walter, M. (2011). ECG on the road: Robust and unobtrusive estimation of heart rate. IEEE Transactions on Biomedical Engineering, 58(11), 3112-3120.

Watling, C. N., Armstrong, K. A., Smith, S. S., & Obst, P. L. (2016a). Crash risk perception of sleepy driving and its comparisons with drunk driving and speeding: which behaviour is perceived as the riskiest? Traffic Injury Prevention, 17(4), 400–405.

Watling, C. N., Armstrong, K. A., Smith, S. S., & Wilson, A. (2016b). The on-road experiences and awareness of sleepiness in a sample of Australian highway drivers: a roadside driver sleepiness study. Traffic Injury Prevention, 17(1), 24–30.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. Psychological Bulletin, 98, 219–235.

Westerman, S.J & Haigney, D. (2000). Individual differences in driver stress, error and violation. *Personality and Individual Differences*, 29(5), 981–998.

Whiteside S. P., Lynam, D. R., Miller, J. D., & Reynolds, S. K. (2005). Validation of the UPPS impulsive behaviour scale: a four-factor model of impulsivity. European Journal of Personality. 19:559-574

Widen, S. C., & Russell, J. A. (2008). Young children's understanding of other's emotions. Handbook of emotions, 3, 348-363.

Wiethoff, M., Oei, H. L., Penttinen, M., Anttila, V., & Marchau, V. A. W. J. (2002). Advanced driver assistance systems: An overview and actor position. IFAC Proceedings Volumes (IFAC-PapersOnline), 35(1), 1–6.

Wijesoma, W. S., Kodagoda, K. R. S., & Balasuriya, A. P. (2004). Road-boundary detection and tracking using ladar sensing. *IEEE Transactions on Robotics and Automation*, *20*(3), 456–464.

Williamson, A. (2007). Fatigue and coping with driver distraction. In I.J. Faulks, M. Regan, M. Stevenson, J. Brown, A. Porter & J.D. Irwin (Eds.). Distracted driving (611-622). Sydney, NSW: Australasian College of Road Safety.

Williamson, A., Lombardi, D. A., Folkard, S., Stutts, J., Courtney, T. K., & Connor, J. L. (2011). The link between fatigue and safety. Accident Analysis & Prevention, 43(2), 498-515.

Windmiller, J. R., & Wang, J. (2013). Wearable Electrochemical Sensors and Biosensors: A Review. *Electroanalysis*, *25*(1), 29–46.

Worner, T. M., & Prabakaran, J. (1985). The accuracy of breath alcohol analysis using the breathalyzer. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, *20*(3), 349–350.

Wu, Y., Kihara, K., Takeda, Y., Sato, T., Akamatsu, M., & Kitazaki, S. (2019). Effects of scheduled manual driving on drowsiness and response to take over request: A simulator study towards understanding drivers in automated driving. Accident Analysis & Prevention, 124, 202-209.

Xiao, L., & Gao, F. (2010). A comprehensive review of the development of adaptive cruise control systems. *Vehicle System Dynamics*, *48*(10), 1167–1192.

Yannis, G., & Antoniou, C. (2000). *State-of-art on Advanced Driver Assistance Systems* (pp. 4–18). pp. 4–18. Athens: National Technical University of Athens - Aristotle University of Thessaloniki.

Yannis, G., Golias, J., & Antoniou, C. (2004, January). Combining traffic simulation and driving simulator analyses for Advanced Cruise Control system impact identification. In *Proceedings of the 83rd Annual Meeting of the Transportation Research Board* (pp. 11-14).

Zhang, C., Wang, W., Chen, C., Zeng, C., Anderson, D. E., & Cheng, B. (2018). Determination of optimal electroencephalography recording locations for detecting drowsy driving. IET Intelligent Transport Systems, 12(5), 345-350.

Zhang, H., Yan, X., Wu, C., & Qiu, T. (2014). Effect of circadian rhythms and driving duration on fatigue level and driving performance of professional drivers. Transportation Research Record: Journal of the Transportation Research Board, (2402), 19-27.

Zhang, M., Ihme, K., & Drewitz, U. (2019). Discriminating drivers' emotions through the dimension of power: evidence from facial infrared thermography and peripheral physiological measurements. Transportation Reserach Part F, 63, 135-143.

Zhang, M., Ihme, K., & Drewitz, U. (2018). Discriminating driver's fear and frustration through the dimension of power. Proceedings of the 6th Humanist Conference, The Hague.

Zheng, Z. (2012). Empirical analysis on relationship between traffic conditions and crash occurrences. Procedia-Social and Behavioral Sciences, 43, 302-312.

Zhong, X., Hilton, H. J., Gates, G. J., Jelic, S., Stern, Y., Bartels, M. N., De Meersman, R. E., & Basner, R.C. (2005). Increased sympathetic and decreased parasympathetic cardiovascular modulation in normal humans with acute sleep deprivation. Journal of Applied Physiology, 98, 2024–2032.

Zimasa, T., Jamson, S., & Henson, B. (2019). The influence of driver's mood on car following and glance behaviour: Using cognitive load as an intervention. Transportation Research Part F, 66, 87–100.

Zuckerman, M. (2007). The sensation seeking scale V (SSS-V): Still reliable and valid. Personality and Individual Differences, 43(5), 1303-1305.

Annex A: Study review on external context on task demand

Table 26: Effects of external context on task demand

Year	Author	PubType	Specific Type	Context	Focus parts	Indicators	Methods	Measurement tools	Equipment	Transferability to other modes
2018	Foy & Chapman	Analysis		road type 4 road types (arterial A-roads, city-centre multi-lane routes, suburban roads, dual carriageway)		Change in concentration of oxygenated (HbO) and deoxygenated (HbR) haemoglobin mean skin conductance number of skin conductance responses per minute, Heart rate (bpm) respiration rate (breaths per minute), NASA-TLX workload scores (1-20), Eye movement measures of mean fixation duration (ms) horizontal and vertical spread of search (degrees) mean speed (mph) SD of lane position (m) acceleration signatures in each of four directions (m/s/s)			NITES 2 driving simulator fNIRS device FaceLAB 5.0 remote eye tracker BIOPAC equipment	

Year	Author	PubType	Specific Type	Context	Focus parts	Indicators	Methods	Measurement tools	Equipment	Transferability to other modes
2017	Bongiorno et al.	Analysis	workload, road context, traffic flows, weather	two-lane rural road with different traffic flows, road geometry, traffic flows, visibility, vehicle ergonomics	visual behaviour, driver characteristics, fixation time	eye movements, Galvanic Skin Resistance, steer rotation, head movements, ECG, EEG, GSR (max, min, ΔGSR (max-min)), Δt (time gap between the max and min GSR), m (angular coefficient), latency		GPS, GSR sensor, OBD port		similar study in a simulated environment, verify how proper corrections to the road context or the traffic flows can improve the driver's MWL
2015	Auflick	Analysis	driver workload driver distraction	highway	gender differences age differences task type differences task duration, speed, range, range rate, time headway, time to contact	lane keeping longitudinal control eye glance behaviour object-and-event detection data was captured for mean, median, standard deviation, minimum and maximum distances, and time durations (at the minimum or maximum)	Exploratory factor analysis (EFA) techniques Maximum Likelihood Factor Analysis (MLFA) Multivariate Analysis of Variance (MANOVA) Real driving experience	Driver Workload Metrics (DWM)	sensors cameras on-board instrumentation recorded	Confirmatory Factor Analysis (CFA)
2015	Marquart et al.	Review	eye measurement parameters and drivers' mental workload	road/environmental variables (e.g. illumination, temperature)	eye tracking eye activity (blinks, fixations, and saccades) pupillometry	variability of the inter-beat-interval Blink rate Blink duration gaze distributions	electrooculography (EOG) measurement method index of cognitive activity (ICA) questionnaires instantaneous self-assessment (ISA) experiment driving simulator	NASA Task Load Index (NASA-TLX) Rating Scale Mental Effort (RSME) in-vehicle information systems (IVIS) PERCLOS measure of fatigue mean pupil diameter change rate (MPDCR)	head-mounted eye tracker	

Year	Author	PubType	Specific Type	Context	Focus parts	Indicators	Methods	Measurement tools	Equipment	Transferability to other modes
2015	Stojmenova & Sodnik	Analysis	mental or cognitive workload		blink duration blink frequency	eye blinks eyes movements (electrooculogram, EOG) heart rate Pupil diameter electroencephalogr am (EEG) magnetoencephalo graphy (MEG) galvanic skin responses (GSR)	driving simulator Detection response task (DRT) Peripheral detection task (PDT)	electrocardiogr am (EKG)	in-vehicle devices and systems global positioning systems (GPS) interactive displays DRT device	estimate the level of cognitive workload of drivers and sustain it in the predefined and safe range
2014	Stuiver et al.	Analysis	low/high traffic density, weather (fog/no fog),	motorway with three lanes in both directions	mental workload, lane change, short-term response patterns	heart rate variability, blood pressure, cardiovascular reactivity, heart rate, systolic blood pressure variability	ST Software driving simulator, short-term analysis demographic questionnaire electrocardiogra m (ECG) General Linear Model Repeated Measures tests (SPSS) MANOVA	traffic generating protocol (lane scenarios and driver behaviour), navigation system FIN.A.PRES device (Finometer)	three 32-inch diagonal HD plasma screens, three Ag–AgCl electrodes CARSPAN	
4	Schwarze et al.	Analysis	workload of different age groups	real, right-hand traffic	two left-turn lanes at an intersection turning right when pedestrians, cyclists and cars were signaled ‘‘go’’ simultaneously driving through a complex intersection turning left without traffic signs	subjectively perceived effort psychophysiologica l measures age inter-beat-interval between two heart beats (IBI) variability of the inter-beat-interval (RMSSD)	questionnaires test drive t-statistic	electrocardiogr am	ViewCar (Audi A6)	study in a driving simulator examine whether the results can be found for other sequences of situations differing time intervals between two situations involve gender

Year	Author	PubType	Specific Type	Context	Focus parts	Indicators	Methods	Measurement tools	Equipment	Transferability to other modes
2011	Benedetto et al.	Analysis	driver workload	In-vehicle information usage (IVIS)	Eye blink duration	Blink rate Blink duration Average Pupil Size (APS) Reaction time (Lane Change Delay) IVIS performance NASA-Task Load Index Rating Scale for Mental Effort (RSME)	driver simulator Lane Change Test		Oktal SCANer II driving simulator SMI iView X HED head-mounted eye tracker 13x17cm touchscreen display	
2008	De Waard et al.	Analysis	mental workload and behaviour	3 traffic scenarios to investiagte the effects of an increase in HGV. 1) only passenger cars 2) mix: common mix of HGVs and private cars 3)HGV column: A column of HGVs in the slower lane weather: a) clear (bright weather), b) fog: (visibility 150m)		Average speed, SD Speed, Average Lateral position, SD Lateral position, Minimum THW, Minimum TTC, Location Lane Change Heart rate (average BPM), Heart rate variability Experienced risk	driver simulator ECG questionnaires	ST Software The R-peak in the ECG signal was detected with 1ms accuracy. Inter-beat intervals were analysed and the power spectrum of heart rate variation in the 0.10 Hz band were calculated by the programme CARSPAN Rating Scale Mental Effort (RSME)	three frontal 32inc. LCD screens and one additional screen on the left-hand side behind the participant three small Ag/AgCl electrodes attached to the chest	
2006	Patten et al.	Analysis	cognitive workload	driver experience traffic: route complexity		PDT reaction times (ms) PDT miss rates	peripheral detection task (PDT) method field study		Volvo 850S	

Year	Author	PubType	Specific Type	Context	Focus parts	Indicators	Methods	Measurement tools	Equipment	Transferability to other modes
91	De Waard & Brookhuis	Analysis	driving performance impairment	road type, traffic		lateral position of the car steering wheel position speed distance to the cat in front speed of the car in front EEG ECG	road experiment The activation of the subjects as measured by the relative energy parameter [(theta + alpha)/beta]	DEC LSI 11/23 computer at 4Hz DEC LSI 11/23 computer at 4Hz DEC LSI 11/23 computer at 4Hz DEC LSI 11/23 computer at 4Hz DEC LSI 11/23 computer at 4Hz instrumented FM tape recorder, at 125 Hz inter-beat-intervals as intervals between R-tops in milliseconds	Volvo 245 GLD	

Annex B: Study review on indirect effects of road layout, traffic, weather, day and time on task demand

Table 27: Effects of exogenous factors on task demand (Negative effect: positive correlation or increase of task difficulties/accident/risk, Positive effect: negative correlation or decrease of task difficulties/accident/risk)

Year	Author	Effect (positive/negative/percentage)	Context	Road Layout	Traffic	Weather	Day & Time
2016	Da Costa et al.	negative	crash risk	narrow lanes			
2015	Stoker et al.	positive	crash risk	denser street			
2015	Wood et al.	positive	crash risk	larger lane width reductions			
		negative	crash risk	small lane width reductions			
2014	Russo et al.	negative	crash risk	narrow lanes			
2014	Stephan and Newstead	negative	crash risk	primary state arterial road			
2013	Manuel et al.	positive	crash risk	wider lanes/low traffic volumes			
		negative	crash risk	wider lanes/high traffic volumes			
2013	Rangel et al.	negative	accidents, injuries, fatalities	higher number of lanes			
2013	Chengye et al.	negative	crash risk	higher number of lanes			
2012	Ukkusuri et al.	negative	crash risk	higher number of lanes and wider roads			
2011	Ahmed et al.	positive	crash risk	higher number of lanes			
2011	Pulugurtha and Nujjetty	negative	crash frequency	minor right-turn lanes			
2011	Guo et al.	negative	crash frequency	through-traffic per lane on minor roads			
2010	Bergel et al.	negative	crash risk	main roads, secondary roads, motorway, minor roads			
2009	Chen et al.	negative	crash frequency	deceleration lane lengths			
2008	Jones et al.	positive	injured casualties	roads classed as mirror			
2002	Valent et al.	negative	risk of fatal and non-fatal injury	provincial/state road within an urban area			

Year	Author	Effect (positive/negative/percentage)	Context	Road Layout	Traffic	Weather	Day & Time
1999	Bared	positive	crash frequency	acceleration and deceleration lane lengths			
1993	Blower et al.	negative	injury crash rate	major artery road for heavy tractor trailers			
1991	Zegeer et al.	negative	superelevation and crash	road width, spirals			
2016	Shi et al.	negative	crash frequency		congestion, speed variation		
2012	Zheng	negative	crash frequency		free flow, transition, congestion		
2010	Guo et al.	positive	crash frequency		major through-traffic road		
2010	Haleem and Abdel-Aty	positive	crash severity		annual average daily traffic (AADT) on the major road		
2008	Golob et al.	negative	severity		congestion		
2016	Focant et al.	71% negative	accident risk	motorway		frost	
2016	Martensen et al.	negative	injuries			precipitation, frost/snow days, sun, wind	
2014	Focant and Martensen	positive	injury/fatal crashes			fog and rain	
2013	Elvik et al.	negative	injury crashes			rain	
2011	Abdel-Aty et al.	negative	crash risk			fog	
2011	Sabir	negative	injury accident			fog	
2008	Brijs et al.	negative	injury crash			rain intensity, duration	
		negative	accident risk			rainfall	
		positive	injury/fatal crash			snow	
2004	Bergel	negative	injury crashes/fatalities	motorways and main roads		rainfall height	
1998	Edwards	positive	severity			rain	
1991	Fridstrøm and Ingebrigtsen	negative	fatalities/injury crashes			rainfall	
2015	Olszewski et al.	negative	fatality risk/severity of crash				dark/no street lighting, twilight
2013	Gaca and Kiec	negative	risk of crash	national road			morning (05-06h)
		negative	risk of crash	regional road			morning (05-06h)
		negative	risk of crash	national road			evening

Year	Author	Effect (positive/negative/percentage)	Context	Road Layout	Traffic	Weather	Day & Time
							(17-19h)
		negative	risk of crash	regional road			evening (17-19h)
2013	Wang et al.	negative	crash frequency				travel time delay
2009	Johansson et al.	30% negative	crash risk	urban area			darkness
		0% negative	crash risk	rural area			darkness
		40% negative	crash risk	rural/urban area			darkness

Annex C: Study review on measuring attention and distraction

Table 28: Detailed results of literature review on distraction and inattention

Year	Author	PubType (Review/Analysis)	Specific Distraction Type	Focus parts	Indicators	Methods	Measurement tools	Equipment	Distraction Activities
2019	Khan & Lee	Review	Olfactory		Gaze patterns		EEG	commercial sensors include Drypad Sensors, Imotive Headset, MindWave Headsets, NeuroSky's Dry Sensor, Quasar Sensors, Flex Sensors [89,111–115]	
			Gustatory		Head movements				
			Visual		Temperature at the tip of nose				
			Auditory		Skin temperature				
			Biomechanical distraction						
			Cognitive						
2019	Costa et al.	Analysis	Visual	Saccades	SEP, PERCLOS,Blink rates	Eye movement encoder	Eye tracking	Commercial Eye and Head Tracker	Mobile phone
			Manual	Fixation	no hands, one, both hands	Wordbooks manager	Car telematics	Ancho radar for physiological measurements	External events
			Cognitive	Blinks	Standard Deviation of (Left/Right Gaze direction, Head direction), average of (L/R Pupil Dialation, Avg Heart Rate)	Feature extraction & ML	Heart measurements		Interaction with passengers or infotainment
					mean/variance of position and rotation, time for head nose vector directed to each of four quadrants in the fov		Head pose		
2019	Huang et al.	Analysis	Manual	Driver's hands	bimanual motions on the steering wheel, head turn angles, and off-wheel detection	Hand, Head tracking and modelling (Kalman Filters)	magnetometer, gravity sensor, and accelerometer measurements	Smartwatch	
			Visual	Head motions				Hand magnetic ring	

			Incorrect steering control					Head magnetic eyeglasses clip	
2019	Koohestani et al.	Analysis	Visual/General		EDA, heart rate, perinasal perspiration,breathing rate		Palm EDA, Visual Facial Cam, Thermal Facial Cam, Eye Tracking sensor, Adrenergic Sensor, Operational Theater Cam		
2019	Botta et al	Analysis	Visual		Eyes position	Speed, Steering angle, lateral position, yaw rate, lane width, road curvature, heading anlge, accelerator pedal, brake pedal, xy coordinates of car in front, speed of car in front	Camera, CAN		
2019	Aksjonov et al.	Analysis	General		Road Radius, Speed limit, Lane keeping, speed deviation, lane keeping offset	Machine Learning & Fuzzy Logic	GPS, speed sensor		Text messages/mobile phone
2019	Billah et al.	Analysis			Hand, Lips, Forehead & combinations of those with coordinates	Feature classification	Camera	KLT POINT TRACKER, Sony Cyber Shot 14.1 MP camera	Talking, Texting, Eating, Inattention
2019	McDonald et al.	Analysis	Visual, Cognitive, Manual		Driving behavior measures included instantaneous measures of acceleration, brake force, distance, lane offset, lane position, speed, and steering signals	Feature extraction, machine learning classification			
2019	Lohani	Review			driver physiological measures included perinasal perspiration, palm electrodermal activiy, heart rate, breathing rate and eye tracking data			Driving Simulator	
2019	Chui et al.	Analysis	Visual	Head motion	Motion coefficient (1-difference of correlation between two images from a video feed)	Video feeds		NeuroDyne Medical,	Nodding, Head Shaking Moving head Head dropping down Blinking
2019	Khandakar et al.	Analysis	Mobile phone		OBD-II, Accelerometer, Smartphones				
2019	Dehzangi	Analysis			EEG	Dimensonality reduction, feature selection	EEG,ECG,Motion, CAN BUS		

2019	Yadawadkar et al.	Analysis			Lateral acceleration, Low speed, Right Marker, Speed, Lateral distance, Longitudianl distance, Variance of speed	Dimensionality reduction, feature extraction,time-series classificaiton	Video, Vehicle data		
2019	Kanaan et al	Analysis	General/Visual	GPS speed and steering wheel position	Long off path glances >2sec, Secondary task engagement	HMMss			
2018	Kim & Yang (I)	Analysis	Visual		Steering control response			FaceLab 4.6	
			Cognitive					Bio measurements: Zephyr	
					Steering wheel angle				
					Eye Tracking (Gaze: Percentage, Area/Saccades),				
					Heart rate				
					deviation from centerline				
2018	Kim & Yang (II)	Analysis	Visual		Steering control response		Camera, CAN	Logitech Surveillance camera	
			Cognitive		Lateral vehicle motion				
					Heart rate				
					Gaze				
2018	Tarabay et al.	Analysis	Auditory/Vocal		speed, lane position, pedal depression, brake, and reaction time, in addition to physiological measures, such as heart rate and skin conductance level were analyzed at each of the road situations			MEDAC System/3 instrumentation	
2018	Kuo et al.	Analysis	Visual	head pose, gaze and pupil metrics and eyelid opening	Perclos			Seeing Machines automotive-grade driver monitoring system, Mobileye, Blackvue forward-facing camera, steering wheel angle sensors	
2018	Li et al.	Analysis	Visual		Micro Sleep Time, Proportion of Vehicle's lateral offset	Dempster Schafer Evidence Theory and Neural Networks		Two monocular cameras (lane scenarios and driver behaviour)	
2018	Dumitru et al.	Analysis	Visual		Number of glances, Glance duration	ANOVA		Tobii X120 eye tracking system	Facebook distraction
2018	Bakhit et al.	Analysis	Visual	glance behaviour	glance duration, eccentricity penalty function, distraction index, Renewal cycles				See Table 3 in manuscript

2018	Tran et al.	Analysis	General (All)	In Lit Review describes technologies for distraction studies		Image Classification, Deep Learning	, Camera for driver's face Camera for driver's distraction Camera for external environment		Texting, Talking phone, Radio Operations, Drinking, Reaching behind, Hair & Makeup, Talking to passengers
2018	Ali & Hassan	Analysis	General distraction		Rotation angle of head	Machine Learning	Video, Facial components, areas of triangles, angles and motion vectors		
2018	Taherisadr et al.	Analysis	General		ECG	Mel-frequency cepstrum and time-frequency	Video images		Phone
					EEG	Images Feature extraction and classification based on CNN			Texting, Talking phone, Radio Operations, Drinking, Reaching behind, Hair & Makeup, Talking to passengers
					GSR				Q&A
					Motion				
2017	Papantoniou et al.	Review	General distraction		Longitudinal control measures Speed Headway Lateral control measures Lateral posiition Std Dev of Lateral position Steering wheel Reaction time measures Gap acceptance Eye movement measures Physiological measurements Cardiac (heart rate, heart rate variability, blood pressure) Respiratory Eye (horizontal eye movements, eye blink rate, interval of clossure) Speech (pitch,rate,loudness, jitter and shimmer)		GPS, Video-camera, accelerometers, radar and video lane tracking, EEG, EOG		
2017	Seppelt et al.	Analysis	Visual	Glance duration	Mean single glance duration		Atten-D algorithm	camera, glare sensor	
					Count of glances		Eye tracking		
					total glance time				
2017	Hari & Sankaran	Analysis	None (general distraction)	Head position and orientation		Dimensionality reduction		dashboard camera	Non - Distracted (approximately the pose angles are varying from −15◦ to +15◦), Small Distraction to the left side (ap-

									proximately the pose angles are varying from −30◦ to −16◦), High Distraction to the left side (approximately the pose angles are varying from −90◦ to −31◦), Small Distraction to the right side (approximately the pose angles are varying from +16◦ to +30◦) and High Distraction to the right side (approximately the pose angles are varying from +31◦ to +90◦)
						Pose estimation			
2017	Hansen et al.	Analysis	Visual	Head motion and gaze	audio, video, GPS, and IMU sensor signals	Machine learning classifiers	Smartphone application, audio, video, GPS, and IMU sensor signals	Smartphone	Radio, GPS operating, GPS following, Phone operating, Phone Talking, Converstation
2016	Sendra et al.	Analysis	General distraction					2 pressure sensors, 2 temperature sensors on steering wheel, proximity sensors based on LDR installed in the headrest, shock sensors	
2014	Almahasneh et al.	Analysis	Cognitive	lane keeping, crashes	alpha, beta, theta bands		EEG		
2008	Hammoud et al.	Analysis	Visual	Eye positions	head turns	eye tracking and head monitoring	raw video images		

Annex D: Study review on measuring fatigue and sleepiness

Authors	Year	Fatigue or Sleepiness	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
Barua, Ahmed, Ahlström & Begum	2019	Sleepiness	• EEG, • EOG • KSS	EEG and EOG electrodes	Road safety; simulator experiment; proposing an automatic sleepiness classification system; two conditions alert (day drive) and sleep deprived (night drive); three road scenarios; self-assessment on KSS; n = 30	EEG power Blink duration PERCLOS	• Sleepiness detection based on EEG, EOG and contextual information demonstrated using four established classifiers • Presented an automatic driver sleepiness detection system based on EEG, EOG and contextual information	+ Using multiple indicators to detect sleepiness in simulated drives - Applicability of electrodes for real world driving - Intrusive
Buendia et al.,	2019	Sleepiness	• ECG • KSS	ECG using Vitaport 2 and Vitaport 3	Road safety; analysis of motorway driving data; recording ECG to monitor differences in physiology between sleepy and awake drivers; self-assessment on KSS; n = 76	Heart rate variability, average heart rate, NN intervals,	• HR decreased with increasing sleepiness, HRV overall increased, and HRV parameters representing with parasympathetic branch of ANS increased • The parameters representing the sympathetic branch only increased with increasing KSS, which could be due to stress	+ Using ECG to detect sleepiness + Multiple HRV indices + Relationship between HRV and subjective sleepiness - Paper focusing on different methods of detecting outliers - Applicability of ECG electrodes to real world driving
Cori, Anderson, Soleimanloo, Jackson & Howard	2019	Sleepiness (drowsiness)	• Eye blink parameters	EOG, video, infrared oculography	Review paper	Blink frequency, blink duration, PERCLOS, eyelid speed	• Most eye blink parameters varied with drowsiness • Blink duration and PERCLOS most robust • All blink parameters were associated with and predicted conventional drowsiness measures (PVT/subjective/driving tasks)	+ Review of eye blink parameters as an assessment of drowsiness + Association of eyeblink parameters with drowsiness measures - Interindividual differences - Further validation required - Variety of techniques used - More robust for extreme levels of sleepiness rather than mild sleepiness?

Authors	Year	Fatigue or Sleepiness	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
Diaz-Piedra, Gomez-Milan & Di Stasi	2019	Sleepiness	• EEG • Nasal skin temperature • Driving performance • SSS	SOMNOwatch + EEG-6 electroencephal ograph, infrared camera,	Road safety; simulator experiment; assessed validity of nasal temperature to monitor changes in arousal levels; experiment was conducted during the morning; self-assessment on SSS; n = 12	Nasal skin temperature, frontal delta EEG activity, speeding time	• Results suggest nasal skin temperature as a valid method for recognising changes in arousal (from alertness to drowsiness)	+ Non-invasive technique - Focus of camera is on tip of nose, may be issues if people move their head - Research conducted in the morning when alertness is generally high - Small sample size
Liang et al.,	2019	Sleepiness (drowsiness)	• EEG • Driving performance • Eyelid measures	EEG, EOG electrodes, in-vehicle investigator	Road safety; field test driving; 2h instrumented drive, after a night shift and after a night of rest; aimed to build a predictive model of drowsiness events; n = 16	EEG characterised sleep episodes, lane crossings, PERCLOS	• Overall the best models for both measures of drowsiness were those that considered driver individual differences and eyelid measures. • These measures should be considered when predicting drowsiness events.	+ Using multiple indicators to detect sleepiness in simulated drives + Real scenario driving with shift workers + Results can benefit the development of real time drowsiness detection - Focusing on developing models to predict driver sleepiness - Intrusive (electrodes)
Sparrow, LaJambe & Van Dongen	2019	Sleepiness (drowsiness)	• EEG • Ocular measures • Cardiac measures • Performance measures	EEG, EOG, eyeglasses, ECG, FIT,	Review paper	EEG activity, PERCLOS, blink rate, blink duration, pupil size, HRV, PVT, driving performance	• EEG and ocular parameters can be limited in detecting lower levels of drowsiness • Cardiac measures can be confounded by other influences and not widely used in operational settings for sleepiness • Vigilance attention and performance measures require driver involvement • Subjective measures relatively easy to get, however associations with objective performance can vary • Individual differences	+ Review of different drowsiness measures used in commercial vehicles + Systems to capture drowsiness may need to combine different measures - Interindividual differences - Variety of techniques used - EEG and ocular measures more robust for extreme levels of sleepiness rather than mild sleepiness? - Cardiac measures not widely used as a measure of drowsiness

Authors	Year	Fatigue or Sleepiness	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
Wu et al.,	2019	Sleepiness (drowsiness)	• EOG • KSS • Driving performance	EOG electrodes	Road safety; simulator experiment; investigating the effects of manual driving on driver drowsiness and performance; three conditions (automated driving for 3min; 31min; 10min -10min manual driving-automated driving for 10min); self-assessment on Japanese version of KSS; n = 115	Eyeblink duration, reaction time, time to steer, time to brake, standard deviation of steering wheel angle, minimum Time-to-Collison, KSS	• Driver sleepiness decreased when scheduled manual driving began, but effects only lasted for a short duration (4-6min) • Older drivers reacted significantly more slowly in steering and braking with scheduled manual driving	+ Using multiple indicators to detect sleepiness in simulated drives + Potential differences in age - Applicability of electrodes for real world driving - Intrusive - Short task switching durations - No information on inducing sleepiness or time of day
Ahlström, Anund, Fors & Åkerstedt	2018	Sleepiness	• EEG • EOG • Driving performance • KSS	EEG and EOG electrodes	Road safety; simulator experiment; to investigate the effect of light conditions on driver sleepiness; conditions daylight vs darkness and daytime vs night-time; self-assessment on KSS; n = 30	EEG activity, eyeblink duration, line crossings, lateral position, speed	• KSS and blink durations increase with sleep deprivation. Darkness also has an effect • Light had an independent effect KSS, lateral position, EEG activity, blink duration • The day/night condition had profound effects on lane crossings, KSS, blink duration, speed	+ Using multiple indicators to detect sleepiness in simulated drives + Differences in day/night - Applicability of electrodes for real world driving - Intrusive
Ahlström, Anund, Fors & Åkerstedt	2018	Sleepiness	• EEG • EOG • KSS • Driving performance	EEG and EOG electrodes	Road safety; simulator experiment; aim to compare two road environments and their effects on driver sleepiness; conditions rural road low traffic density vs suburban road with higher traffic density, daytime vs night-time; n = 30	EEG activity, blink duration, line crossings, relative speed, steering activity, pedal activity	• Only minor effects of road environment • Increased subjective sleepiness, longer blink duration and increased EEG alpha content due to time on task and night-time driving	+ Using multiple indicators to detect sleepiness in simulated drives + Increased sleepiness from time on task and night-time driving, relevant factors to be considered - Applicability of electrodes for real world driving - Intrusive

Authors	Year	Fatigue or Sleepiness	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
Ahlström, Anund & Kjellman	2018	Sleepiness, fatigue, stress	• EOG • ECG • EDA • Eye tracking • KSS • Video recording	Smart Eye Pro 7.0, EOG, Vitaport 2, EDA wrist device	Road safety; real world driving, exploratory study to investigate fatigue and stress levels of city bus drivers; no manipulation of stress of sleepiness levels; data collected during morning shift and during the afternoon shift; fatigue function of time on task; self-assessment on KSS; n = 15	Blink durations, KSS, HRV, EDA signals	• Without manipulation, during ordinary daytime bus route there are instances of sleepiness and stress in some individuals • Low reported KSS levels – understanding or comfort with verbal reporting? • Individual differences relate to use of personalised algorithms	+ Bus drivers + Paper states the importance of context when analysing visual behaviour + importance of establishing 'on-road region' when using eye tracking other false reports of distraction + Importance of considering external factors alongside physiological measures - Sleepiness function of KSS, fatigue function of time on task - Exploratory study
Aidman et al.,	2018	Sleepiness (drowsiness)	• Ocular parameters • Driving performance	Optalert glasses and Alertness Monitoring System - infrared sensor measuring ocular parameters and converting to JDS every 60s	Road safety; simulator experiment; study aimed to examine the effects of repeat dose caffeine on sleepiness and driving performance; 50h sleep deprivation; placebo or caffeine group; n = 11	JDS (Johns drowsiness scale); lane keeping, speed maintenance	• Sleepiness increased and driving performance declined during the study • Lateral lane position and speed variability associated with JDS scores and sleepiness • Caffeine mitigated impairments in driving performance – reducing sleepiness and weakening impact on driving errors	+ Sleepiness measures associated with driving performance +Optalert glasses more usable in real world driving compared to electrodes - Requires the driver to wear Optalert glasses - 50h sleep deprivation not likely a real-world scenario
Anund, Ahlström, Fors & Åkerstedt	2018	Sleepiness	• EOG • KSS • Driving performance • Reaction time	EOG electrodes, PVT reaction time task	Road safety; simulator experiment; investigate differences in sleepiness in professional and non-professional drivers; day vs night drive to induce sleepiness; three road scenarios driven in succession (rural low demand daylight, rural low demand darkness, suburban	Blink duration, KSS, line crossings, speed, mean PVT reaction time and percentage of lapses	• Professional drivers self-report significantly lower sleepiness than non-professional drivers • Professional drivers showed longer blink durations, more line crossings and drove faster	+ Sleepiness measures associated with driving performance + Differences in professional and non-professional drivers may be relative factors to consider + Differences in self-reported sleepiness and objective measures of sleepiness

Authors	Year	Fatigue or Sleepiness	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
					high demand daylight; self-assessment on KSS; n = 30			- Applicability of electrodes for real world driving - Intrusive
Anund, Fors, Ihlström & Kecklund	2018	Sleepiness	• EEG • EOG • KSS • PVT	Vitaport 2 for electrophysiological measurements, electrodes, Vbox for speed and GPS position, video cameras	Road safety; real road driving with bus drivers; investigated the effect of spilt shift working on sleepiness and driving performance during the afternoon; one drive during a split shift and one drive after being off duty for the morning; self-assessment on KSS; n = 18	Blink duration, KSS, reaction time, EEG based Karolinska Drowsiness Score	• Increased sleepiness associated with split shift schedules • Strong individual differences	+ Real road driving with a sample of shift workers + Individual differences in development and experience of sleepiness + Bus drivers - Applicability of electrodes for real world driving - Intrusive
Ariansyah, Caruso, Ruscio & Bordegoni	2018	Fatigue (mental workload)	• Cardiac activity • Skin response • Driving performance • NASA-TLX • SOFI-20	Thought Technology physiological sensor biograph infinity system, BVP sensor	Road safety; simulator experiment; aim to investigate whether a measure of cardiac activity could show differences of different driving conditions on workload; car following task with visual ADAS; monotonous (constant speed) vs active (variable speed) condition; self-assessment on NASA-TLX and SOFI20; n = 14	Mean heart rate, HRV, skin conductance, respiration rate, blood volume pressure, lateral position, steering wheel movement, NASA-TLX, SOFI-20	• Workload increased over time regardless of driving condition • Main effect of driving condition resulted in higher level of sympathetic activation during variable speed driving	+ Obtaining heart rate measures through a sensor on the index finger + Focus on fatigue rather than sleepiness + Increased heart rate measures linked to increased workload - Likelihood of drivers consistently wearing finger sensor?
Balasubramanian & Bhardwaj	2018	Fatigue (cognitive fatigue)	• cECG • EEG	EEG electrodes, eECG electrodes in seat (non-contact ECG system)	Road safety; simulator experiment; estimating driver fatigue based on cECG and EEG; correlate eECG and EEG signals to established method to analyse driver fatigue; driving in simulated mild traffic for 120min continuously; n = 35	HRV, EEG power activity	• Estimated coherence between ECG and EEG signals found to be good • During simulated driving, the changes in EEG power bands is proceeded by changes in cardiac activity	+ Estimations of fatigue using eECG and EEG + eECG correlated with EEG + eECG placement more applicable for real world + Focus on fatigue rather than sleepiness - Drivers only completed one 120min drive

Authors	Year	Fatigue or Sleepiness	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
Caponecchia & Williamson	2018	Sleepiness (drowsiness)	• Eye tracking • Driving performance • Sleep patterns	Smart Eye Pro tracking system, actiwatch	Road safety; simulator experiment; investigated the effects of mild sleep deprivation on driving performance; three groups (no deprivation, 2h deprivation, 4h deprivation); 45min drive in morning and evening; n = 45	PERCLOS, blink duration, lane deviation, speed	• Measurements of eye closure didn't show sleepiness in drivers despite performance impairments • During afternoon, drivers reported more sleepiness • Increased lane deviations during the morning following sleep deprivation	+ Sleep deprivation resulted in increased lane deviations - No significant effects of sleep deprivation on measures of blink duration or PERCLOS. - Sensitivity of the system?
Choi, Koo, Seo & Kim	2018	Sleepiness (drowsiness), fatigue and stress	• PPG • GSR • Driver movement • Driver facial measures	Wrist worn manually built wearable device using commercially available parts, camera, additional ear clip for PPG	Road safety; simulator experiment; investigating a system to detect and distinguish driving conditions; drivers drove four conditions (normal, stressed, 'drowsy', fatigued – stress induced by increased traffic, drowsy induced by monotonous driving, fatigue last drive) n = 28	Changes in blood volume, skin conductance, temperature, facial signs of sleepiness, acceleration, rate of rotation	• Device distinguished stress, fatigue and sleepiness from normal driving condition • Classification accuracy 98.43% for cross validation on the data • Accuracy 68.31% for the four conditions • Reported accuracy of 84.46% if drowsy and fatigue same condition	+ Non-intrusive wearable device + Distinguished differences in driving conditions - Were the conditions enough to induce the different states? - Unlikely to measure sleepiness - Fatigue and sleepiness/drowsiness not the same state - Required pre-processing of the data
Darzi, Gaweesh, Ahmed & Novak	2018	Sleepiness (drowsiness), fatigue (with distraction, stress and high workload)	• ECG • GSR • Driving performance • Personality stress • Mood • Workload	ECG electrodes, GSR sensors on palm of hand, respiration sensor on nose, temperature sensor on little finger	Road safety; simulator experiment; investigate whether cause of driver's hazardous state can be identified by combination of vehicle data, driver characteristics and physiological measures; self-report on IPIP, PSS-10, STAQ, NASA-TLX; eight different scenarios inc. weather, traffic density, cell phone use; mild sleep deprivation induced by less than 6h sleep; drivers had	Mean heart rate, SD of IBI, ECG signal mean and gradient, skin conductance, skin temperature, mean respiration rate, lane position, velocity, throttle force, slip level of front tyres	• Classification for sleep deprivation 98.8%, traffic density 91.4%, cell phone use 82.3% and weather 71.5% • Vehicle data most useful for classification of weather and traffic density • Physiology and driver characteristics most useful for classification of sleep deprivation and cell phone use	+ Various measures to detect driver state + Different types of info lead to higher classification accuracy - Applicability of electrodes for real world driving - Intrusive - Results showed driver characteristics e.g. mood useful to classify drowsy state rather than physiological measures

Authors	Year	Fatigue or Sleepiness	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
					four sessions (2x drowsy 2x alert); n = 21			
Fors, Ahlström & Anund	2018	Sleepiness	• EOG • KSS • Driving performance	EOG electrodes, Vitaport 3, Smart Eye Pro 5.7 system	Road safety; simulator and real road comparison; daytime vs night-time driving condition and real road vs simulator condition; self-assessment on KSS; n = 16	Blink duration, percentage of road centre gaze, speed, line crossings, KSS	• Simulator driving resulted in higher KSS, longer blink durations, lower percentage gazes to centre of road and higher speed • Night drive showed increased lane crossings, increased KSS, higher blink durations	+ Differences in simulator and real-world driving + Smart Eye Pro system non-intrusive + Well use measures of sleepiness - Applicability of electrodes for real world driving - Electrodes intrusive
Lees et al.,	2018	Fatigue and sleepiness	• EEG • PSQI • KSS • ESS • CIS20	QuikCap EEG electrodes	Safety; simulator experiment; aimed to investigate the capability of monopolar EEG analysis as a prediction of fatigue/sleepiness; data collected between 1000-1400; n = 63	EEG activity, KSS, PSQI, ESS, CIS20	• Self-reported sleepiness mainly associated with EEG delta, theta and alpha variables • Self-report sleepiness predicted to varying degrees of success by changes to monopolar EEG variables	+ Train drivers - EEG data collection conducted during active phase (10min of driving) - Applicability of EEG cap to real world train driving - Was time of day enough to elicit sleepiness? Length of task enough to elicit fatigue?
Ma, Gu, Jia, Yao & Chang	2018	Fatigue and sleepiness	• EEG • Eye tracking • Driving performance • SOFI	Tobii Glasses II eye tracking system, EEG cap	Road safety; simulator experiment; investigated the effects of speed variability on driver fatigue; monotonous following 60min drive; self-assessment on SOFI-C (Chinese version); n = 21	EEG activity, pupil diameter, lane position, speed, car following distance, subjective fatigue	• 60min monotonous driving elicited driver cognitive fatigue (underload) • Differences in speed variability groups resulted in differences in physiological measures of sleepiness	+ Eye tracking glasses non-intrusive - Applicability of EEG cap for real world driving - EEG cap/electrodes intrusive - Only one eye tracking measure - Fatigue mainly measured using self-assessment - May interfere with glasses

Authors	**Year**	**Fatigue or Sleepiness**	**Indicators**	**Technical equipment**	**Context & Design**	**Outcome variables**	**Results**	**Conclusion** (for i-DREAMS)
Schmidt, Laarousi, Stolzmann & Karrer-Gauss	2018	Sleepiness (drowsiness)	• EOG • Eye tracking • KSS	AntiCAP, head mounted eye tracker Dikablis, electrodes	Road safety; simulator experiment; aim to evaluate the performance of an EOG and camera-based blink detection process; between subjects manual vs automated condition; 6pm or 10pm drives; comparison of eye blink detection algorithms; self-assessment on KSS; n = 30 (14 monotonous manual, 16 monotonous automated driving sessions)	EOG signals, real eye closure events, KSS	• Blinking behaviour significantly affected by drowsiness • Automated driving also impacts blinking behaviour	+ Eye tracking measures as a detection of drowsiness - Paper aimed to compare difference algorithms - Applicability of EOG and head mounted eye tracker for real world driving - EOG cap/electrodes intrusive
Shiferaw et al.,	2018	Sleepiness (drowsiness)	• Eye tracking • Driving performance • KSS • PVT	Cap mounted eye tracking system (SensoMotoric Instruments)	Road safety; real driving on a closed track; investigate gaze behaviour as an indicator of drowsiness; one drive following sleep deprivation and one drive following normal sleep; self-assessment on KSS; n = 9	Gaze behaviour, blink duration, saccade amplitude, lane departure, KSS, reaction time	• Sleep deprivation resulted in increased blink duration and saccade amplitude; rate of fixations reduced • Increased stationary gaze entropy associated with increased odds of lane departure	+ Association of eye tracking measures and driving impairment + Use of gaze direction and scanning behaviour as indicators of drowsiness - Small sample - Applicability of cap to real word driving
Wang et al.,	2018	Fatigue	• Wireless, dry EEG • Driving performance	Dry EEG cap	Road safety; simulator experiment; aim to develop a method of fatigue detection based on dry EEG signals; two sessions of 90min drive no break; self-assessment on NASA-TLX; n = 10	EEG activity, reaction time	• Dry EEG method with NASA-TLX response indicated changes in mental fatigue • Changes in reaction time consistent with fatigue prediction	+ Wireless, dry electrodes - Small sample - Applicability of cap to real world driving - Results based on two 90min drives
Zhang et al.,	2018	Sleepiness (drowsiness)	• EEG • Driving performance	EEG cap	Road safety; simulator experiment; car following task; study conducted between 1000-1700; n = 22	EEG activity	• Sensitive index to detect EEG changes during monotonous driving • Differences in EEG activity before and after driving • Differences found in temporal and frontal locations	- Applicability of cap to real world driving - Was time of day enough to elicit drowsiness?

Authors	Year	Fatigue or Sleepiness	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
Ahlström, Jansson & Anund	2017	Sleepiness	• EEG • Driving performance • KSS	EEG electrodes	Road safety; exploratory simulator study; investigate whether lane departures are associated with local sleep measured by EEG; drives on six occasions, three drives (including changes in road, traffic density and daylight and night-time), self-assessment on KSS; n = 30	EEG activity, lane departures, KSS	• The number of lane departures increased exponentially with KSS • Global theta power lower for KSS ≤5 compared to KSS 9 • No correspondence between global theta and lane departures • Unit increase in theta power in certain brain regions associated with increase in odds of departing the road	+ Changes in local sleepiness associated with driving impairments - Applicability of electrodes to real word driving - Exploratory study
Anund, Fors & Ahlström	2017	Sleepiness	• EOG • Driving performance • KSS	EOG electrodes, Vitaport system	Road safety; simulator study; investigate differences in daytime and night-time driving during self-reported sleepiness and long blink durations, in terms of line crossings; two drives; n = 16	Blink duration, lateral position, KSS, time on task	• No difference in % line crossings during day and night when reported high LSS • Significant difference in percentage line crossings between day and night during long blink durations • KSS as predictor of line crossings most promising measure	+ Associations between eye tracking measures and subjective sleepiness with driving performance + Driver awareness of own sleepiness - Applicability of electrodes to real word driving - Subjective measure most promising predictor of driving impairment (line crossings)
He et al.,	2017	Sleepiness (drowsiness)	• Eye tracking • Driving performance • KSS • SSS	Google glass drowsiness detection system	Road safety; simulator study; development and testing of detection system; comparison of driving performance and eye blinks; study conducted between 0800-2000; three driving sessions continuously; self-assessment on KSS and SSS; n = 23	Blink frequency braking response time, headway, lane position, lane excursions KSS, SSS	• Drivers classed as drowsy had longer braking responses, lower braking response rates, increased lane deviations and lane excursions	+ Uses sensors of commercially available product to monitor eye blinking frequency + Associations between drowsiness and driving impairments - Device interferes with glasses - Only used subjective ratings to measure sleepiness - Needs further validation

Authors	Year	Fatigue or Sleepiness	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
Ahn, Nguyen, Jang, Kim & Jun	2016	Sleepiness	• EEG • ECG • EOG • fNIRS	Electrodes, webcam, custom built fNIRS system	Road safety; simulator study; two conditions, well rested and sleep deprived (stay up all night); driving occurred before 9am for 30min; self-assessment on subjective questionnaire; n = 11	EEG activity, heart rate (R-peak and RR-peak), changes in oxy-haemoglobin and haemoglobin concentration, eye blinking	• Significant features in EEG, ECG and fNIRS data between the conditions • Heart rate lower in sleep deprived drivers • EOG showed high variability in the data	+ Detecting sleepiness through several measures + Combination of several measures improves the classification accuracy - Applicability of electrodes to real world driving - Paper focuses on developing an algorithm - Small sample size
Alvaro, Jackson, Berlowitz, Swann & Howard	2016	Sleepiness	• Ocular measures • Driving performance • PVT • KSS • ESS	Video recording, PVT	Road safety; simulator study; aimed to describe the duration and frequency of eyelid closure during acute sleep deprivation; 24h sleep deprivation; test battery occurred seven times during 24h period; self-assessment on the KSS and ESS; n = 20	PERCLOS, eye closure duration, lane position, speed, braking reaction time, number of crashes, KSS, ESS	• Frequency and duration of eyelid closure increased with acute sleep deprivation • Eyelid closure increased and became more frequent after 17h and 20h of wakefulness lasting up to 18 seconds • Lateral lane position, breaking RT, crashes, vigilance and subjective sleepiness correlated moderately to high with length of eyelid closure	+ Ocular measures associated with increased wakefulness + Automated systems often use averaged durations which may conceal instances of prolonged eyelid closure - Manual scoring of ocular measures - 24h sleep deprivation quite extreme, although shift workers may experience close to this on first night shift
Jackson et al.,	2016	Sleepiness (drowsiness)	• Slow eyelid closure • PVT • Driving performance • KSS	Copilot video based system, PVT,	Road safety; simulator study; aimed to determine whether changes in eyelid closure occurred following acute sleep deprivation; two conditions (normal night sleep 8h TIB and 24h total sleep deprivation); driving at 1000; self-assessment on KSS n = 12	PERCLOS, reaction time, lapses, lane position, speed, braking, crashes, KSS	• 24h sleep deprivation resulted in significantly more eyelid closure, greater lane position variability, increased attentional lapses • PERCLOS moderately associated with variability in vigilance performance and lane position variation	+ Ocular measures detected impairment due to sleep deprivation + Automated measure of sleepiness - Acute (24h) sleep deprivation - Small sample

Authors	**Year**	**Fatigue or Sleepiness**	**Indicators**	**Technical equipment**	**Context & Design**	**Outcome variables**	**Results**	**Conclusion** (for i-DREAMS)
Borghini, Astolfi, Vecchiato, Mattia & Babiloni	2014	Fatigue, sleepiness (drowsiness), mental workload	• EEG • EOG • Heart rate	Wireless wearable EEG	Review paper	EEG power, HRV, blink rate, blink duration	• Changes in EEG, EOG and heart rate associated with changes in workload, fatigue and drowsiness • Drowsiness characterised by increased blink rate and decreased heart rate	+ Reviews several measures and mental states - Paper focuses on pilots and a few drivers - HR more associated with mental workload?
Filtness et al.,	2014	Sleepiness	• EOG • Driving performance • KSS	EOG using Vitaport 3, Smart Eye Pro 5.7	Road safety; simulator and real road driving; four drives on two occasions; simulator and real road; daytime (afternoon/eve) and night-time; eye symptoms informed from focus groups and Accumulated Time with Sleepiness scale; self-assessment on KSS; n = 16	Blink duration, ten eye symptoms (sore, itching, gravel, pain, strain, difficulty focusing, tearful, heavy eye lids, struggle to keep eyes open, dry), KSS, line crossings	• Four symptoms reflective of driver sleepiness (eye strain, difficulty focusing, heavy eyelids and difficulty keeping the eyes open) • Sore eyes and tearful associated with sleepiness in simulator • Eye symptoms associated with increased subjective sleepiness and driving impairments	+ Simulator and real road study + Range of ocular measures + Differences in real road and simulator driving + Subjective reports of sleepiness and associated eye symptoms + Associations of eye symptoms with subjective sleepiness and impaired performance - Applicability of EOG for real world driving - Eye symptoms rated after each drive
Åkerstedt et al.,	2013	Sleepiness	• EEG • EOG • Driving performance • KSS	EEG, EOG, EMG, ECG using Vitaport 3	Road safety; real road driving; aimed to describe the development of sleepiness indicators; day (afternoon) and night motorway drive of 90mins; self-assessment on KSS; n = 18	EEG activity, blink duration, line crossing, KDS, KSS, speed, lateral position,	• Those that terminated the drive showed high sleepiness ratings, and higher levels of sleep intrusions on EEG/EOG • Night drive showed significant increases in all indicator's vs day drive • Blink duration differed between night and day conditions	+ Real road driving + Multiple measures of sleepiness - Study not designed to identify variables to detect drowsiness - Focused on the pattern of development of sleepiness in variables - Applicability of electrodes to real world driving

Authors	Year	Fatigue or Sleepiness	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
Golz, Sommer, Trutschel, Sirois & Edwards	2010	Sleepiness	• EEG • EOG • KSS • Driving performance	22 devices to monitor driver sleepiness, three video-based devices	Review paper and evaluation of three video-based devices; overnight driving simulator study; eight test runs with 10min breaks in-between; self-assessment on KSS; n = 14	PERCLOS, EEG, KSS, lane position	• PERCLOS associated with higher KSS and SD of lateral position in lane • PERCLOS has difficulty differentiating between mild and extreme sleepiness • Devices may not provide valid predictions of subjective sleepiness and driving performance on individual level	+ Reviewed several 'fatigue monitoring devices' + Evaluated three technologies in simulator study - Technologies available in 2006 - Experimental findings - Interindividual differences

Annex E: Study review on measuring emotions

Table 29: Studies reviewed on measuring emotions through behavioural and physiological indicators. Sample size is considered small when n<20

Authors	Year	Theoretical constructs	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
Agrawal et al.	2013	• Happiness • Anger • Sadness • Surprise	Facial features (eyes, lips)	DSC-S5000 camera, Nikon Coolpix-L21 camera	HFE; suggestion of a fuzzy system to detect emotions and facial gestures in combination with attentive state	n.a.	• 90% accuracy for emotion detection • 94.58% accuracy when simultaneously detecting facial expression	+ Combination of emotional and attention state – Lacking theoretical concept of emotions – Not clear which visual material used; emotions seem not induced but facially expressed
Al Machot et al.	2012	• Sadness • Anger	Speech recognition signals (12 features)	Microphones (no type or brand reported)	HFE; Berlin emotional speech database for data training; Bayesian Quadratic Discriminate Classifier	n.a.	Total accuracy of three emotions (including ‘neutral’) is 86.67%	– Lower accuracy rates compared to more recent studies – Not tested for robustness
Ali et al.	2016	Arousal and valence (four quadrants)	• ECG • EDA • Skin temp.	(No new data acquired)	Road safety; ADAS driver emotion recognition, test of neutral network (CNN); n=30; induction of emotions via video clips; manipulation check with Mankins scales of valence and arousal	n.a.	• 92.4% accuracy of CNN based emotion detection • EDA signals provide best classification results	+ EDA signals extracted by non-intrusive method, seems to gain good results – Prototype to be developed and tested with real persons
Balters et al.	2019	Stress (autonomic arousal as proxy for stress)	• Respiration • HR, HRV • EDA	• ECG: Zephyr BioModule • EDA: Empathica E4 bracelet	Road safety; experiment in real traffic (experimental car); stress induced driving behaviour (changes in speed, acceleration, braking, lane keeping, steering reversal rates), n=16 commuters; stress rated on ‘Perceived Stress Scale’	n.a.	Gained knowledge of in-car real-time stress management intervention. No results available so far	+ Real world driving – No results available yet – Small sample size

Authors	Year	Theoretical constructs	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
Barnard & Chapman	2016	Fear (Arousal as indicator of fear/threat)	• EDA • HR • Eye movements (saccadic amplitude)	• Biopack variable response transducer (GEL101, TSD115) • iView X RED eye tracker	Road safety; measuring situational threat (arousal) including participants with different levels of *trait* anxiety; arousal induced via videos of accidents; mixed design, two within subject factors, n=57	• Eye movem. • Physiol. indicators	Perception of fear increased with increased accident level and skin conductance can be used as reliable measurement	+ SC could serve as somatic marker for fear – Participants were students, university personal
Bailenson et al.	2007	• Amusement • Sadness	• EDA (SCL) • Body temperature • HR, Pulse • Blood pressure • Facial expression	12-channel Grass Model 7 polygraph	Human-computer studies; automated real-time emotion-recognizer model for data analysis; videos from n=41; emotions induced via films	Facial features Physiological indicators	• Good statistical fit of algorithms to predict emotions from facial expressions and physiological measurements • Physiological measures alone perform better than facial expressions alone	– Dated study – More advanced algorithms may be available – Model includes only two emotions
Chan & Singhal	2015	• Negative emotion • Positive emotion	EEG sign.	Geodesic Sensor Net	Road safety, driving simulator experiment (two-lane, bidirectional highway in rural setting); emotions induced by words from the Affective Norms for English Words database (valence value, arousal value); n=25	• Speed • Lateral control • RT	• Mean speed and lateral control sign. reduced in with negative words compared to other cond. • Negative emotional auditory content may reduce safe driving • RT was shorter in positive emotion than in neutral emotion cond.	– No manipulation check – No details on measuring EEG signals reported
Dobbins & Fairclough	2018	• Anger • Stress	• ECG, IBI, HR, HRV • PPG, Peak-to-peak interval, pulse transit time	• 5-lead ECG unit from Shimmer3 • Optical pulse ear (PPG) clip from Shimmer3	Road safety; on-road experiment (commuting, camera based: number of lanes, type of road, traffic density); n=21; self-reported stress, anger: STAXI2, UWIST Mood Adjective Checklist	• Self-report of emotions • Physiol. indicators of emotions	• Deriving labels for stress and anger based on physiological indicators for the purpose of machine learning was as accurate as based on self-reports • HR: lowest rate of false alarms • Pulse transit time: lowest rate of misses	+ Real world driving – No details/results on indicators of anger and stress provided – No conceptional definition of anger, stress
Fan et al.	2010	• Happiness • Anger	EEG signal	192 channel digital brain wave measurement system from	HFE; n=9; emotional states induced by videos of traffic situations; manipulation check via	n.a.	Accuracy of Bayesian network model to detect emotions of 78.17%	– Accuracies have increased in more recent years

Authors	Year	Theoretical constructs	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
				NEURO Company	self-assessment (designed by authors)			– Not validated with further material – Small sample size
Gao et al.	2014	• Anger • Disgust as stress related emotions	Facial expression	NIR-camera on dashboard	HFE; data set 1 recorded in office setting, n=21; data set 2 recorded in in-car n=12; participants were asked to facially express the emotions in question	n.a.	Monitoring system to detect emotions works well on two simulated data sets; with detection rate of 85% in-car	– Emotions were not induced but facially expressed – No details on NIR measure – Small sample size
Girardi et al.	2017	Arousal, valence	• EEG • EMG • EDA	• BrainLink headset • Shimmer GSR+Unit • Shimmer EMG3	Health care; recognition of high vs. low emotional valence and arousal by non-invasive low-cost sensors; n=19; emotions induced via music videos (DEAP data); classification via machine learning	n.a.	For arousal GSR and for valence EEG qualifies best	+ Non-invasive sensors + Low-cost sensors + Suitable for in car use – Results have to further validated with a larger sample – Small sample size
Gotardi et al.	2018	Anxiety (low vs. high)	• Eye movements, fixations per AOI, dwell time, transition frequency • HR, BPM	• Eye tracker, head-mounted ASL H6 • RS800CX wristband from Polar	Road safety; simulator experiment (multi-lane highway); anxiety induced by i.a. peer pressure; self-assessment on STAI-S and HR for manipulation check; n=16	• Visual entropy • fixations transition • frequency • HR	• Sign. more random scanning, indicating poorer acquisition of visual information under high anxiety condition • Mean HR in high anxiety cond. 91.41 ± 2.62 BPM vs. 79.43 ± 2.04 BPM • Impact on driving behaviour not reported	– Small sample size – Theoretical concept of anxiety not explained
Guo et al.	2019	Anxiety	Eye movements	Eye tracker classes from Tobii	Road safety; simulator experiment; Anxiety induced by visual scenarios; self-assessment on Beck Anxiety Scale and Self-Rating Anxiety Scale; change in eye movements in data set detected with change-point model; n=36	Eye movements	Differences between anxious and calm episodes: • Mean fixation duration: 0.92s (0.66s when calm) • Mean visit duration: 9.99s (7.34s when calm) Impact on driving behaviour not reported	+ Authors suggest that method is applicable for detecting unexpected events in the road environment – Type of eye tracker only suitable for simulator
Halim & Rehan	2020	Distress (negative valence and high arousal in 2D	EEG	EEG cap from EMOTIV EPOC+	Road safety; simulator experiment; driving-induced stress; classification of EEG	Con-gruence of EEG pattern and	SVM classifier performs best to distinguish between rest and stress state	+ Test of more than one classifier

Authors	Year	Theoretical constructs	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
		arousal valence emotion model)			patterns (classifiers: SVM, CNN, random forest); self-report of emotion; n=86	self-reported emotion	• accuracy=97.95% • precision=89.23% • sensitivity=88.83% • specificity=94.92% Impact on driving behaviour not reported	– Only applicable in simulator – Proposed model not further validated yet
Hu et al.	2018	Anger (no conceptual reference provided)	• ECG, average HR, R-R SD • EEG sign.	MP150 from Biopack	Road safety; simulator experiment (urban scenario); anger induced by video of unfair incident, manipulation check via verbal self-report; 4 anger levels, n=12 young drivers	ECG and EEG parameters	• Sign. differences between four anger groups regarding physiological measures • Impact on driving behaviour not reported	– Small sample size – Theoretical concept of emotions, anger not explained
Katsis et al.	2008	• Stress (high, low) • Disappointment • Euphoria	• EMG • ECG • EDA • Respiration	Multisensory wearable (balaclava, not commercially available)	Road safety, simulator experiment; test of wearable system for car-racing drivers; classification of emotions via SVM and ANFIS; n=10 male car-racing drivers	n.a.	The overall classification rates achieved by using tenfold cross validation are: • SVM: 79.3% • ANIFS: 76.7%	+ Non-intrusive device (in balaclava and around thorax), good results – Small sample size – Transferability to average drivers not clear (placement of the sensors)
Kolli et al.	2011	• Anger • Disgust • Fear • Happiness • Sadness • Surprise	Body temperature (thermal imaging)	Infrared thermal camera 'PathFindIR' from FLIR systems	Road safety; face recognition algorithms and emotion classifiers for ADAS; database with n=35; test of three different algorithms	n.a.	• Morphological operation-based algorithms with best performance • Colour-based detection with inconsistency, region growing with good result (thresholds difficult to set) • Happiness, sadness and disgust classified best • Fear could not be classified	– Algorithm performance for emotions relevant for driving not satisfying (fear and anger)
Lafont et al.	2018	Anger (sympathetic dominance hypothesis: HR, pupil size and long-term variability indexes of HRV increased and short-	• ECG, HRV • Eye measures: sample entropy, saccadic amplitude, pupil	• BIOPAC Systems • Mobile Eye-tracking glasses from Tobii	Road safety; simulator experiment (urban environment); anger induced via film clips; neutral, slight anger and strong anger group;	• Detection of VRU • Scanning strategy	• Difference in visual scanning strategies for three conditions • No effect on VRU detection was reported	+ Very detailed documentation of operationalization, design and procedure

Authors	Year	Theoretical constructs	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
		term indexes decreased)	size (>100 vs. <100)		manipulation check via self-assessment on Emotional Wheel; n=45 HRV: LF/HF-ratio: 0.04-0.15Hz indicating sympathetic activity; 0.15-0.4Hz indicating parasympathetic activity		• Results are inconsistent with previous research as well as with the sympathetic dominance hypothesis (although manipulation seemed successful)	– Inconsistency with previous results as unresolved issue
Landowska	2014	Affective state (not further specified)	• EDA • EMG • Respiration • EEG-Z sign. • Blood Volume (HR) • Temperature	FlexComp from Thought Technology, Canada	Human-computer studies; evaluation of sensors located on forearm and for sensitivity to movements during computer work, controlled experiment; n=31	Subjective assessment of intrusion	• SC sensors can be located on forearm: less intrusive than on finger, less sensitive to movement • Baseline personal average to be considered instead of overall mean	+ Sensor locations can be set on non-intrusive locations on the body – Transferability to driving not clear
Lee et al.	2018	'Aggressive driving emotions' (no further specification)	• Facial expression (facial image analysis of eyes, mouth) • Body temperature (thermal image of forehead, cheeks)	Near-infrared light thermal camera Tau2 from FLIR (NIR band-pass filter attached)	Road safety; simulator experiment, n=15; aggressive driving mode established by causing participants to make mistakes or introducing aggressive competitor in racing video game	n.a.	CNN classifier showed accuracy of 99.96% (combined scores from near-infrared camera and thermal camera)	+ High classification accuracy of CNN classifier + Open source image database, trained CNN – No manipulation check – Robustness not tested – Small sample size
Lisetti & Nasoz	2005	• Frustration, anger • Panic, fear • Boredom, sleepiness	• EDA • ECG, BPM • Body temperature	BodyMedia SenseWear Armband Polar chest strap	Road safety; simulator experiment; n= 41, emotions induced via scenarios; test of 3 algorithms (Marquardt-Backpropagation, k-Nearest Neighbor, Resilient Backpropagation)	n.a.	Accuracy of classifying the three emotions: • KNN 66.3%, • MBP 76.7% • RBP 91.9%	+ RBP showed good performance for classifying emotions + Controlled environment facilitates comparisons between participants
Lotz et al.	2018	• Neutral • Positive emotions (summarized) • Frustration, anger • Anxiety, fear (circumplex model of emotions)	• Speech recognition signals • EDA • ECG	• Dashboard (Shure VP 82 shotgun), headset mic. (Sennheiser HSP-4 EW-3) • Heally from Spacebit: EDA finger sensor, 3-lead	HMI; test track driving experiment (residential area); emotions induced by Wizard-of-Oz technique (alleged encounter of autonomous vehicle), reinforced by staff; manipulation check via Geneva Emotional Wheel,	n.a.	Consistency of two annotation approaches shown	+ On-road experiment (not simulator) – Time for equipping participants: 20 min. – No results on EDA and ECG reported

Authors	Year	Theoretical constructs	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
				ECG wearable and electrodes on chest	Self-Assessment Manikin; data annotated; n=30			
Ooi et al.	2016	• Stress • Anger	EDA	BioRadio 150 by Great Lakes NeuroTechnologies	Road safety; simulator experiment, n=20; Power spectral density (PSD) used to analyse EDA; SVM classifier; 10-fold cross-validation	EDA (PSD)	• Significant differences between neutral-stress and neutral-anger by EDA measurements • Classification accuracy 85%	+ Non-intrusive (system behind the wheel) – Stress and anger distinguished by EDA signals with accuracy of only 70%
Paschalidis et al.	2018	Stress	• EDA • HR	Empathica E4 wristband	Road safety; simulator experiment (urban environment); stress induced by time pressure; n=41	Gap-acceptance behaviour	Increased stress levels significantly increased probability of accepting a gap	– No manipulation check – No details on EDA, HR outcomes
Rebolledo-Mendez et al.	2014	Stress	• EDA • Neural data	• Affectiva´s Q sensor (EDA) • NeuroSky's MindWave	Road safety; feasibility of detecting emotional state in field experiment, n=24; urban vs. highway; drivers classified own emotions (concentrated, tension, tired, relaxed)	n.a.	• SCL correlated sign. with self-reports, preliminary model of a driver state recognition modul for OBU. • Middleware architecture to detect emotions proposed (communicated via OBU)	+ Architecture of measuring emotions while driving and integrating information in the OBU – Reliability of self-reports
Rodrigues et al.	2015	Stress	ECG, HRV	Vital Jacket	Occupational health, on-road experiment; testing mobile sensing to detect physiological and psychological stress; n=36 bus drivers; drivers geofenced stress events by pushing button	n.a.	• HRV correlates with stress events • Less cardiac response of experienced bus drivers • Stress map for bus drivers derived from the data • effect on driving behaviour not investigated	+ Non-intrusive, easy to use system (t-shirt) + Tested in real world
Villarejo et al.	2012	Stress	EDA, GSR	No information provided	Medicine, n=16; experiment including different tests requiring certain degree of effort	n.a.	• GSR device differs between relaxed and situation requiring effort with accuracy of 90,97% • no differentiation between being stressed and making an effort	– Pre-mature device, does not perform sufficiently – Small sample size
Zhang et al.	2019, 2018	• Fear (represented as	Emotional power:	• Infrared thermography	Road safety; simulator experiment (urban	Forehead temperature	• Reduction of forehead temperature after threat	+ Indication that forehead temperature

Authors	Year	Theoretical constructs	Indicators	Technical equipment	Context & Design	Outcome variables	Results	Conclusion (for i-DREAMS)
		dimension of emotional power in Fontaine's four-dim. emotional space) • Emotional arousal	• Body temperature of forehead, nose, finger (facial image analysis) Emotional arousal: • ECG, average IBI • EDA, SCL	camera 'PI640' from Optris • Finger sensor and ECG sensors (chest) 'Heally' from SpaceBit	scenarios); n=18; fear induced by threat and challenge events in driving; within-subject design; manipulation check via self-report on Positive and Negative Affect Schedule and Self-Assessment Manikin		events (M= -0.02°), SD=0.08°C); • Forehead temperature did not correlate sign. with arousal (SC and IBI) • SCL (indicator of arousal) did not differentiate between fear vs. no-fear • Impact on driving behaviour not reported	can be used to measure + Infrared camera non-intrusive, contact-free – Small sample size – Accuracy of facial temp. measure might be sensitive to ambient temp.
Zimasa et al.	2019	• Happiness (positive+high arousal) • Sadness (negative valence/low arousal) • Anger (negative valence/high arousal)	Arousal: • EDA • HR, BPM Valence: • Self-reported	• 'E4' wristband from Empatica • Eye tracker 'faceLAB 5' from Seeing Machines	Road safety; simulator experiment; n=40; emotions induced via music and mental imagery, manipulation check with self-reports on mood assessment grid (valence) and physiology (arousal); within-subject design	• Attention (mean duration and spread of fixations) • Car following behaviour	• Sign. increase in fixation duration while sad • Decrease in fixation duration in neutral mood, compared to happy, angry (high arousal), indicates improvement in attentional shift, inform. processing • Neutral mood: wider visual field compared to high arousal • No clear results on time headway	– No details on EDA and HR parameters provided

Annex F: Technology review on measuring the driver's mental state

Table 30: Overview of devices and technical equipment used in the reviewed studies on measuring attention, fatigue and sleepiness and emotional states and stress

Product/company name and contact	Constructs measured	Equipment, measurement method	Intrusive-ness	Indicators	Relevant information for i-DREAMS Overall assessment and considerations
Seeing machines https://www.seeingmachines.com/	• Fatigue, sleepiness • Distraction	• In-cab sensor • Cameras facing forward and driver	Contact free	Face and eye tracking indicators	• Used in truck fleets • Planned to be used on UK rail network (https://www.smartrailworld.com/driver-anti-sleep-seeing-machines-device-uk-rail-safety-regulator) • Already implemented in trams in Croydon following tram crash • Can be designed to issue alerts • Established product • Would need installation and training on use/analysis
Vigo https://www.wearvigo.com/	• Drowsiness, alertness • Distraction	• Headset • Head, eyes tracking	Medium	Vision based: 20+ eye /head tracking parameter	• Real time performance monitoring • Can be implemented across fleets • Established product • Would need installation and training on use/analysis • Requires driver to wear headset • Minimal information available
Optalert https://www.optalert.com	• Drowsiness, alertness • Distraction	Video cameras on dashboard, steering wheel	Contact free	Vision based: eye tracking, facial features, amplitude and velocity ratio of blinks	• Issues early warnings • Driver does not need to wear or do anything • Established product • Would need installation and training on use/analysis • Licensable software
Cardio wheel https://www.cardio-id.com/cardiowheel	• Sleepiness, drowsiness • Attention • Stress	Sensors on steering wheel	Low	ECG, HRV	• Contact of both hands to steering wheel required • Requires custom steering wheel • Dashboard for fleets of vehicles • Can be integrated with certain third-party systems • Un-intrusive
Smart eye https://smarteye.se/	• Fatigue, sleepiness • Attention, alertness	Eye tracking cameras on dashboard	Contact free	Vision based: eye, face and head tracking	• Non-intrusive • Driver does not need to wear or do anything • Developed for automotive industry • Established product • Would need installation and training on use/analysis
Phasya https://www.phasya.com/en	• Drowsiness • Stress	Any equipment used to measure aspects of driver state	Dependent on equipment used	Vision based, eye tracking: PERCLOS, facial features,	• Software • Range of modules available

Product/company name and contact	Constructs measured	Equipment, measurement method	Intrusive-ness	Indicators	Relevant information for i-DREAMS Overall assessment and considerations
	• Distraction, mind wandering • Cognitive load			HR, blink duration, pupil diameter, facial images	• Offers test and validation of physiological and cognitive state monitoring systems • Software can be used within rail and automotive industry.
Texas Instruments Biometric Steering Wheel http://www.ti.com/tool/TIDA-00292	• 'Driver state'	Sensors on steering wheel, measuring pulse, respiration, HR	Low	ECG heart rate, pulse rate, respiration rate	• Contact of both hands to steering wheel required • Non-intrusive • Unsure of validation
Veoneer LIV – Driver monitoring system https://www.veoneer.com/sites/default/files/Veoneer_Meet%20LIV_Aug29.pdf	• Fatigue • Attention • Environment	Uses external and internal sensors, combined with AI	Contact free		• Is not a single system or component but an equipped car
BioRadio by Great Lakes NeuroTechnologies https://glneurotech.com/bioradio/	• Emotions (anger) • Stress • Attention • Fatigue	Electrodes attached to finger	Medium	EDA	• Used for clinical research • use only with assistance of project team
BIOPAC Systems, BioNomadix® Logger https://www.biopac.com/product-category/research/bionomadix-wireless-physiology/	• Emotions (anger, fear) • Attention • Fatigue	Electrodes attached to the body	Medium	ECG EDA Temperature	• Electrodes are placed once, then connected to the wearable recording system, • can be used with other devices (e.g. eye tracking)
Empatica E4 Wristband https://www.empatica.com/en-eu/research/e4/	• Fatigue • Attention • Emotions (arousal)	Wristband with sensors	Low	EDA Sensor	• Seems easy to use • Battery runs 48h • additional equipment: 3-axis Accelometor to capture motion-based activity, event-mark button
FlexComp from Thought Technology	• Emotions (affective state)	Electrodes attached to fingers	Medium	HRV, EDA	• Mainly used for biofeedback

Product/company name and contact	Constructs measured	Equipment, measurement method	Intrusive-ness	Indicators	Relevant information for i-DREAMS Overall assessment and considerations
http://thoughttechnology.com/index.php/complete-systems.html					
Shimmer 3, including PPG ear clip http://www.shimmersensing.com/ products/shimmer3-wireless-gsr-sensor	• Emotions (arousal, valence, anger) • Stress • Fatigue, • Attention	Electrodes positioned on chest or arms	Medium	ECG	• Electrodes are placed on participant fixed for each trial • used in laboratory research • can be used with other devices
Shimmer 3 GSR unit, EMG3 http://www.shimmersensing.com/	• Emotions (arousal, valence, anger) • Stress • Fatigue • Attention	Electrodes attached to two fingers	Medium	EMG	• Electrodes are placed on participant each trial • used in laboratory research • Cannot be used together with Shimmer 3, which measures ECG
Vital jacket www.vitaljacket.com	• Stress	T-shirt with implantable body sensor networks	Low	ECD, HR	• Seems easy to use • records physiological measures 72h on SD card
Zephyr BioModule https://www.zephyranywhere.com	• Stress (arousal)	T-Shirt or as wristband	Low	ECG	• Is used in sports for training feedback and stress measurements • Customized system can be acquired with different components.
Eye-tracking glasses	• Fatigue, sleepiness • Attention • Emotions (anxiety)	Eye-tracking light classes	Medium	Vision based: mean fixation time	• Calibration of eye tracker might be time-consuming • Not suitable for on-road trials
Near-infrared light thermal camera (e.g. Tau2 or PI640 from Optris)	• Fatigue • Emotions • Stress		Contact free	Thermal imaging	• Non-intrusive