Delivery Report for

MeBeSafe

Measures for behaving safely in traffic

Deliverable Title	Driver profiles and situations
Deliverable	D4.1
WP	WP4 Driver Coaching
Task	Task 4.1 Driver profiling

This project (MeBeSafe) has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 723430.

Disclaimer

Copyright

Grant Agreement No.	723430
Project Start Date	01/05/2017
Project End Date	31/10/2020
Duration of the Project	42 months
Deliverable Number	D4.1
Deliverable Leader	Nicole van Nes, SWOV
Dissemination Level (Confidentiality)	PU: Public
Nature	Report
Status	Final
Due Date	M9 (Jan 2018)
Deliverable Leader/ Responsible Organization	Nicole van Nes, SWOV
Contributing Author(s)	Peter Adams, SHELL Angela Baker, SHELL Bram Bakker, Cygnify Saskia de Craen, SWOV/SHELL Reinier Jansen, SWOV Teri Lillington, SHELL Peter Moertl, Virtual Vehicle Celina Mons, SWOV Norah Neuhuber, Virtual Vehicle Nicole van Nes, SWOV Ingrid van Schagen, SWOV Simone Wesseling, SWOV
Reviewer(s)	MariAnne Karlsson, Chalmers/ SAFER
Formal review and upload to the EU by Coordinator	Institute for Automotive Engineering - RWTH Aachen University (ika)

Abstract
One of the objectives in MeBeSafe is the coaching of drivers, in particular heavy goods vehicle (HGV) drivers, on their driving behaviour. Risky driving behaviour can lead to crashes but by coaching drivers on their driving behaviour we can reduce risky driving behaviour, therefore reducing crashes and as a result increase traffic safety. The deliverable serves as a progress report. The objective was to investigate what data is needed for coaching of heavy goods vehicle drivers, how we can collect these data, what variables are relevant for driver profiling and how we can use these variables for driver profiling. With regards to technology, our recommendation is to collect data on driving behaviour and driving context with a mobile phone, augmented with inward- and outward-facing cameras where possible. In terms of driver profiling we aimed to capture “the tendency to behave a certain way *in a certain situation or context* ” and distinguish meaningfully between different situations or contexts in which a particular type of behaviour occurs. Therefore driver profiles were developed using driving behaviour variables measured by telematics, including context information. “The Traffic Safety Wheel” was developed, a representation of driver profiles where we can compare driver behaviour with fleet behaviour across varying driving contexts. Based on the results further decisions can be made on how to proceed in this MeBeSafe project.

Version	Date	Comment
1	1-11-2017	Template and structure
2	4-12-2017	Draft 1
3	15-12-2017	Draft 2
4	21-12-2017	Draft 3 – Send to reviewer for 1st review
5	12-1-2018	Review comments received on draft 3.
6	17-1-2018	Modifications made based on feedback. Draft 4 send to reviewer for 2nd review
7	23-1-2018	Review comments received on draft 4
8	25-1-2018	Finalised the deliverable and submitted to coordinator

table of document history

Table of Contents

List of Figures

List of Tables

Acronyms

Acronym	Explanation
DBQ	Driver behaviour questionnaire
HGV	Heavy goods vehicle
IVMS	In Vehicle Monitoring System
FOT	Field Operational Trial
KPI	Key performance indicator
MeBeSafe	Measures for Behaving Safely in Traffic
ND	Naturalistic driving
UDRIVE	European naturalistic Driving and Riding for Infrastructure & Vehicle safety and Environment
WP	Work Package

Glossary

Term	Definition/explanation
Coaching	A collaborative solution-focused, results-orientated systematic process, used with normal, non-clinical populations, in which a coach facilitates the enhancement of work performance and the self-directed learning and personal growth of a coachee.
Competences	Dispositions that allow an individual to master variable situations successfully and responsibly and can be seen as fundamentals for learning. This includes motivational and violational aspects.
Concurrency	The amount of overlap between different KPI variables (driving behaviour variables).
Driver profiling	Profiles based on driving behaviour in order to distinguish between different styles of driving (risky versus safe) or between driving behaviours in different situations.
KPI variables	The variables selected for measuring driving behaviour. The Key Performance Indicator (KPI) variables in the current study are: harsh braking, speeding, distraction, drowsiness, close following, harsh cornering, lane departure and possibly fuel consumption.
Naturalistic driving	A research method wherein every day trips by drivers are recorded by unobtrusive data acquisition systems with the aim of providing insights into actual driver behaviour.
Thresholds	A cut-off value used to determine whether or not a particular drivers' behaviour is safe or unsafe.
Traffic safety wheel	A visualization of driver profiles based on driving behaviour and context.

Executive Summary

Objective

One of the objectives of the MeBeSafe project is the coaching of drivers, in particular heavy goods vehicle (HGV) drivers, in order to improve their driving behaviour. This is part of Work Package 4. The objective in this deliverable is to investigate what data is needed for the coaching of heavy goods vehicle drivers, how we can collect these data, what variables are relevant for driver profiling and how we can use these variables for driver profiling. The aim of driver profiling is to create profiles based on driving behaviour in order to distinguish between different styles of driving (risky versus safe) or between driving behaviours in different situations.

This deliverable serves as a progress report and the results can be used as input for the design of coaching schemes and the app that will measure and give feedback on driving behaviour (Task 4.3) and eventually for the field evaluation test in WP5.

Measuring risky driving behaviour

Risky driving behaviour can lead to crashes. By coaching drivers on their driving behaviour the aim is to reduce risky driving behaviour, therefore reducing crashes and consequently increasing traffic safety. The following behaviours related to risky driving, or so called Key Performance Indicator (KPI) variables, are focused on for coaching in this study: harsh braking, harsh cornering, close following, lane deviations, drowsiness/fatigue, distraction, speeding, and optionally fuel consumption. Research has indicated that these behaviours are related to traffic safety (Hanowski, Perez, & Dingus, 2005; FMCSA, 2006; Olsen, Hanowski, Hickman, & Bocanegra, 2009; Saberg, Selpi, Piccinini, & Engström 2015; Dingus et al., 2016; SWOV, 2016) and are therefore relevant for coaching. The technology that was considered for measuring driving behaviour were ***IVMS*** (In Vehicle Monitoring System) and ***mobile phone***. Results show that roughly the same KPI variables can be measured by these devices. Although IVMS can measure certain variables more precisely than the mobile phone, a great disadvantage of IVMS is the fact that not all vehicles have the same system installed. This is a problem, because the software that is used to read the data from the IVMS needs to be adapted for every system, making it a costly and time consuming endeavour. Also, data like speed limits, weather conditions, and type of road are relatively easily combined with data collected by a mobile phone, but not as easily with an IVMS system. Therefore, the mobile phone is a better option to use for collecting data on driving behaviour in the current MeBeSafe project.

Measuring driving behaviour on itself is often not sufficient, as behaviour should be put into context whenever possible. The characteristics of the driving environment can influence how a driver behaves and it is therefore of importance to also take driving context into account when looking at driving behaviour. For example, in an urban environment drivers will most likely have to brake harshly more often compared to when driving on a highway. Drivers that drive more in urban environments will therefore show more harsh braking behaviour compared to drivers that drive more often in rural environments. Consequently, a difference between drivers could be the result of differences in the environment they are driving in, and not so much their actual driving behaviour. It is therefore important to consider the characteristics of the trips and situations the drivers are in, to generate a fair representation of driving behaviour. ***Cameras*** can capture additional driving behaviour as well as information on the environment. Cameras would therefore be of great value for collecting additional information for coaching. Outward-facing cameras would give more information on the conditions a driver is in, for example traffic density can be measured and videos of relevant situations can be saved and shown to a driver. Inward-facing cameras could provide information on distraction and drowsiness, both important factors related to traffic safety. Nevertheless, it is of importance to realise that drivers might have issues with having cameras installed in their vehicle due to privacy, and in some countries outward-facing cameras are not legally allowed. These are still concerns that need to be dealt with while investigating the possibilities of collecting data.

To further understand why a driver is behaving in a certain way we need to look at the characteristics of a driver. The behaviour of a driver can be influenced by characteristics like personality, attitude, age, experience and competences. These factors can be measured by ***questionnaires***. However, several studies point out that the validity of questionnaires measuring personality and driving behaviour is low and that more research is needed before questionnaire data can be used for driver profiling. We therefore refrain from using questionnaires for driver profiling based on personality, competences and attitude. Nevertheless, gaining additional insights into why a driver behaves in a certain manner by looking into competences, personality and attitudes, is difficult to do so other than by questionnaires, interviews or tests (simulator, hazard perception tests).

Driver profiling

In terms of driver profiling we aimed to capture "the tendency to behave a certain way *in a certain situation or context* ", and distinguish meaningfully between different situations or contexts in which a particular type of behaviour occurs. Therefore driver profiles were developed using KPI variables measured by telematics, including driving context information.

The "Traffic Safety Wheel" was developed, a representation of driver profiles wherein we can compare driver behaviour with fleet behaviour across varying driving contexts. As shown in *Figure 0.1* below, we can make a distinction between driving behaviour of a driver compared to average fleet behaviour; and we can make a distinction between driver behaviour in different contexts, like highway compared to city. The driving behaviour variables we measure are shown on the axes. Note that these driver profile representations stays close to the KPI variables we wish to optimize. The shape that results from this mapping gives an overview of how well a driver scores on each of the KPI variables. A larger surface in the safety wheel implies that a driver shows more aberrant behaviour compared to a smaller surface in a safety wheel.

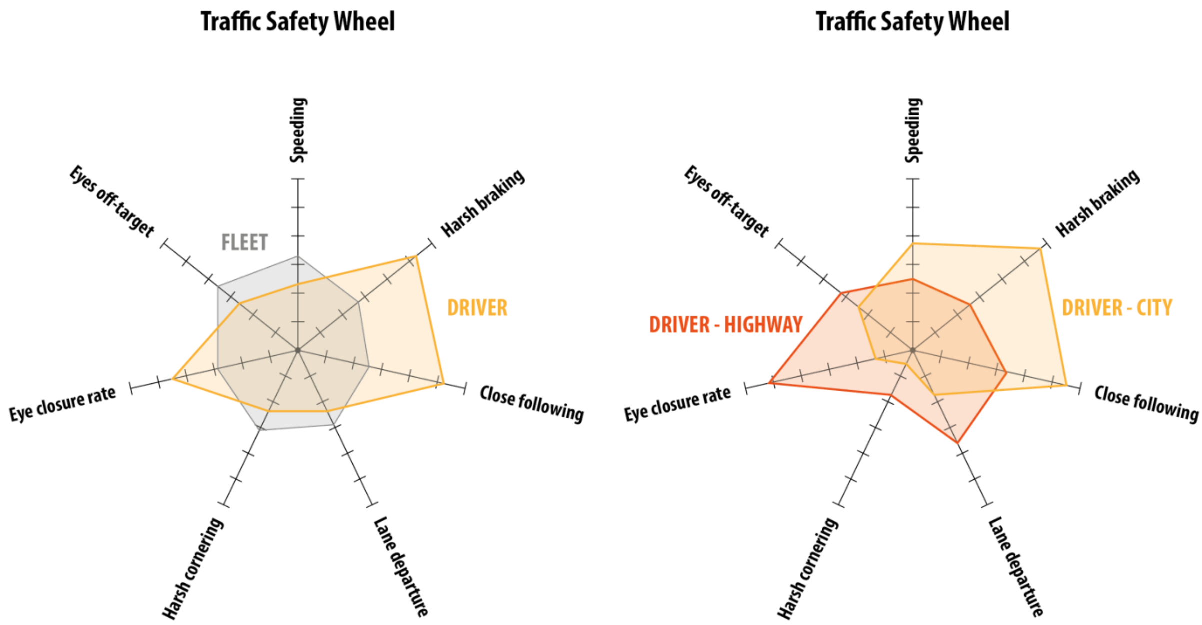


Figure 0.1"The Traffic Safety Wheel". A representation of driver profiles wherein we can compare driver behaviour with fleet behaviour across varying driving contexts showing driving behaviour variables on the axes.

The safety wheel as presented serves as a foundation for driver profiles which could be used for coaching, but it is not yet intended to use directly as visualization for drivers - although it might. For the latter purpose, the safety wheel should first be tested by HGV drivers and usability experts on visual appeal and ease of use, possibly followed by a redesign. Furthermore, in the app used on the mobile phone and for coaching there could be an additional focus on emphasizing positive driving behaviour, next to risky behaviour, the former of which is not most naturally represented in the traffic safety wheel. But could be visualised for example by using a "positive" green colour as background colour, when variable values are closer to the centre of the traffic safety wheel more of the green colour is shown. Another option would be to use a kind of "bull's eye" visualization. Further development of the Traffic Safety Wheel would be needed if it will be incorporated in the project.

Based on the results described in this deliverable further decisions can be made on what driver profiles can be used and how data should be collected for Work Package 4.

1 Introduction

1.1 MeBeSafe

The aim of the MeBeSafe project is to develop, implement and validate measures that direct road users towards safer behaviour in common traffic situations. MeBeSafe is looking to do this by changing habitual traffic behaviour using 'nudging' and coaching, with the aim of improving driving behaviour. Nudging is a technique that subconsciously stimulates drivers to drive safer, while with coaching, drivers are given feedback on their driving behaviour by a coach in order to enhance driving performance and learn about their own driving behaviour. The work in this deliverable is focussed on coaching, in particular heavy goods vehicle (HGV) drivers.

MeBeSafe is organised in altogether six work packages (WPs), as shown in *Figure 1.1*. The coaching of drivers on their driving behaviour is part of Work Package 4.

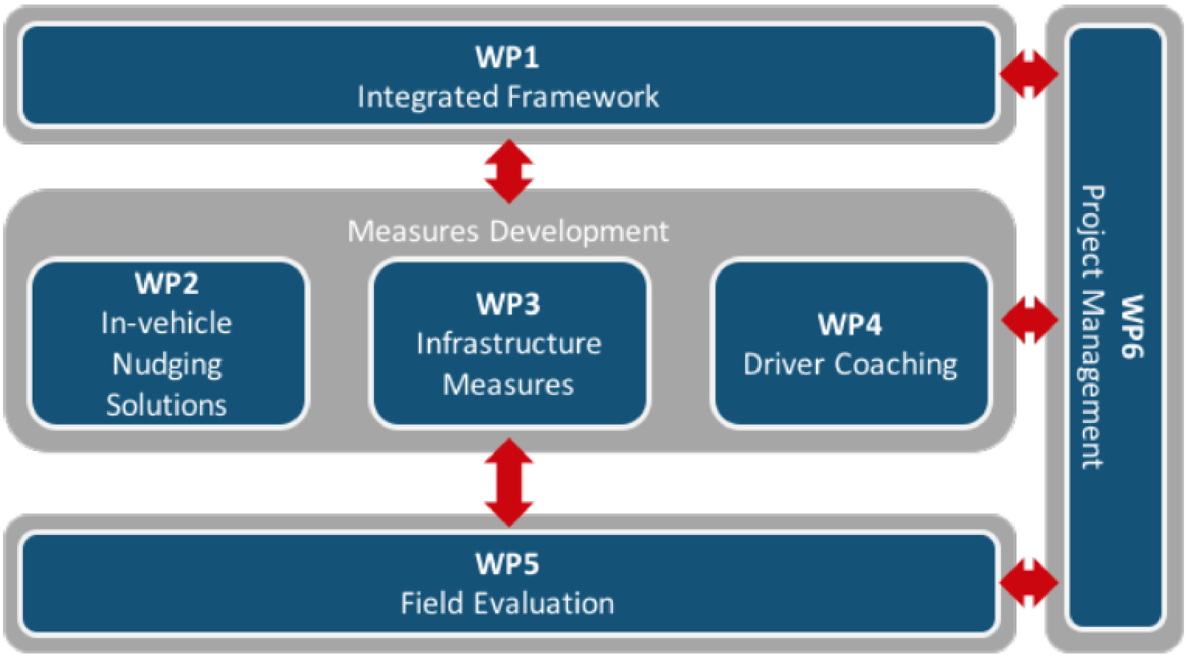


Figure 1.1 Work packages in MeBeSafe.

Work Package 4 focuses on the development of a driver coaching scheme, a supporting coaching app that can be used to coach HGV drivers and an evaluation of the coaching scheme. HGV drivers that will be approached to participate in the study are working for hauliers contracted by Shell. The work package consists of the following 4 tasks:

1) Profiling drivers for whom, and situations wherein the coaching scheme could be used and tested (task 4.1);
2) Defining the methodology for the evaluation of the scheme (task 4.2);
3) Development of the coaching scheme and coaching app and a small-scaled pilot (task 4.3);
4) Evaluation of the coaching schemes and app based on the pilot results (task 4.4).

The end result of WP4 consists of a coaching scheme and a coaching app that can be used for a large-scale field evaluation in WP5. In this deliverable the work in task 4.1 is described.

1.2 Objective of this deliverable

Risky driving behaviour can create dangerous situations in traffic and these dangerous situations can then result into crashes. Preventing unsafe behaviour can therefore lead to a decline of the number of deaths and injured in traffic. In 2014, 25,939 fatal road accidents took place in Europe. 15% of these fatal accidents were HGV-related (Volvo Trucks, 2017). Furthermore, non-fatal accidents can lead to severe injuries, especially since HGV are large vehicles with a gross weight above 3.5 tonnes. Giving drivers insights into their driving behaviour by coaching can result in safer driving behaviour (Karlsson et al., 2017). To coach the drivers, coaching schemes need to be developed which describe how drivers will be coached on what behaviours, with which frequency, etc. In addition a “coaching app” will be developed that gives drivers feedback on their driving behaviour. The objective in this deliverable is to investigate what data is needed for coaching of heavy good vehicle drivers, how we can measure these data, what variables are relevant for driver profiling and how we can use these variables for driver profiling. The aim of driver profiling is to create profiles based on driving behaviour in order to distinguish between different styles of driving (risky versus safe) or between driving behaviours in different situations. The results serve as input for task 4.3: the development of coaching schemes and suggestions made in this deliverable will need to be further developed.

In the following paragraphs of the introduction the underlying approach for this deliverable is described. Starting with discussing the risky driving behaviours that will be the focus for coaching.

1.3 Risky driving behaviour by HGV drivers

In MeBeSafe deliverable 1.1 (Karlsson et al., 2017) it was concluded that HGV drivers are experienced drivers that generally know how to drive safely and efficiently, though they do not always use their safe driving skills to its full extent. Increasing drivers' motivation to drive safer is therefore mentioned as an option for coaching. This can be done by giving insights into risky driving behaviours that could be improved. The following variables related to risky driving are of focus for coaching in this study: harsh braking, harsh cornering, close following, lane deviations, drowsiness/fatigue, distraction, and speeding. Research has indicated that these behaviours are related to traffic safety (Hanowski, Perez, & Dingus, 2005; FMCSA, 2006; Olsen, Hanowski, Hickman, & Bocanegra, 2009; Saberg, Selpi, Piccinini, & Engström 2015; Dingus et al., 2016; SWOV, 2016) and are therefore relevant for coaching. We call these driver behaviour variables Key Performance Indicator (KPI) variables throughout the deliverable. In addition, fuel consumption could also be monitored, since drivers and hauliers could be interested in the financial and environmental aspects of reducing fuel consumption. This could give an added motivation for drivers and hauliers to join the project.

Driving behaviour can be influenced by several factors, amongst others, the design of the road system, road user competences or skills, road user states and personality and attitude (Karlsson et al., 2017). As suggested in deliverable 1.1 the driving context might not always support HGV drivers to drive safely. It is therefore of importance when looking at driving behaviour to also take driver characteristics (such as age, attitude and personality), situational and environmental factors into account. In this deliverable we therefore examine how driving behaviour, driver characteristics and context can be measured.

1.4 Measuring driving behaviour, driver characteristics and driving context

To be able to coach drivers on their driving behaviour, we need to know how they generally behave. For this purpose data needs to be collected on their driving behaviour. These data can then be processed into an overview of a driver's behaviour. This overview can for example show how often a driver has been speeding or how often the driver has been braking harshly during a particular trip. The overview of driving behaviour then serves as the basis for feedback to drivers and which, in turn, can generate insight to the drivers about how risky they are behaving in traffic. Coaches can use the overview to give concrete feedback to drivers, which can help with gaining insights.

There are several ways to measure driving behaviour. In this deliverable we examine different options to collect data on driving behaviour, driver characteristics and context of the driving situation, which could be used as a basis for coaching. The aim is to give an overview of the possibilities. One way to measure driving behaviour is by using telematics; variables that can be measured and registered by a monitoring device located in the vehicle. This can be an internal in vehicle monitoring system (IVMS), but can also be an app on a mobile phone. Variables like GPS position, acceleration and speed can be measured. Telematics can be enriched by video data and map data. Video and map data can provide information that can be used to identify the influence of the context on driving behaviour. For example, inward-facing cameras in a truck could capture drowsiness, distraction and inattention, while outward-facing cameras on a truck could capture the infrastructure and other road users. Map data could provide information on the type of road (rural, urban, highway), presence of intersections and speed limit. Another way to collect information about driving behaviour, but also on driver and/or organisation characteristics and demographics is by means of questionnaires or interviews. These data can give further insights into *why* a driver behaves in a certain manner. For example by asking questions about the drivers' attitude towards safe or risky driving. In addition, competences could be measured by questionnaires, possibly in combination with a simulator or other tests (a hazard perception test for example).

Besides examining how driving behaviour variables, driver characteristics and driving context factors should be measured, we will also look at using driver profiling to see if we can distinguish between risky driving behaviour and safe driving behaviour and distinguish between different situations.

1.5 Driver profiling

When we measure variables by telematics, questionnaires or video and map data we can get more or less detailed insight into driving behaviour of an individual driver. With coaching we can give tailored feedback to drivers on their behaviour. We can give positive feedback when they are driving safely (or safer than others) and we can give feedback to improve their risky driving behaviour. This can be done through an app, by face-to-face coaching or a combination of both. The intended result of coaching is that their driving behaviour improves or that they are triggered to continue to drive safely. In this deliverable we examine how we can use driver profiling to further tailor the coaching efforts.

With driver profiling we aim to design driver profiles based on driver behaviour, the environment, demographics, competences or other driver characteristics in order to distinguish between different styles of driving (risky versus safe) or between driving behaviours in different situations. This can be done in several ways.

One way is to distinguish sub groups or grade drivers in terms of how risky they drive using questionnaires, with the aim of distinguishing between risky and safe drivers.

These sub groups of drivers could be coached differently and these different groups could respond differently to coaching. We can choose for example to coach risky driving behaviour more extensively than drivers with safe driving behaviour. Because different sub groups could respond differently to coaching it is valuable to look at this while evaluating the coaching schemes. When drivers are already driving safely most of the time, coaching might not have any effect on them, while drivers that show more risky driving behaviour might benefit more from coaching. If we are able to make a distinction between drivers using driver profiling we can look at what effect coaching has on the different sub groups. It should be noted that these are options to be considered for how we can use driver profiling for coaching, but still needs to be investigated further.

Another way to use driver profiling would be to look at driving behaviour in different contexts, like urban and rural roads and design driver profiles to distinguish between situations. To be able to distinguish between drivers and driving behaviour we need to look at the variables that can be measured and are relevant for distinguishing between different situations categories. By categorizing driving behaviour for different contexts we can visualise how driving behaviour differs between types of context. Or we could visualise driving behaviour of a specific driver in relation to average driving behaviour of a population.

1.6 Report structure and contribution by partners

In this deliverable we focus on:

- How to collect data on driving behaviour, driver characteristics and driving context and;
- Driver profiling based on driving behaviour, driver characteristics and the driving context or situation.

We work towards a model of representing driving behaviour and driver profiles with a clear relationship to the KPI variables identified earlier, which can be used to compare individual drivers to others, to distinguish meaningfully between driving in different contexts (situations), *and* which can be visualized in such a way to allow for meaningful interpretation.

The chapters are written by different authors. The structure of the deliverable is as follows: (see *Table 1.1* for an overview):

- *Chapter 2* provides input on how to collect data on driving behaviour focussing on capturing driving behaviour with IVMS and with a mobile phone.
- *Chapter 3* explores whether to include driving context in driver profiling and what thresholds could be used for braking and speeding events. To do so, naturalistic driving data has been used to study actual driving behaviour. Furthermore, the relation of data collected by questionnaires with actual driving behaviour is investigated.
- *Chapter 4* focusses on how we can include information on the environment to generate a better understanding of the context that influences (risky) driving behaviour. This chapter provides input on how to collect data on driving behaviour and driving context using cameras installed in the vehicle and software developed by the company Cygnify.
- *Chapter 5* describes the added value of knowledge on driver competences for coaching and how to measure these. This chapter provides input on how to collect data on driver competences and suggests a driver competence model.
- *Chapter 6* is a literature review focussed on the possibilities of driver profiling based on data collected with questionnaires. Specifically, driving behaviour, demographics, driving experience, personality and attitude are looked at.
- The conclusion in *Chapter 7* summarises the main findings on how we can collect data on driving behaviour, driver characteristics and context; which methods are advised to use for the current goal and how we can use this data to create driver profiles.

Chapter	Title	Main focus		Author
		Collecting data	Driver profiling	
Ch. 2	Measuring driving behaviour with IVMS and mobile phones	Driving behaviour and context		Saskia de Craen, Shell/SWOV
Ch. 3	Driver profiling based on Naturalistic Driving data	Driving behaviour	Driving characteristics and context	Reinier Jansen, SWOV
Ch. 4	Collecting data on driving behaviour and context based on automated and video-based situation analysis	Driving behaviour and context		Bram Bakker, Cygnify
Ch. 5	Measuring driver competences	Driver competences		Norah Neuhuber, Virtual Vehicle
Ch. 6	Driver profiling based on questionnaires		Driver characteristics	Simone Wesseling, SWOV

Table 1.1 Overview of the chapters, the focus of the chapter on either collecting data or driver profiling and their author

This deliverable can be seen as a progress report and different possibilities of collecting data and driver profiling have been explored. No definitive decisions are made in this deliverable yet. Further steps will be taken in the project to decide on how to proceed. The findings in this deliverable are a basis for these decisions.

2 Measuring driving behaviour with IVMS and mobile phones

MeBeSafe aims to change driver behaviour as means of reducing road related deaths. Technology is a mode by which such a change can be delivered, monitored and measured. This chapter describes the technology that already exists within the Shell fleet[1] that can be used to collect data for the MeBeSafe project[2]. It is focused on which data can be measured with 'In Vehicle Monitoring Systems' (IVMS; *Section 2.1*) and with a mobile phone (using a driving app; *Section 2.2*).

In the MeBeSafe project we need information to feed the driving app, as input for the coaching scheme and information for the evaluation of app and coaching scheme (i.e. Key Performance Indicators). As already indicated in the introduction, variables that are considered most relevant for the project are:

- (Unnecessary) harsh braking
- Speeding
- Distraction
- Close following
- Harsh cornering
- Lane departure warnings
- Fuel consumption

In addition this chapter will assess if existing technology can be used to collect contextual data:

- Location (route), date and time of a trip

[1] The definition Shell fleet includes vehciles owned and operated by Shell which are typically light vehicles and also the vehicles that are operated on our behalf by Contractors which are typically the heavy vehicles and buses. For MeBeSafe, and this chapter, we only consider heavy goods vehicles (HGV's)

[2] This information is based on existing expertise and experience with measuring driving behaviour within the Shell fleet.

- Characteristics of the roads (e.g. urban/rural roads)
- Conditions (weather, congestion, etc.)

The following questions will be answered:

1. Can IVMS and/or mobile phone data be used to calculate (some of) these variables? What are the pros and cons?
2. What are the advantages and disadvantages of each method?

2.1 IVMS data

Existing fleet and professional driver behaviour monitoring technology assists companies to understand and improve safety and cost performance outcomes. The technology, commonly referred to as telematics, is a combination of hardware that is installed within the vehicle (NB: there are a variety of providers and data stream processes/output and version changes/updates) and telecommunications (capabilities to share data in real time or at pre-determined intervals) that enable the data to be accessed and analysed via a web browser.

In the Shell fleet the telematics used for driver and vehicle monitoring is via 'In Vehicle Monitoring Systems' referred to as IVMS. However, not all fleet vehicles have IVMS installed, and of those that do some hauliers only activate a proportion of what is actually available from the vendors, while other hauliers use the data streams and actively further analyse the data offline. Some hauliers even map locations with relatively high harsh braking events to identify potentially dangerous locations that can be used to prepare drivers for a particular route or trip.

Furthermore, the feedback received by a driver may vary from a detailed monthly printed report that includes graphs of braking, speeding and accelerations against fuel consumption with information on how to improve their behaviour (all of which is supported by a face to face engagement with their supervisors), to very little formal feedback but rather an annual discussion with their line manager regarding overall behaviour for the year. Some businesses provide incentives whilst others do not. So within the Shell global business there is a large range of availability and use of IVMS by the different hauliers.

Data streams that can be typically captured from IVMS systems in the Shell fleet include:

- Working hours per shift
- Driving hours per shift
- Engine start and stop times
- Engine running time
- Engine idling time
- Engine revolutions
- Brakes On
- Brakes Off
- Unauthorised routes
- Number of trips
- Second by second rep/post-accident data

There are a number of different variables that can be obtained via IVMS and that are relevant for the MeBeSafe project. These are described in more detail in the next section.

2.1.1 IVMS variables

Braking & Harsh Braking
Braking the vehicle in a smooth and progressive manner demonstrates that the driver has good control of the vehicle and is anticipating the road ahead. Braking and harsh braking is recorded if it exceeds pre-defined limit.
IVMS Limitation: IVMS data cannot determine if the harsh brake was avoidable or not. However, there are two things that can be derived from analysis of harsh braking data:

1) Identify those drivers who harsh brake often, on many locations: that is a driver behavioural trait that needs to be addressed. 2) Identify those locations where all drivers harsh brake, that is an infrastructure issue. In this case opportunities for planning alternative routes should be considered and driver briefings should alert all drivers to the risks associated with these locations.
Facts: o Braking wastes fuel since the engine has used fuel to get vehicle to the desired speed o Harsh braking wastes even more fuel because it will usually require more gear changes to get vehicle back to the desired speed o Braking causes wear to the braking components which will affect maintenance costs o Harsh braking is dangerous to yourself and other road users, especially on wet roads o Use of the exhaust brake will save fuel and reduce wear on brake components

Speed
Speeding – either not complying with speed limits or having a too high speed for the current traffic situation - is dangerous. It increases the risk of a crash involving the truck and other road users. Additionally, speeding can have a negative effect on fuel economy due to aerodynamic drag.
IVMS Limitation: IVMS can measure speed of the vehicle at any given time, however it cannot determine if the speed was suitable for the driving conditions.

IVMS may be able to determine if the vehicle was traveling within the speed limits if the limits are incorporated into the GPS map of the area, or if not available within the GPS mapping system then within a speed defined geo fenced region.
Facts: o The higher the driving speed, the more fuel is consumed o A significant fuel saving can be seen when reducing speed o The higher the driving speed, the more time is required to stop the vehicle safely and the greater the momentum that may be carried into a collision o Excessive speeding can put unnecessary stress on the engine and gearbox

Harsh cornering
Newer IVMS units can measure harsh cornering through XYZ accelerometers. This is now widely available on new IVMS products and has been adopted as a minimum standard for all new devices bought after February 2017 within the Shell fleet. The means of capturing the vent is the same as for harsh acceleration and harsh braking.
IVMS Limitation: Similar to harsh braking, IVMS data cannot determine if the harsh cornering was avoidable or not. However also with harsh cornering, there are two things that can be derived from analysis of data: 1) Identify those drivers who harsh corner often, on many locations: that is a driver behavioural trait that needs to be addressed. 2) Identify those locations where all drivers harsh corner, that is an infrastructure issue. In this case opportunities for planning alternative routes should be considered and driver briefings should alert all drivers to the risks associated with these locations.

Facts:
o Harsh cornering increases wear on tyres and suspension components o Frictional losses during harsh cornering contribute to increased fuel consumption o G-forces increase the risk of loads moving if not correctly secured and hence creating greater risk of further incidents (lost loads, rollovers) o Data from within the Shell fleet suggests a correlation between harsh cornering and rollover rate.

Eco-Band Driving or Smooth driving
Although running outside the Eco-Band primarily has a negative effect on fuel economy there is also an (indirect) relationship to safe driving. For smooth driving it is important for drivers to look ahead, predict and anticipate on how traffic situations will evolve (i.e. to have a good 'Situation Awareness' (SA; (Endsley, 1995)).
IVMS Limitation: IVMS can measure a number of factors which combined result in smooth driving and anticipation (harsh acceleration, braking, turning, engine revolutions). Although there is no overall 'metric' to measure smooth driving, this should be strongly inferred from events recorded for the four factors mentioned.
Facts: o The engine has an optimum speed range where it produces the most torque and this is known as the "Eco-Band" or "Green Band"

Location (route), date and time of a trip
IVMS can report route / location through GPS and date and time of trip. Time and space geofences can be set up if needed to flag deviations from desired route / route timings (e.g. night driving).
IVMS Limitation: The IVMS system can determine road type only crudely through the driven speed, but cannot determine specific denotations of road classification and any context that goes with that. With 'add ons' such as maps (local or remote), that carry information about the type of roads (e.g. urban or rural, speed limits, etc.) on the route, this information can be available.
Facts: o The driven routes are recorded for each trip and this can be used against other trip data to make safety judgements about location and time of certain routes

A number of other variables are measured with IVMS. They are primarily related to fuel consumption and reducing maintenance costs, but are not directly related to road safety. These variables are therefore less relevant for the MeBeSafe project. These are:

- **Harsh Acceleration**

  The acceleration of the vehicle should be steady and progressive where ever possible. Harsh accelerations will be recorded if it exceeds a pre-defined limit.
- **Engine Idling**

  Running an engine on idle with the vehicle stationary will impact the fuel economy since it is consuming fuel and not travelling any distance. This measure cannot determine why the vehicle was idling over the period of time, only that is was.

- **Over-Rev**

  Over-revving the engine also puts unnecessary strain on the engine components.

Relevant variables that cannot be measured with IVMS are the following:

- **Distraction**

  IVMS can only measure driving characteristics of the vehicle, not the underlying reason (e.g. distraction) of certain driving behaviour.

- **Close following**

  Although most Shell trucks are equipped with Adaptive Cruise Control (ACC), the following distance is not captured by IVMS.

- **Lane departure warning**

  Some trucks in the Shell fleet have a lane departure warning system installed as standard. Information on how often or on which locations this system gives a warning cannot be read or used for analysis.

- **External conditions (weather, congestion, etc.)**

  As the IVMS system only measures vehicle characteristics, data on traffic and weather conditions are not available

2.1.2 IVMS thresholds

In order to determine whether or not a particular driver's behaviour is 'good' or 'bad' relative to company requirements and expectations, a set of thresholds are used for Shell contractors. Each vehicle/driver is monitored and assessed against this set of thresholds for the different output measures. *Table 2.1* displays the thresholds for a number of IVMS variables.

Variables	Light Vehicles			Heavy Vehicle		
	Metric	Imperial	'G' Force	Metric	Imperial	'G' Force
Acceleration Threshold (HA)	10-12 kph/sec	6-7.5 mph/sec	0.27-0.3 g	6-8 kph/sec	4-5 mph/sec	0.17-0.2 g
Deceleration Threshold (HB)	10-12 kph/sec	6-7.5 mph/sec	0.27-0.3 g	10-12 kph/sec	6-7.5 mph/sec	0.27-0.3g
Harsh Turning (HT)	3.0 m/s 2	-	0.31g	3.0 m/s 2	-	0.31g
Maximum Speed	Follow published country or asset maximum speed limits					

Table 2.1 Thresholds for driver behaviour measured by IVMS that are used by Shell.

2.1.3 Conclusion on IVMS data

Not all vehicles currently utilised within the Shell Fleet use and report on IVMS data. Furthermore the quality and detail of available data differ between contractors, hauliers, country and region. The thresholds are consistent overall but some hauliers have set thresholds in a different format that is not necessarily easy to compare across fleets.

Thus, IVMS can reliably measure some of the relevant variables for the MeBeSafe project; but not all of them.

2.2 Smartphone data

Shell has developed both telematics and smartphone based technology to support its commercial and digital services offerings for businesses and consumers. Smartphones have recently been seen not only as a communication interface, but as a measurement device as well. The ability to capture more data points at a fraction of the cost compared to IVMS is attractive, but we need to establish what information can be captured directly and indirectly to enable us to evaluate the effectiveness, applicability and suitability.

In the fleet environment, the smartphone is considered as a secondary device to any other telematics device because of potential smartphone security (in terms of theft) and the fact that it can be switched off (i.e. driver chooses when to record or not record). Typically, large fleet companies have sophisticated and necessary telematics to help manage their operations, but smartphone technology is still a contender as an alternative and can appeal to both these large fleets, where complimenting information is available, as well as the smaller fleets, whose investment levels into telematics are not viable.

2.2.1 Smartphone / app variables

An app that uses mobile phone measurements can define a combination of driving behaviours against defined targets and weightings in order to monitor basic driving styles. These driving behaviours are:

- Harsh Acceleration – number of times thresholds are exceeded which vary with speed
- Harsh Braking – number of times thresholds are exceeded which vary with speed
- Smooth Driving – how well the driver controls changes of speed (both increasing and decreasing)

Each individual behaviour rating (e.g. 0-100% of the driving time, or number of events) can be calculated at the end of each journey. The individual measurements are aggregated to produce an overall, weighted performance score. For any measurement the following attributes could be included:

- Date and time
- Vehicle speed
- GPS Lat/Long

In theory, this will allow us to correlate and integrate other information with common attributes (i.e. type of roads, darkness, weather), allowing to measure driving behaviour in a certain context.

Other measurements could be included in this weighting as more inputs are considered. In theory the following variables can be measured using a smart phone:

- **Speeding**

  It should be possible to measure speeding behaviour by combining the drivers' speed with local speed limits. However, it is not possible to determine if the speed was suitable for the driving conditions. The measurement of driving speed can be enhanced with accelerometers, but this would mean that the smartphone would need to be docked. In addition, the responses are known to vary device to device and brand to brand, so the accuracy of these devices will need to be evaluated accordingly.
- **Harsh cornering**

  This can be measured using the accelerometers in the smartphone, but again this creates a docked device dependency and differences between devices. Alternative options include use of map data attributes to help determine harsh cornering; but this could be quite expensive.
- **Fuel consumption**

  Whilst we do not have a connection to vehicle CANbus to measure this directly, the impact of driving style on fuel consumption can be estimated on a relative basis. This can help drivers better understand which manoeuvres demand most fuel consumption. A comparison to actual fuel consumption could provide information on the accuracy of this estimate (algorithm developed by Shell).

- **Location information (i.e. urban/non-urban)**

  Currently it is not possible to determine road type. In theory it is possible to combine GPS information with public map data; but costs and accuracy need to be evaluated.

- **Conditions (weather, congestion, etc.)**

  Currently we cannot determine local weather conditions to make any judgements against any resulting impacts. In theory it is possible to combine GPS information with public map data; but costs and accuracy need to be evaluated.

Relevant variables that <u>cannot</u> be measured with a mobile phone or an app are the following:

- **Distraction**

  A mobile phone / driving app cannot capture the underlying reason (e.g. distraction) of a certain driving style.

- **Close following**

  At the moment, a mobile phone cannot measure the distance to a lead vehicle. We can explore the use of proximity sensors, vehicles are increasingly being produced with such safety systems built-in.

- **Lane departure warning**

  This cannot be measured with a mobile phone.

2.2.2 Smartphone thresholds

In order to determine whether or not a particular driver is 'good' or 'bad' relative to company requirements and expectations, a set of thresholds for driving behaviour variables have been established by Shell. Each vehicle/driver is monitored and assessed against this set of thresholds for the different output measures. *Table 2.2* displays the thresholds for a number of variables.

Parameter	Light Vehicles
Acceleration Threshold (HA)	1.0 m/s2
Deceleration Threshold (HB)	1.0 m/s2
Smooth Driving	+/- 0.75 m/s2

Table 2.2 Thresholds for driver behaviour measured by mobile phone that are used by Shell.

2.2.3 Conclusion on Smartphone data

Using a smartphone as a measurement device to collect data on driving behaviour enables quick installation for many vehicles, because no expensive (IVMS) technology has to be installed in the vehicle. A smartphone can be seriously considered as a data logging system as well as an audio-visual interface for the driver. Through careful and clever development of algorithms, the smartphone will measure a high proportion of fleet operations and driving behaviours.

The security of the smartphone and making mandatory use of this are challenges that should be evaluated specifically. It is assumed that there is a level of general adoption of smartphones within targeted fleets, which enables easier integration and familiarisation of drivers with smartphones.

2.3 Conclusions

IVMS and an app/mobile phone can measure (roughly) the same variables (see *Table 2.3*). IVMS is more accurate; but not all vehicles are equipped with (the same) systems, which makes it very hard to compare drivers (for evaluation). Furthermore when the final app that will be used for coaching and coaching scheme that will be developed in the MeBeSafe project rely on a certain type of IVMS data, this will exclude many companies from implementing the app and coaching scheme in their organisation. A disadvantage of using a mobile phone to measure driving behaviour is that it relies on how many drivers own and turn on their mobile phone while driving.

Variables	IVMS	Mobile phone
(unnecessary) harsh braking	+	+
Speeding	+ / - Only if we connect it to map data; but cannot give information real time	+ Has to be connected to map data
Distraction	--	--
Close following	--	--
Harsh cornering	+	+/-
Lane departure warning	--	--
Fuel consumption	++	+
Location (route), date and time of a trip	+	++
Urban/rural route	Possibly if combined with public map data	Possibly if combined with public map data
Conditions (weather, congestion, etc.)	--	Possibly if combined with public map data
Advantages	Accurate information	All drivers can use it (no need to install expensive equipment)
Disadvantages	Not all vehicles are equipped with (the same) systems	We need to ensure drivers turn the device on

Table 2.3 The KPI variables that can be measured with IVMS and mobile phone.

In the next chapter thresholds that are used for detecting risky behaviour events are explored as well as the influence of driving context on driving behaviour,

3 Driver profiling based on Naturalistic Driving data

This chapter explores whether to include driving context in driver profiling and what thresholds could be used for braking and speeding events. To do so, naturalistic driving data has been used to study actual driving behaviour. Furthermore, the relation of data collected by questionnaires with actual driving behaviour is investigated.

A harsh braking manoeuvre is sometimes unavoidable, for example when a child suddenly emerges from between two parked cars to cross a street (note: this does not mean that the driver is not to blame). Other harsh braking manoeuvres may, in hindsight, be avoidable, such as when a driver chooses to drive above the speed limit just before entering an intersection, the traffic light suddenly turns orange, and the driver does not wish to risk running a red light. Ideally drivers anticipate the occurrence of potentially dangerous situations, so that they do not have to brake (harshly) at all.

Therefore, in a coaching setting one should not focus on the dichotomous question of whether drivers perform a harsh braking manoeuvre or not (i.e., sometimes they are unavoidable). Instead, one should review to what extent a driver differs from his/her colleagues (or compared to a previous measurement of the same driver) in terms of how often, how harsh, and in which situations harsh braking manoeuvres were performed.

With regard to speeding, professional truck drivers have to balance the fine line of driving as fast as possible within tolerable limits on the one hand (i.e., time is money), and driving as safe as possible on the other hand. Analogous to harsh braking, one should review how often speeding occurs across drivers, with which magnitude speeding occurs, and if speeding occurs in specific situations (e.g., at roads with particular speed limits). Harsh braking events may co-occur with speeding events, for instance if the earlier described traffic light scenario occurs often. Co-occurring events may be of particular interest for a coaching session, since an improvement on the one event type is likely to yield an improvement on the other event type (here: a lower driving speed will reduce the harshness of braking events).

Naturalistic Driving (ND) data offers an opportunity to study harsh braking and speeding events as they occur in a truck driver's natural, everyday environment. The challenge of analysing ND data is that events of interest (here: harsh braking and speeding) have to be found in a relatively large database. It is common practice in ND studies to implement triggers to identify events, for instance by comparing the momentary value of longitudinal acceleration with a pre-defined negative threshold value (note: a negative acceleration values corresponds with deceleration). The start of an event is then marked by the moment when the longitudinal acceleration drops below the threshold, and the end is marked by the moment when the signal increases rises above the threshold again.

However, the actual threshold value that has been chosen to identify harsh braking events differs greatly across previous ND studies with trucks. For example, the NDTS project (Olson et al., 2009) uses a threshold of -1.96 m/s^2 to identify safety critical events, whereas the DDWS FOT project (Olson et al., 2009) uses a speed-dependent threshold of either -3.43m/s^2 (driving speed ≥24km/h) or -4.91m/s^2 (driving speed <24km/h). The EuroFOT project (Malta et al., 2012) also uses a speed-dependent threshold that increases linearly from -5.4m/s^2 to -3.6m/s^2 when the speed increases from 50km/h or less to 150km/h. Furthermore, personal communication with staff at Shell revealed that they use a threshold of -2.65 to -2.94m/s^2 in their in-vehicle monitoring systems (also see *Section 3.1.2*). To summarise, there is no agreement on the acceleration threshold beyond which one speaks of a harsh braking event.

As previous studies have not provided sufficient input for a choice of threshold values, a focused study has been performed on the database of the UDRIVE project (Van Nes et al., 2018). In the UDRIVE project, a fleet of trucks from four Dutch transport companies has been equipped with multiple video cameras and sensors, through which continuous driving data (e.g., acceleration, local speed limits) has been collected. The primary aim of this study is to explore the impact of trigger threshold values on the frequency of emerging events across driving contexts. The results may inform the conceptualization and/or implementation of driver profiles. The ND studies reported earlier used either a fixed or a speed-dependent threshold value. Consequently, the chosen threshold value could not be evaluated compared to alternative threshold values. In contrast, the threshold value was systematically manipulated in the present study. First, a liberal value was chosen as initial trigger threshold (e.g., a deceleration relatively close to zero), resulting in a large number of events. Second, a categorization of the resulting events was performed according to the peak value registered in each event (e.g., the maximum deceleration value), and in terms of driving context (e.g., speed limits).

The secondary aim of this study is related to self-reported attitudes on driving, personality, and age. Data collected by questionnaires are typically related to self-reported and registered traffic offenses to determine validity (see *Chapter 6* of this deliverable). In contrast, the UDRIVE database offers the opportunity to relate questionnaire data to actual driving behaviour, including unreported and unregistered behaviour. Therefore, this study also explores the relation between attitudes and personality on the one hand, and the frequency of harsh braking and speeding events on the other hand.

3.1 Method

3.1.1 Description of truck drivers, trucks, and records in UDRIVE

Behavioural and subjective data has been collected from 43 drivers in the UDRIVE truck database, using Volvo FL and Volvo FM trucks. There were 42 males and 1 female, with ages between 22 and 71 years, of which three drivers had an age below 30 years (M = 49.0, SD = 11.2). The dataset consists of 54658 records with a minimum distance of 1km (maximum: 561km, M = 10.8, SD = 17.0), and a total distance of

592700km. Most records covered a relatively small distance. For example, 39803 records (i.e., 73% of the sample) had a distance below 10km, and only 1725 records (i.e., 3% of the sample) had a distance above 50km. The distribution of record distance for the Volvo FL and Volvo FM trucks is similar. Therefore, no further distinction will be made between the truck types in subsequent analyses.

Figure 3.1 shows that the number of records and the total distance driven differ across drivers. Although the majority of records fall within the range of 1-10km for all drivers (left panel of *Figure 3.1*), the total distance covered per driver (right panel) is typically accumulated through records longer than 10km. In case of one driver, the distance covered originates mainly from records with a distance longer than 50km.

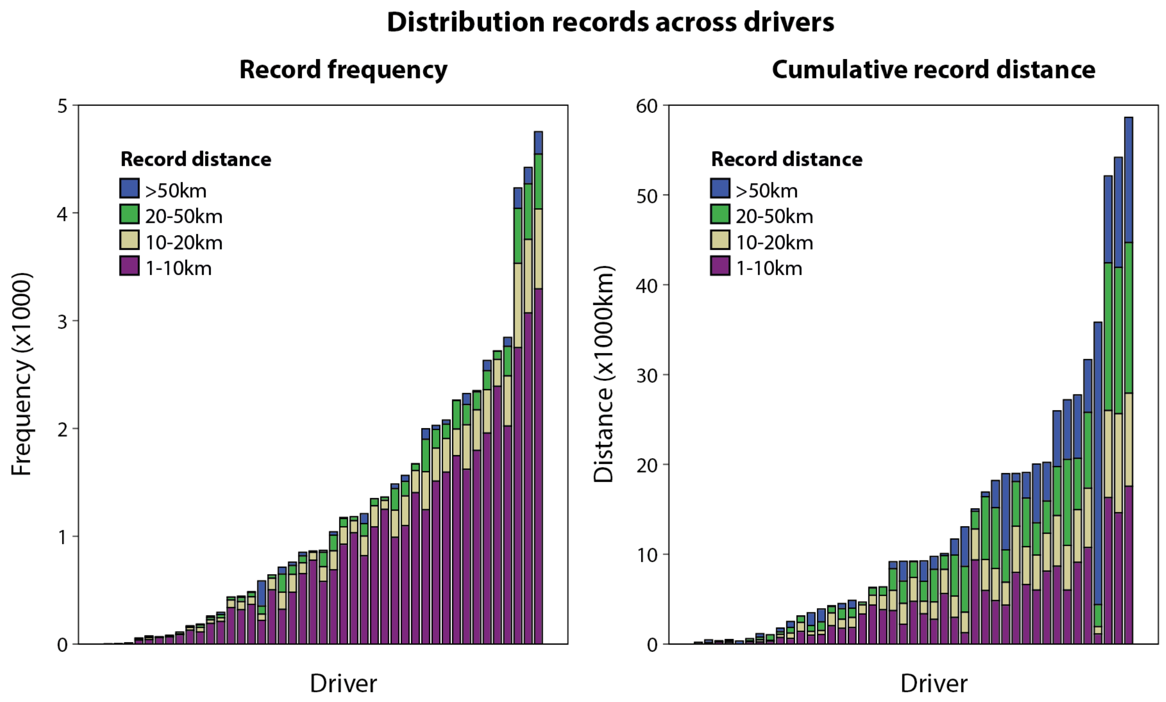


Figure 3.1 Frequency of records (left) and cumulative distance (right) per driver, as function of record distance category. NOTE: the order in which the drivers are represented differs between the panels.

3.1.2 Trigger implementation

Three event triggers have been implemented in the UDRIVE truck database for the purpose of the present MeBeSafe investigation. Two of them correspond with harsh braking (i.e., through longitudinal acceleration and longitudinal jerk), and one with speeding. *Table 3.1* summarizes the trigger implementation. Triggers for harsh braking have only been implemented for records with Volvo FM trucks (N = 26), because the Volvo FL trucks did not provide a longitudinal acceleration signal.

Trigger type	Description
Harsh braking (1)	Brake pedal depressed, longitudinal acceleration smaller than or equal to -2.94m/s^2, with driving speed at the event onset larger than or equal to 5km/h.
Harsh braking (2)	Brake pedal depressed, longitudinal jerk (i.e., first time-derivative of longitudinal acceleration) smaller than -2m/s^3 and longitudinal acceleration smaller than -2m/s^2, with a minimum duration of 0.2s and driving speed at the event onset larger than or equal to 10km/h.
Speeding	Driving speed larger than the posted speed for a minimum duration of 15s. Four categories of speeding above posted speed: light (0-10%), moderate (11-15%), severe (16-20%), and extreme (>21%).

***Note:** Multiple triggers within a 2 seconds window have been joined into a single event. Only triggers in records with a minimum total distance of 1km have been included.*

Table 3.1 Initial thresholds for five event triggers.

Several map matching attributes have been collected for each of the trigger types in *Table 3.1*, such as the posted speed at the onset of an event, and whether the event took place at an intersection. In addition, the peak value of longitudinal acceleration has been collected to explore the relation between threshold values and event frequency.

3.1.3 Subjective data on attitude and personality

Within the UDRIVE project, subjective data on (driving-related) attitudes and personality has been collected for each participant through five questionnaires with Likert scales. The Driver Attitude Questionnaire (DAQ; Parker et al., 1996) explores attitudes toward speeding and close-following. The Driving Behaviour Questionnaire (DBQ; Lajunen et al., 2004) involves errors, ordinary violations, and aggressive violations. The Driving Style Questionnaire (DSQ; French et al., 1993) contains 15 questions related to, among others, speeding, calmness, and focus. The Traffic Locus Of Control (T-LOC; Özkan & Lajunen, 2005) explores perceived causes of road accidents, and finally, the Arnett Inventory of Sensation Seeking (AISS; Arnett, 1994) concerns personality in terms of the sub-scales intensity and novelty.

3.1.4 Procedure

The harsh braking (2) and speeding (1) triggers have been implemented in the UDRIVE database through Matlab scripts. Triggers where the speed limit was unavailable were excluded. Next, the mean frequency of harsh braking events across speed limits was calculated. As will be shown, there appear to be three speed limit clusters with distinct harsh braking event frequencies. Therefore, subsequent analysis of harsh braking events on the level of individual drivers has been performed separately for each speed limit cluster. Triggers related to speeding events have also been examined across speed limits. However, no analysis on the level of individual drivers has been performed, because none of the observed speeding events could have resulted in a financial penalty according to Dutch legislation.

Harsh braking events have been categorized according to their maximum acceleration or jerk value. The correlation between event frequencies across speed limit clusters has been calculated for each acceleration threshold category, using SPSS v.24. Based on the correlations, an acceleration threshold was chosen for subsequent linear multiple regression analysis on the relation between driving attitudes and event frequencies. Predictors in this analysis included only attitude sub-scales with a Cronbach alpha larger than 0.6.

3.2 Results

Three drivers have been excluded from analysis. Two Volvo FM drivers were excluded because speed data were missing and one Volvo FL driver because questionnaire data was missing. Hence, the remaining sample corresponds with 40 participants, of whom 24 drove a Volvo FM truck.

3.2.1 Harsh braking events across speed limits

Only Volvo FM trucks have been considered in the analysis on harsh braking events, as Volvo FL trucks did not provide a longitudinal acceleration signal. Of the 24 Volvo FM drivers in the sample, 17 featured at least one harsh braking event. All 24 Volvo FM drivers featured one or more harsh braking events. In total, 8303 harsh braking trigger events were found based on longitudinal acceleration. In 3028 events, the brake pedal was activated, and each of those events occurred in a unique record. The average distance in these records was 17.6km (*SD* = 24.0km, minimum: 1.0km, maximum: 277km), which is slightly larger than, but within one standard deviation of, the average distance of entire sample of records. The speed limit was unknown in 973 cases, and incorrect in 24 cases (i.e., a speed limit of 90km/h is normally not used in the Netherlands, where the data were collected). Furthermore, the longitudinal jerk trigger yielded a dataset consisting of 201 events (including the use of the brake pedal), of which 62 cases featured an unknown or incorrect speed limit. Consequently, 2031 longitudinal acceleration events and 139 longitudinal jerk events have been used in subsequent, separate analyses. *Table 3.2* provides an overview of the driving contexts in which the harsh braking trigger events have been found.

Trigger signal	Location	Speed limit (km/h)							Total
		30	50	60	70	80	100	120&130	
Longitudinal Acceleration	Intersection	68	256	6	29	83	17	8	467
	Elsewhere	78	631	47	359	310	91	48	1564
	Total	146	887	53	388	393	108	56	2031
Longitudinal Jerk	Intersection	2	10	1	2	16	1	0	32
	Elsewhere	4	41	2	25	24	9	2	107
	Total	6	51	3	27	40	10	2	139
Total distance (x1000km)		3.3	24.5	2.9	12.1	33.6	82.7	68.0	227.0

***Note:* Volvo FL trucks have been omitted from the total distance, because no harsh braking trigger could be implemented. Speed limits 120km/h and 130km/h have been merged due to their similar road design.**

Table 3.2 Overview of harsh braking events as function of speed limit.

Most harsh braking events occurred at a speed limit of 50km/h, followed by 80km/h and 70km/h. However, the distance covered at roads with a speed limit of 50km/h was shorter than the distance at 70km/h and 80km/h roads. Furthermore, *Figure 3.1* shows that for some drivers substantially more data has been collected than for others. For each participant, the number of events per km has been calculated to account for such differences in distance. Also, the average number of events per km was calculated across drivers, where a weight was applied to each driver proportional to the distance driven by that driver at each speed limit. The result of these calculations is displayed in *Figure 3.2*, in which the horizontal axis contains seven speed limits, and the vertical axis displays the event frequency.

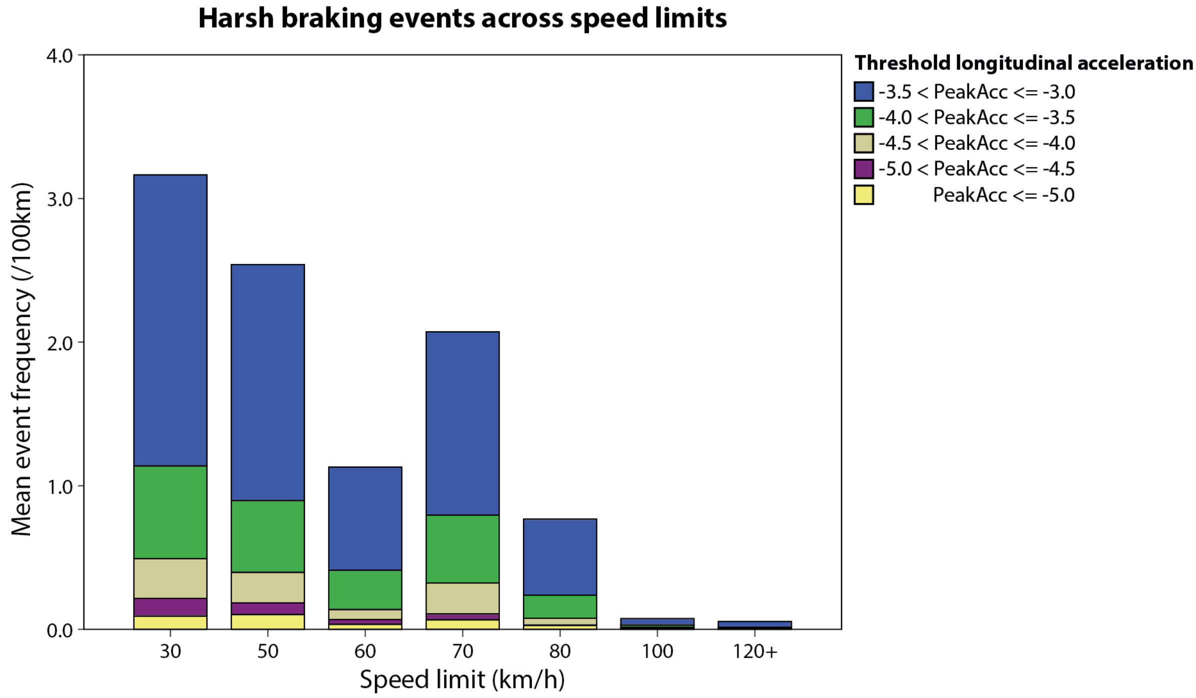


Figure 3.2 Harsh braking event frequency as function of speed limit and threshold for longitudinal acceleration. NOTE: PeakAcc = lowest longitudinal acceleration value within event, in m/s2. See text for more details.

The minimum value of longitudinal acceleration ('PeakAcc' in *Figure 3.2*) was recorded for each harsh braking event to enable a comparison of event frequencies at five threshold values: -3.0m/s^2, -3.5m/s^2, -4.0m/s^2, -4.5m/s^2, and 5.0m/s^2. These values span the range found in previous ND studies with trucks (Malta et al., 2012; Olson et al., 2009). At each speed limit, most events had an acceleration value between -3.0m/s^2 and -3.5m/s^2. In *Figure 3.2* this finding is represented by the blue part of each bar, where each 'stack' represents the event frequency in the threshold range set by two subsequent threshold values.

At the most liberal threshold (i.e., -3.0m/s^2), it appears there are three clusters of speed limits with a similar event frequency (i.e., total bar height). The highest event frequencies are found in urban areas (speed limits: 30, 50, 70km/h). In rural areas (speed limits: 60, 80km/h) the average event frequency is approximately 2-3 times lower than in urban areas. Finally, at highways (speed limits: 100, 120+km/h) the average event frequency is about one tenth that of rural areas (note: trucks were restricted at 85km/h). The distributions at rural roads and highways were significantly different from the normal distribution. Therefore, a Friedman ANOVA was performed, which yielded a significant effect on speed limit cluster, $\chi^2(2) = 36.55$, $p < .001$. Two Wilcoxon signed ranks tests were used for post-hoc comparisons. A Bonferroni correction was applied, such that the significance was tested against an alpha $\alpha = .025$. The event frequency at urban roads (*Mdn* = 1.92 events/100km) proved to be significantly higher than at rural roads (*Mdn* = 0.48 events/100km), $T = 1$, $p < .001$. Also, the event frequency at rural roads was significantly higher than at highways (*Mdn* = 0.018 events/100km), $T = 0$, $p < .001$.

With regard to threshold values, *Figure 3.2* shows that a threshold of -3.0m/s^2 at rural roads yields approximately the same event frequency as a threshold of -3.5m/s^2 at urban roads. Likewise, a threshold of -3.0m/s^2 at highways yields approximately the same event frequency as -5.0m/s^2 at urban roads, and -4.5m/s^2 to -4.0m/s^2 at rural roads.

The above approach has also been applied to harsh braking events based on longitudinal jerk (i.e., the first time-derivative of longitudinal acceleration, expressed in m/s^3), where high jerk values provide an indication of the speed with which a harsh brake is accomplished. *Figure 3.3* displays the frequency of harsh braking events per 100km, in which the mean frequency has been weighted according to the distance travelled by each participant (i.e., the same procedure as with harsh braking based on longitudinal acceleration). Speed limits found in urban areas (i.e., 30, 50, 70km/h) show a similar event frequency at the most liberal threshold of longitudinal jerk (i.e., -2m/s^3). Likewise, clusters are found for speed limits in rural areas (i.e., 60, 80km/h), and highways (100, 120, 130km/h). The clusters correspond with those found with harsh braking events based on longitudinal acceleration, see *Figure 3.2*. However, the rate at which longitudinal jerk triggers occur is approximately one order of magnitude lower than the longitudinal acceleration triggers.

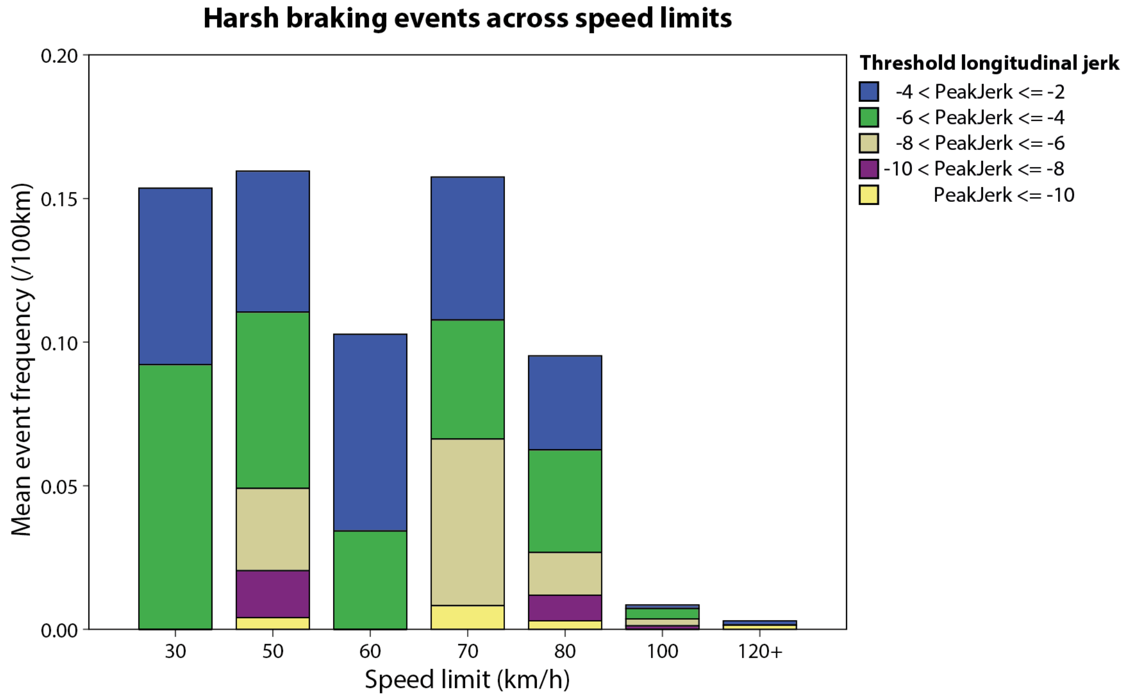


Figure 3.3 Harsh braking event frequency as function of speed limit and threshold for longitudinal jerk. NOTE: the longitudinal acceleration throughout each event was smaller than or equal to -2.0m/s2.

3.2.2 Harsh braking based on longitudinal acceleration across drivers

The previous section revealed three speed limit clusters (i.e., urban, rural, highway) with similar event frequencies. The present section incorporates those clusters when exploring how often drivers perform harsh braking events, and at which thresholds harsh braking events occur. Due to relatively low frequency of harsh braking events based on longitudinal jerk (see *Table 3.2*), the analysis at the level of individual participants has been performed only with events based on longitudinal acceleration.

Figures 3.5, 3.6, and *3.7* show event frequencies across drivers for urban, rural, and highway speed limits, respectively. Cleary, most drivers (but not all) perform harsh braking manoeuvres, yet the frequency at which they occur varies across drivers. Currently, the drivers have been ordered based on their event frequency at the most liberal threshold value (i.e., -3.0m/s^2). However, the ordering would change if a more conservative threshold value had been chosen.

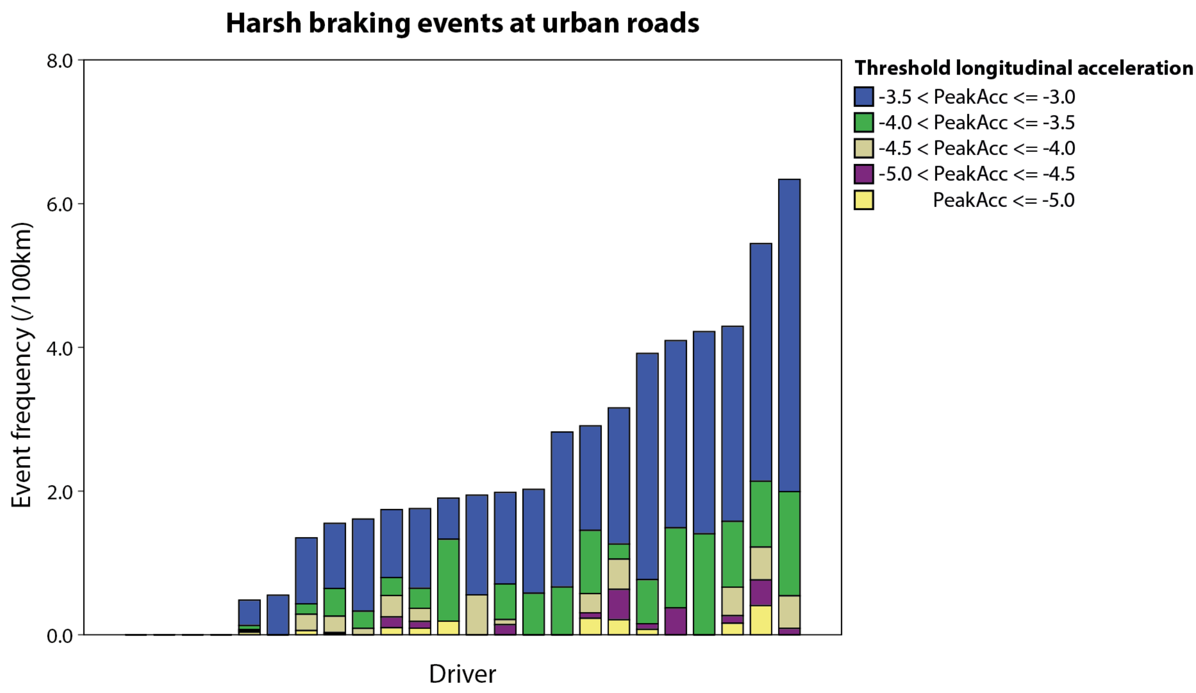


Figure 3.4 Harsh braking events across drivers within urban areas (speed limits: 30, 50, 70km/h).

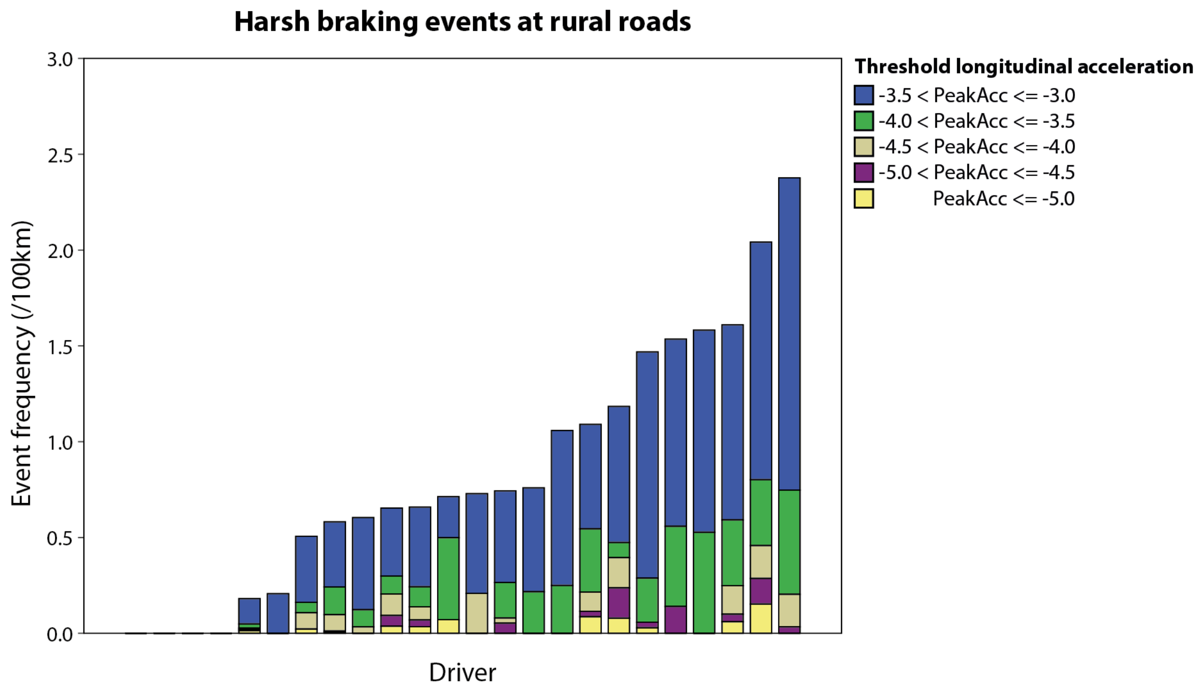


Figure 3.5 Harsh braking events across drivers within rural areas (speed limits: 60, 80km/h).

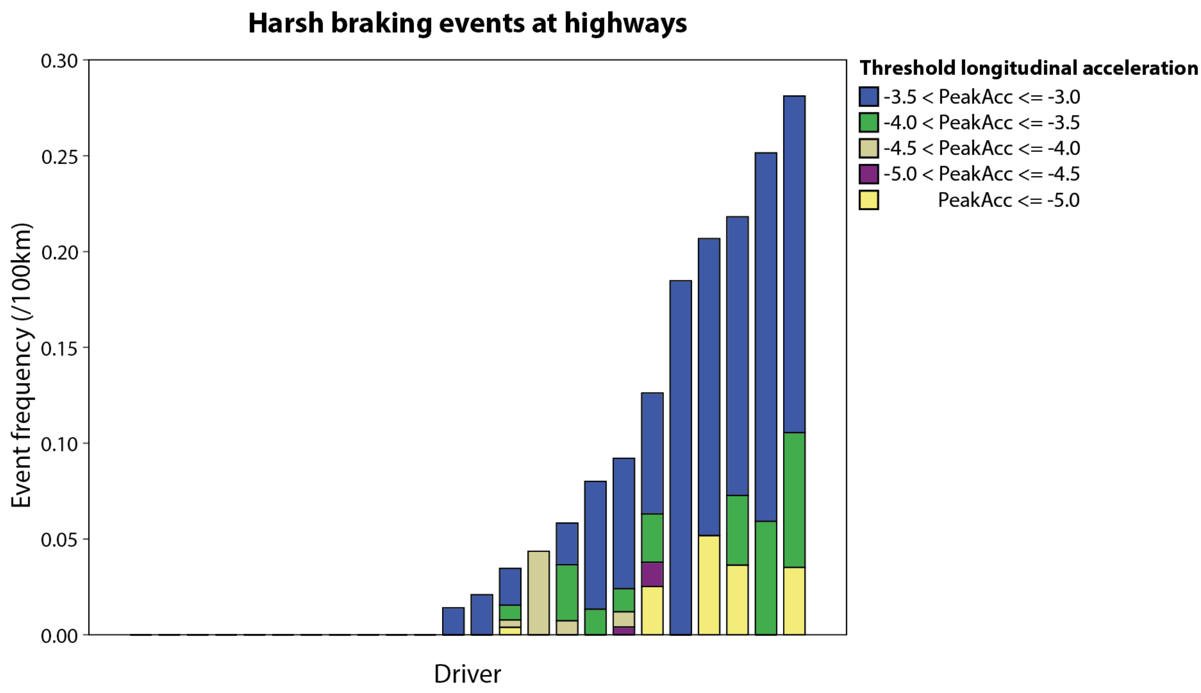


Figure 3.6 Harsh braking events across drivers at the highway (speed limits: 100, 120, 130km/h).

In *Figures 3.5* to *3.7*, the drivers are ordered by event frequency. The ordering was performed independently for each figure. Consequently, the order of the drivers differs somewhat across the figures. Nonetheless, the similarity in the overall shapes of the distribution raises the question of, do drivers who show more events in urban speed limits also show more events at rural and/or highway speed limits? Therefore, the correlation was calculated between event frequencies across the speed limit clusters, once for each of the trigger thresholds that were used previously in *Figure 3.2* significant positive correlations were found between each speed limit cluster at the most liberal trigger threshold (i.e., PeakAcc ≤ $-3.0 m/s^2$), see *Table 3.3.* When the threshold is set to an increasingly conservative value, the magnitude of the correlations and their significance declines. This finding is a preliminary indication that drivers differ in where they perform very harsh braking manoeuvres, at least when the driving context is operationalized in terms of urban, rural and highway speed limits.

Location	PeakAcc ≤ -3.0			PeakAcc ≤ -3.5			PeakAcc ≤ -4.0			PeakAcc ≤ -4.5			PeakAcc ≤ -5.0		
	Threshold longitudinal acceleration (m/s^2)														
	U	R	H	U	R	H	U	R	H	U	R	H	U	R	H
Urban	1	.68**	.67**	1	.51*	.58**	1	-.15	.46*	1	.25	.38	1	-.09	.18
Rural	.	1	.73**	.	1	.39	.	1	-.10	.	1	-.15	.	1	-.13
Highway	.	.	1	.	.	1	.	.	1	.	.	1	.	.	1

NOTE: U = Urban, R = Rural, H = Highway. * p < *.05*, ** p < *.01.*

Table 3.3 Pearson correlation on event frequency between speed limit clusters, stratified across trigger thresholds.

Summed over all records, the median number of intersections that drivers passed was 7286 (minimum: 313, maximum: 58163). According to *Table 3.2*, on average approximately 23% of the harsh braking events took place at an intersection (urban: 25%, rural: 20%, highway: 15%), regardless of whether the trigger was based on longitudinal acceleration or jerk. Using the former measure, *Figure 3.7* displays the frequency of harsh braking events taking place at an intersection, expressed as proportion of the total number of intersections passed by each driver. At the most liberal threshold, most drivers perform between 0.5 and 1.5 harsh braking manoeuvres every 1000 intersections. One driver performed 4 manoeuvres every 1000 intersections. However, this driver did not show the highest event frequency when the most conservative threshold was used. In other words, when ordering drivers based on event frequency, the order strongly depends on the threshold that is chosen. This was also the case when the event frequency was expressed as function of distance.

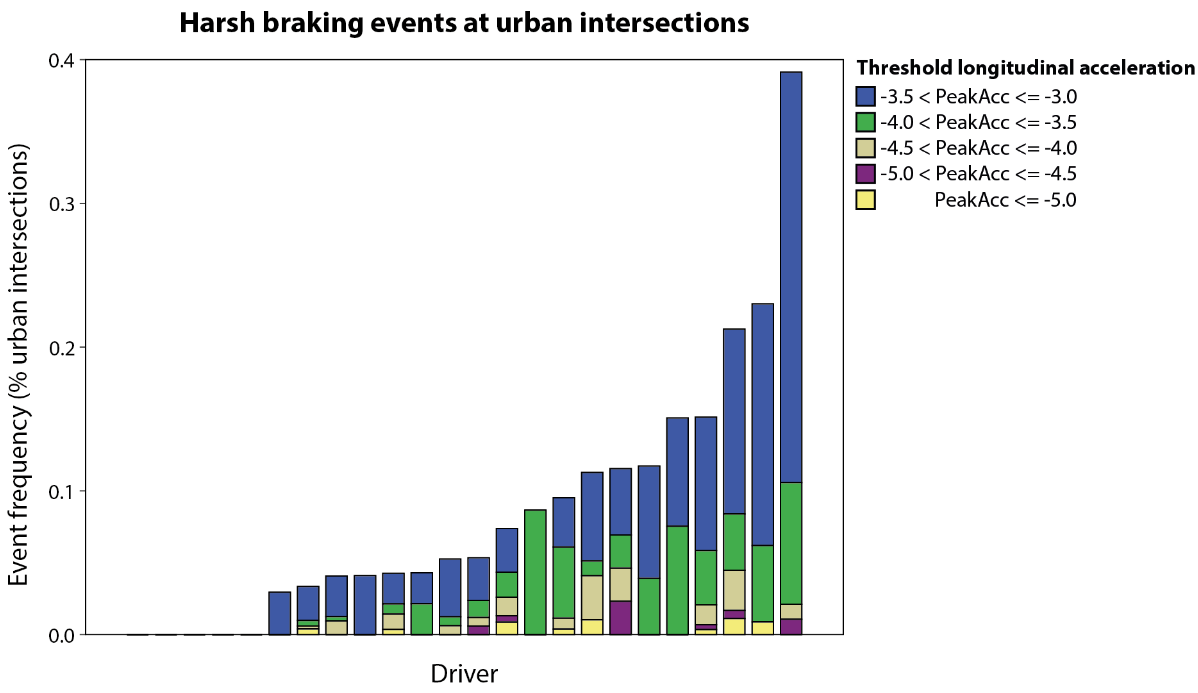


Figure 3.7 Harsh braking events (longitudinal acceleration) across drivers at intersections.

3.2.3 Speeding behaviour

We focus at speed limits of 80km/h or less, because the trucks were equipped with a speed limiter. Looking at both Volvo FM and Volvo FL drivers, no moderate, severe, or extreme speeding events were found. In contrast, 38051 light speeding events were found, in which the driving speed was up to 10% above the posted speed limit. The light speeding events had an average duration of 39.0s (minimum: 15.1, maximum: 1169.6, *SD* = 36.9), and the average distance covered in those events was 728.5m (minimum: 131.3, maximum: 26459.0, *SD* = 852.9). For each driver, the distance driven while speeding in a certain speed limit category was divided by the total distance driven in that speed limit category (regardless of speeding). The resulting proportion was averaged across the participants, where the weight of each participant in the calculation was inversely proportional to the distance covered by each participant. *Figure 3.8* shows that the driving speed in areas with a speed limit of 30km/h is too high (i.e., speed range 30<v ≤33km/h) in 29% of the distance that was covered. At a limit of 80km/h, the proportion of speeding (i.e., speed range 80<v ≤88km/h) has decreased to 13% of the distance.

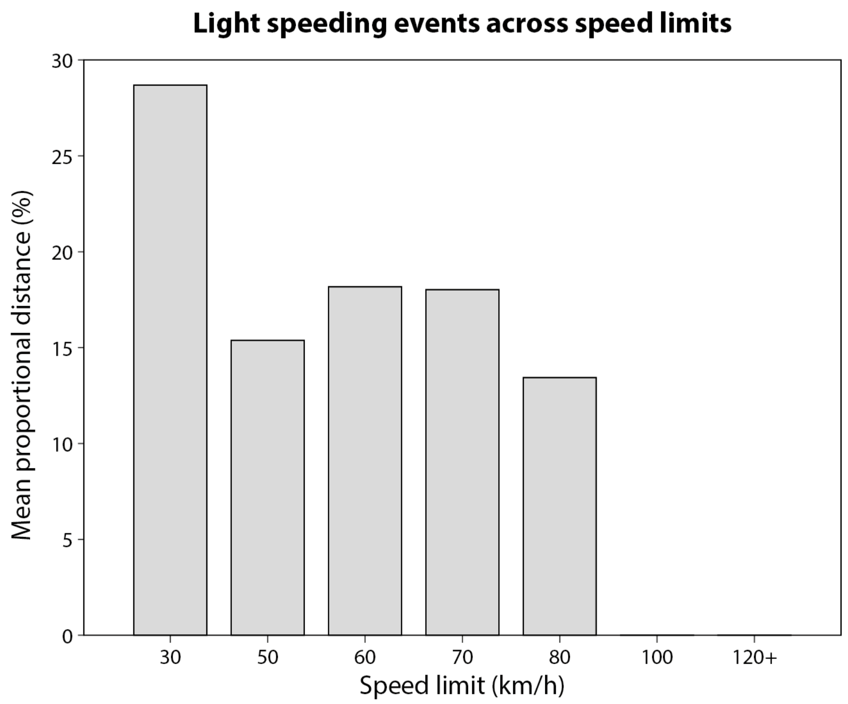


Figure 3.8 Light speeding events as function of speed limit.

Although it is technically speeding, in the Netherlands (i.e., where the data were collected) no fines are given when the driving speed is less than 7km/h above the speed limit, given that the speed limit is below 100km/h. Within the light speeding category, only speeding events at a speed limit of 80km/h could in theory exceed the fine threshold of 7km/h. An evaluation of a selection of the records indicates that the Volvo FM and Volvo FL trucks were capped at 85km/h. Thus, it appears that participants often sought to push the boundaries of maximum speed choice, but it is unlikely that small violations of the speed limit could have resulted in a fine. Due to the latter, no further analysis has been performed at the level of individual drivers.

3.2.4 Relation between triggers

This section examines the co-occurrence of harsh braking events and light speeding events. The examination of light speeding events involves only Volvo FM trucks (*N* = 16690 events), because a harsh braking trigger could not be constructed for Volvo FL trucks.

Of the 139 events with longitudinal jerk, 74 events appear in records that also have an event with longitudinal acceleration, 66 of which appear within 1.0s of the start or end of a longitudinal acceleration event (i.e., 47% of all events). This finding suggests that harsh braking events based on longitudinal jerk are not independent from events based on longitudinal acceleration.

Of the 2031 events with longitudinal acceleration, 1139 appear in records that also have one or more light speeding events. Of these 1139 events, 69 events overlap with a light speeding event, and another 76 events appear within 5.0s following a light speeding event. When combined, this accounts for 7% of all events with longitudinal acceleration; a finding which suggests that harsh braking events occur independent of light speeding events.

3.2.5 Relation between attitude, personality, and harsh braking

The questionnaires on attitudes and personality have not been used on Dutch truck drivers in previous studies. Therefore, a reliability analysis using Cronbach's alpha was performed on the sub-scales of each of the questionnaires. Items with a negative corrected item-total correlation were removed. After removal, sub-scales with a Cronbach's alpha >0.6 have been selected for further analysis, these being: DAQ_Speeding, DAQ_Close_Following, DBQ_Errors, DSQ_Speeding, TLOC_Self, TLOC_Others, and AISS_Intensity. The descriptive statistics in *Table 3.4* indicate that many drivers believed it was acceptable to engage in speeding, that they did not often commit errors, and that the cause of accidents was generally other drivers, as opposed to themselves.

Questionnaire	Sub-scale	Likert scale range	*M*	*SD*	Number of items	Cronbach's alpha
DAQ	Speeding	5	2.78	.49	8	.66
DAQ	Close following	5	2.44	.50	10	.73
DBQ	Errors	6	1.61	.43	8	.78
DSQ	Speeding	6	2.36	.86	3	.68
TLOC	Self	5	1.74	.49	5	.77
TLOC	Other drivers	5	3.14	.65	6	.76
AISS	Intensity	4	2.08	.44	7	.69

Table 3.4 Means of the selected sub-scales (Cronbach's Alpha>0.6) after removal of items with a corrected item-total correlation below zero.

The correlations between the participants' attitudes and personalities on the selected sub-scales, as well as participant age, are depicted in *Table 3.5.* In regards to the relation between the sub-scales, the strongest significant relationships were found between DAQ_Speeding and DAQ_Close_Following (r = .51), and between DBQ_Errors and TLOC_Self (r = .51), closely followed by the relations between DAQ_Close_Following and TLOC_Others (r = .47) and between DSQ_Speeding and AISS_Intensity (r = .47). With regard to sample characteristics, the only significant relation was between age and AISS_Intensity (r = -.41).

Sub-scale	1	2	3	4	5	6	7	8
1 DAQ Speeding	1	.51**	.082	.43**	.31*	-.32*	.23	-.022
2 DAQ Following		1	.15	.21	.33*	-.47**	.20	-.021
3 DBQ Errors			1	.24	.51**	-.12	.14	-.021
4 DSQ Speeding				1	.45**	-.078	.47**	-.221
5 TLOC Self					1	-.076	.42**	-.21
6 TLOC Others						1	-.15	-.21
7 AISS Intensity							1	-.41**
8 Age								1

NOTE: * $p < .05$, ** $p < .01$.

Table 3.5 Pearson correlations between attitudes, personality, and age.

A linear multiple regression analysis was performed on Volvo FM drivers ($N = 24$), with the selected sub-scales and age as predictors, and harsh braking frequency as dependent variable. The lowest threshold was used for harsh braking frequency. At this value ($-3.0 m/s^2$) the significant positive correlations between each speed limit cluster (see *Table 3.3*) indicates that the clusters can be grouped into one factor. Thus, the overall harsh braking frequency per participant was calculated by summing the clusters. Next, a stepwise regression procedure was used with backward elimination of the predictors.

The variance in harsh braking frequency was significantly explained by a model with DSQ_Speeding and AISS_Intensity as predictors, $F(2,23) = 4.38$, $p < .05$, $R^2 = .29$. In this model, an increase in AISS_Intensity corresponds with a significant increase in harsh braking frequency, $\beta_{std} = .65$, $t(21) = 2.86$, $p < .01$. In contrast, an increase in DSQ_Speeding corresponds with a significant decrease in harsh braking frequency, $\beta_{std} = -.52$, $t(21) = -2.28$, $p < .05$. The remaining sub-scales, as well as driver age, were excluded from the model as part of the backward elimination procedure.

It was expected that a positive attitude towards speeding would have been associated with a higher harsh braking frequency, but instead an association was found with a decreased harsh braking frequency. This raises the question to what extent the self-reported attitude toward speeding is valid for predicting actual speeding behaviour. For this reason, the relation between DSQ_Speeding (i.e., attitude) and light speeding (i.e., actual behaviour) was examined for Volvo FM drivers. The total speeding distance was summed for roads with speed limits between 30km/h and 80km/h. Higher speed limits have been omitted, because a speed limiter prevented driving faster than 85km/h.

A linear regression model with DSQ_Speeding as predictor and the proportional speeding distance as dependent variable proved to be significant, although the total variance explained by the predictor was limited, $F(1,23) = 4.37$, $p < .05$, $R^2 = .17$. According to the model, an increase in DSQ_Speeding corresponds with a significant increase in actual speeding, $\beta_{std} = .41$, $t(22) = 2.09$, $p < .05$. We discuss the apparent paradox between DSQ_Speeding, actual speeding, and harsh braking frequency in the next section.

3.3 Discussion

The primary aim of this study was to explore the impact of trigger threshold values on the frequency of unsafe driving behaviour across driving contexts. A secondary aim was to relate age, personality, and attitudes on driving styles to actual driving behaviour. These aims are discussed separately, followed by a reflection on the limitations of the present study.

3.3.1 Trigger thresholds for harsh braking and speeding

Most truck drivers in the UDRIVE database have been found to perform harsh braking manoeuvres, yet the event frequency varies across drivers. When drivers are ordered according to their harsh braking event frequency, the ordering changes when the trigger threshold is shifted from a liberal to a more conservative value. The implication of this finding with regard to driver coaching is that the interpretation of individual driver performance compared to fleet performance depends on the threshold that is chosen to identify harsh braking events.

With regard to driving context, it was found that the momentary speed limit significantly influenced the frequency of harsh braking events. At urban roads (speed limits: 30, 50, and 70km/h) the event frequency was approximately twice as high compared to rural roads (speed limits: 60 and 80km/h). In turn, the event frequency at rural roads was approximately ten times higher than events found at highways (speed limits: 100, 120, and 130km/h).

The differences in event frequency across speed limits may be explained by the fact that highways and rural roads, due to their absence of pedestrians and cyclists, are generally more predictable than urban roads. In addition, highways generally feature intersections where traffic merges in the same driving direction, whereas rural roads more often feature crossing traffic. A survey of the event video data showed that harsh braking events often take place in front of a traffic light. Thus, one would expect more events at intersections at rural roads. Indeed, the proportion of events at intersections was higher at rural roads (i.c., 20%) than at highways (i.c., 15%), but the difference is modest, and it has not been tested on significance. A larger difference may have occurred, though, when the trajectory directly preceding a traffic light would have been included (i.e., when approaching a queue). Additional analysis is required to explore this hypothesis.

Differences across drivers in event frequency as a function of trigger threshold are a preliminary indication of distinct driving styles. Some drivers perform many harsh braking events, but the magnitude of deceleration in each event is modest (e.g., above $-3.5 m/s^2$). Other drivers perform relatively few harsh braking events, but for those drivers the magnitude of deceleration is much larger (e.g., below $-5.0 m/s^2$). Furthermore, drivers appear to differ in where they perform the harshest of braking manoeuvres. When the threshold for lateral acceleration was set at a liberal value (i.c., $-3.0 m/s^2$), it was found that drivers with a high event frequency at urban roads also had a high event frequency at rural roads and highways. At more conservative threshold values, however, the following relation dissolved: drivers who perform their harshest braking events at urban roads are different drivers that those who perform them at rural roads or highways. Additional analyses and theory are required to explore how driving styles can be operationalized and validated.

Speeding was the second category of unsafe driving behaviour that has been investigated. Four categories of speeding have been constructed, based on the proportion of driving speed in excess of the momentary speed limit: light (0-10%), medium (11-15%), severe (16-20%), and extreme (>21%). In total, 38051 speeding events have been found, all of which corresponded with light speeding, but none of which showed driving speeds that could have resulted in a speeding ticket in the Netherlands, where the data were collected.

In transportation, drivers may be inclined to drive as fast as possible, because time is money. The consequence of a speeding ticket, however, stretches further than a monetary penalty: it may cost professional drivers their job, at least at hauliers with a relatively strict company policy. The UDRIVE project included only Dutch truck drivers from four transport companies in the Netherlands. The results based on this sample suggest that the added value of speeding triggers in driver coaching is limited from a safety perspective. Additional analysis with companies and drivers from other nationalities (and consequently, other driving cultures) is required to generalize the findings beyond the Netherlands.

3.3.2 Relation between age, attitudes, personality and unsafe driving behaviour

The relation between age, attitude, personality, and harsh braking event frequency was investigated through a linear multiple regression analysis. Variance in harsh braking was significantly explained by a model with AISS_Intensity (i.e., intensity of sensory experience) and DSQ_Speeding as predictors. In this model, an increase in AISS_Intensity results in an increased event frequency, whereas an increase in DSQ_Speeding results in a decreased event frequency.

With regard to DSQ_Speeding, it was expected that drivers who reported a positive attitude toward speeding would show more harsh braking events. However, the opposite effect was found in the regression model. This raises the question whether the DSQ_Speeding results are valid for predicting harsh braking events. In terms of internal consistency, the subscale had a Cronbach alpha of .68. This is an acceptable value for a subscale with relatively few items (Hair, Black, Anderson and Tatham, 2006), as is the case for DSQ_speeding (i.c., 3 items). Furthermore, a separate regression analysis confirmed that self-reported DSQ_Speeding was congruent with actual light speeding behaviour. Thus, there is no reason distrust the DSQ_Speeding results.

A causal relation between speeding and harsh braking events is unlikely, because only 7% of the harsh braking events overlapped with or directly followed light speeding events. Possibly, drivers have compensated for their relatively high driving speed by increasing their distance to a lead vehicle, as a result of which they had more time to anticipate and react to potential hazards (cf. Fuller, 2005; Summala, 1997). Alternatively, drivers only engaged in light speeding when no other traffic was around. Additional analysis is necessary to verify whether such compensatory behaviour indeed took place, for example by evaluating time headway.

Driver age did not show up as significant predictor in the regression model on harsh braking. Possibly, this was because age was significantly correlated with AISS_Intensity, where older drivers showed a lower AISS_Intensity score. Another possibility is that the sample consisted mainly of drivers around 40-50 years old. A replication of the study with younger drivers could address the latter concern. However, personal communication with staff at Shell revealed that the age distribution in UDRIVE was similar to the age distribution of drivers for one of their Dutch hauliers (M = 47.6 years, minimum: 30, maximum: 68) and for one of their Norwegian hauliers (M = 49.0 years, SD = 10.0, minimum: 22, maximum: 65).

Therefore, in the context of MeBeSafe driver age does not appear to be a relevant parameter for the construction of driver profiles.

3.3.3 Limitations

The main limitation of this study is that the number of harsh braking events was too large to perform a manual validation on each of them. The purpose of harsh braking triggers in ND studies is typically to identify (near-) crashes, usually with other road users in front of the subject vehicle. A subset of the resulting events may turn out to be false alarms (i.e., unrelated to potential crashes and actual crashes). In the SHRP2 project with instrumented passenger cars, Hankey (2016) reports a longitudinal acceleration trigger validity of 22% (i.e., the proportion of events that were actual safety critical events), despite using a threshold value (i.c., -6.38m/s^2) that was relatively conservative compared to other ND studies with cars (e.g., -2.94m/s^2 in Dotzauer et al., 2017). In the present study, a large portion of the events has been removed by selecting only events where the brake pedal was used, so the validity of the events is expected to be higher than the 22% reported in SHRP2, where no such filtering was applied. Nonetheless, the analysis could benefit from additional quality assessment, for example by using automated computer vision. Such an approach could also help to distinguish between necessary and unnecessary harsh braking, which has not been part of the present analysis. Moreover, an evaluation of the validity of the events across speed limits will inform the choice between a fixed threshold value and a speed-dependent threshold value.

Another limitation is related to trip distance. Most records in the UDRIVE database covered a short distance, which potentially reduces the generalizability of the findings to long-range truck trips. Previous experimental research suggests that long, monotonous trips increase fatigue and reaction time (Ting et al., 2008). In long trips a relatively large proportion of the distance is typically covered on highways, which, if reaction time is affected, may increase the number of harsh braking manoeuvres. However, all harsh braking events in the present study have been found in unique records, implying that even records covering a large distance have yielded maximally one event. Furthermore, the harsh braking event frequency has been stratified across speed limits, and expressed as proportion of the distance driven. Therefore, a potential bias introduced by trip distance has been mitigated to the best possible extent. In the next chapter collecting data on driving context is further explored.

4 Collecting data on driving behaviour and context based on automated and video-based situation analysis

4.1 Introduction

In the previous chapters, various methods have been discussed on how to collect data on driving behaviour, what thresholds to use and to explore using driving context for driver profiling, based on different data sources. To varying extents, these methods are able to take into account some information on the *situation* in which the driver finds him- or herself when performing a particular type of (safe or unsafe) behaviour. Taking into account the situation (or *context*) is important; one could argue that profiling should ideally consider the driver profile as "the tendency to behave a certain way *in a certain situation or context* ", and distinguish meaningfully between different situations or contexts in which a particular type of behaviour occurs. In other words, for an ideal driver profile it would be useful to distinguish between, for example:

- drivers braking harshly often because they are tend to be distracted and therefore have their eyes off the road;
- drivers braking harshly often because they tend to approach intersections with high velocity;
- drivers braking harshly often because they are prone to close following (tailgating);
- drivers braking harshly often because they fail to do proper left or right shoulder checks when turning a corner and therefore sometimes fail to see bicycles or pedestrians approaching.

Using questionnaires one may ask the driver, for example, to describe whether he/she is sometimes speeding in urban environments and get some idea about situations and context in which undesirable behaviour may occur; but then one has to rely on self-reporting, and one cannot investigate specific situations on a more fine-grained level. Using in-vehicle sensor (IVMS) data one may look at specific, logged occasions where the driver is speeding, match simultaneously logged GPS coordinates to a digital map and investigate in more detail on which roads and at which times such speeding events (or harsh braking events or whatever) occur.

More fine-grained still, naturalistic driving data such as the UDRIVE data allows for investigating the events in more detail —including looking at outward- and inward-facing video data logged simultaneously. In that way, one can get a better "picture" of the "complete situation" in which a particular behaviour (undesirable or not) occurred — as described in the previous chapter.

This chapter follows up on that approach. In the previous chapter, the analysis was necessarily limited because it took too much time to visually inspect (with human eyes) all videos associated with all harsh braking events. In the current chapter, we describe ways of automating that process to a large extent by using AI-based computer vision algorithms and software. Furthermore, we describe how we will use that technology to do further analyses on the UDRIVE data and contribute to the driver and situation profiling task, allowing fine-grained profiles distinguishing specifically between different relevant situations or contexts.

4.2 Contents of this chapter

Neither UDRIVE data nor MeBeSafe-internal data from actual driving by MeBeSafe participating drivers have so far become available for the AI-based computer vision software that we will use in the MeBeSafe project. In the remainder of this chapter, therefore, we focus on:

- Description of the technology we have developed and what can be done with it;
- Experiments and demonstrations on third party data illustrating the principles, as Proof of Concept (PoC);
- Description of the analysis and profiling that will be done once the UDRIVE data and MeBeSafe-internal data from actual driving by participating MeBeSafe drivers become available in the near future.

4.3 Description of technology and PoC demonstrations on third party data

4.3.1 Building a complete picture of the situation: our approach

The goal of our approach is to build up a complete picture of the situation surrounding a particular type of behaviour (such as a harsh brake). For that we need to integrate a variety of data sources. *Figure 4.1* illustrates this. By using in-vehicle (IVMS) sensor data we get information on the GPS location, speed, pedal actions, etc. By using map matching to rich digital map information we know where an event took place, for example in an urban environment, and/or at an intersection, which may have been marked as 'dangerous' due to a previous record of accidents at that place, etc.

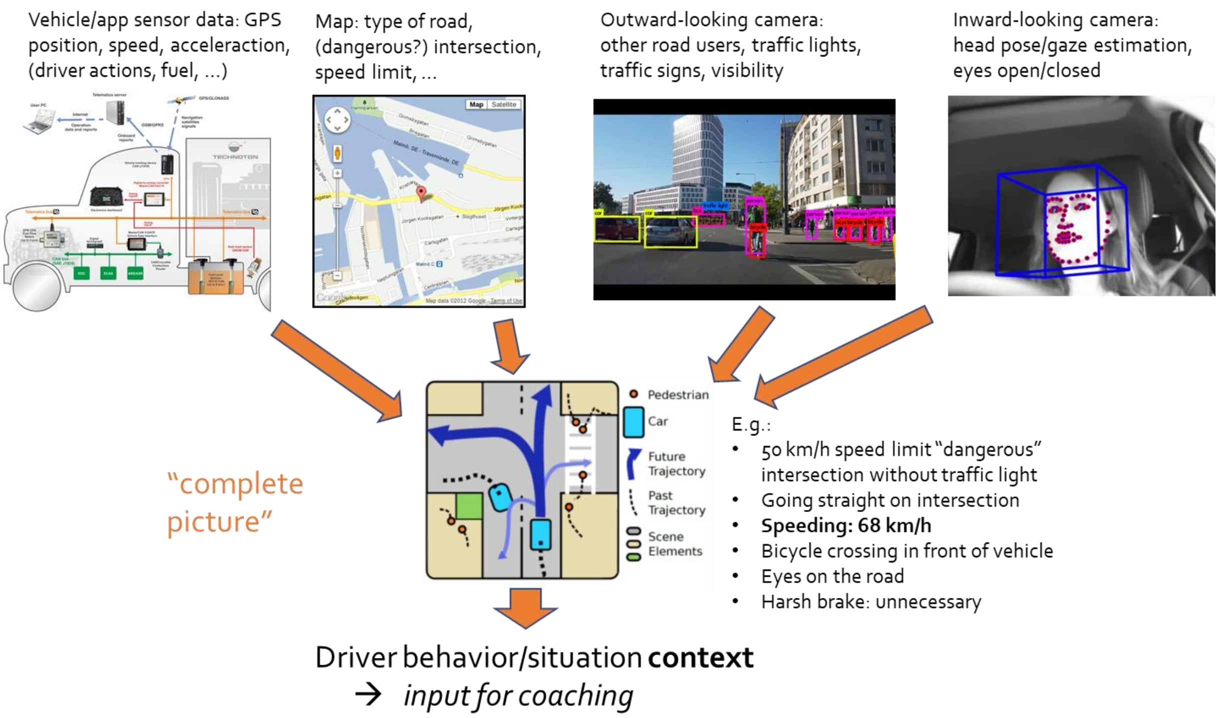


Figure 4.1 Building up a complete picture of the situation surrounding a particular type of behaviour by integrating a variety of data sources; including outward-looking and inward-facing camera data which capture information which cannot otherwise be captured. Video source: YouTube and public UDRIVE data.

However, only by incorporating data from outward-facing cameras (or, to a lesser extent, radar) can we get information on the outside traffic situation: whether there was, for example, a bicycle suddenly crossing in front of the vehicle, requiring the driver to brake.

Furthermore, only by incorporating data from inward-facing cameras (or an eye tracking device) can we get information on the particular driver situation: whether he/she appeared to have his/her eyes on the road (which is indicative of paying attention to the traffic situation), or perhaps appeared to be distracted by a cell phone or something else; and/or whether he/she did a proper right or left shoulder check before making a turn.

In the next sections, we describe our technology for automatic AI-driven analysis of these types of outward and inward-facing camera data.

4.3.2 Off-the-shelf commercial technology versus in-house, open technology

It should be noted that some of the functionality described here could be acquired off-the-shelf, using commercially available systems. For example, MobilEye devices can be acquired and used as outward-facing cameras, giving information on detected road users. Seeing Machines sells smart cameras for the automotive industries which can be used as inward-facing cameras, giving information on head pose and eye gaze and eye closure. However, there are multiple reasons to develop and use our own, in-house developed technology:

- **Cost**: Commercially available systems are very expensive, especially if the licensing conditions must allow the type of large scale use and analysis capabilities we require in the MeBeSafe project
- **Closed systems vs. open systems**: Commercially available systems are virtually always very closed, in that they do not allow deep access to the type of 'raw', semi-processed data which we in the MeBeSafe project require for our analyses. For example, a Mobil Eye device can output a certain amount of object data with position information, but not more underlying and more fine-grained data (including, for example, more specific image and other data available for individual road users which we intend to exploit in our analysis and profiling).

- **Flexibility**: In the MeBeSafe project, we will carry out custom server-based analysis using the AI technology on naturalistic driving (UDRIVE) data, using custom interfaces to the databases and custom outputs. Furthermore, we intend to use the same technology on data (including outward- and inward-facing video) acquired for (a subset of) MeBeSafe drivers, coupling to coaching app software using additional custom interfaces. Commercially available systems are designed for one or two specific, pre-designed use cases, and do not legally and technically allow for this type of flexibility. For example, a MobilEye device or Seeing Machines device can be used in a vehicle, but cannot be adapted to do server-based analyses.
- **Adaptability**: In the MeBeSafe project, we adapt core technology (e.g. for object detection, tracking, gaze and eye closure detection, etc.) for our particular purposes, for example for a specific type of trajectory analysis or for a specific type of eyes-off-target analysis; requiring deep access to the code. This type of access and adaptability is not available for commercially available systems.
- **State of the art**: Artificial Intelligence (AI) and in particular deep-learning based vision algorithms have been and are still making tremendous progress, each year, in terms of general capabilities, accuracy, speed of processing, and robustness. The same holds for the software and hardware used to implement those algorithms. Commercially available systems, due to their production, testing and sales cycles, necessarily lag behind compared to state of the art. Within MeBeSafe, due to its research and innovation nature and the available expertise within the consortium, we are able to use and exploit more recent innovations.

4.3.3 Outward-facing camera-based road user detection

For a much more automatic (compared to using human eyes) analysis and profiling using outward-facing camera video data as additional situation information, we use an approach based on modern AI-driven computer vision. In particular, we use modern deep learning-based object detection algorithms (cf. Sermanet et al., 2014; Girshick, Donahue, Darrel, & Malik, 2014; Girshick, 2015; Ren, He, Girshick, & Sun, 2017; Redmon, Divvala, Girshick, & Farhadi, 2016; Redmon & Farhadi, 2017) which are able to detect, in images, *what* type of objects are in the image, and *where* in the image they are located. In *Figure 4.2*, for example, different object classes (road user types as well as other relevant objects such as traffic lights) are identified by different colours, labels indicating the class, and bounding boxes indicating the size and position in the image.

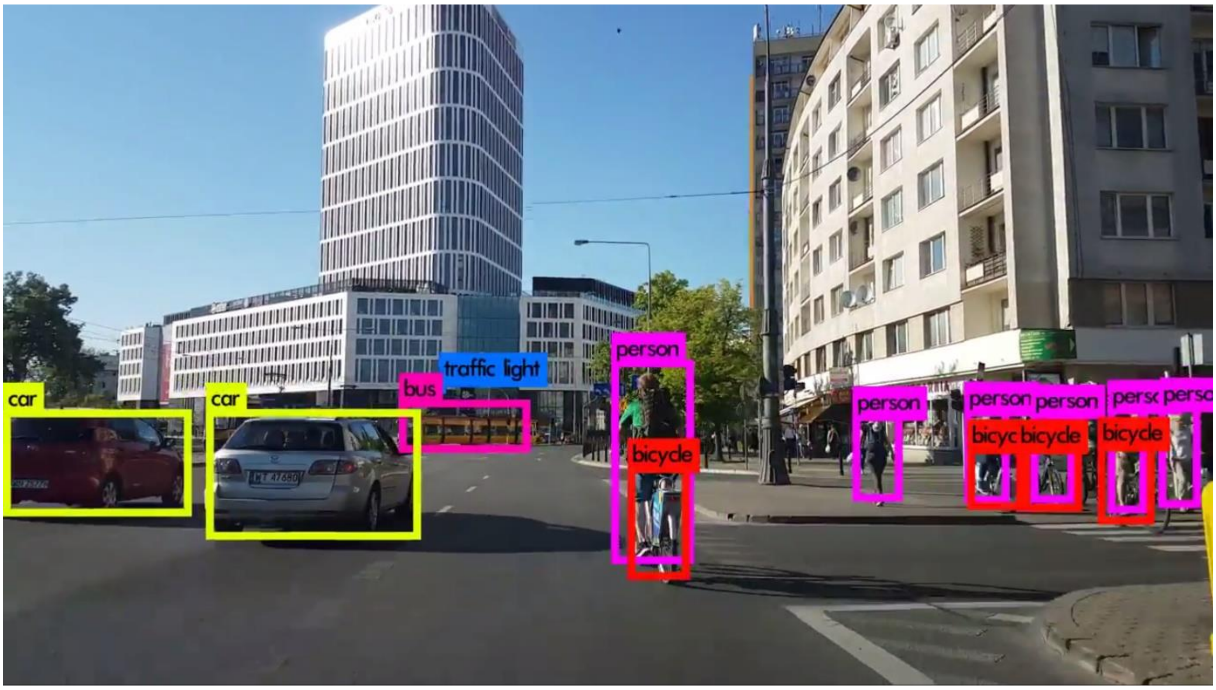


Figure 4.2 Deep learning-based road user detection in an image. Note the accurate identification of road user class (such as car vs. bicycle vs. person), and the accurate localization in the image. Video source: YouTube.

Specifically, we use YOLOv2 (cf. Redmon & Farhadi, 2017). YOLO ("You Only Look Once", Redmon et al., 2016) and YOLOv2 (Redmon & Farhadi, 2017) distinguish themselves from most other approaches by *not* having a separate, serialized process in which candidate object "region proposals" are determined and examined sequentially (a time-consuming process). Instead, they use a single regression process (hence, You Only Look Once) in which possible object regions and classifications and bounding boxes are predicted and learned in one sweep, in parallel. This makes it a very fast during model application ("inference"), while at the same time maintaining high accuracy and robustness. Experiments on third party data with our implementation confirm this.

We use a fine-tuned and adapted version of the algorithm that is particularly suitable for *road user* detection and tracking under challenging visibility conditions. That is, we use particular training data and categorization (MS COCO, Lin et al. 2014, with certain categories discarded or merged for our outside traffic estimation application); and we have adapted the method to include a novel technique of extracting data from the deep network which is very suitable for tracking (see below).

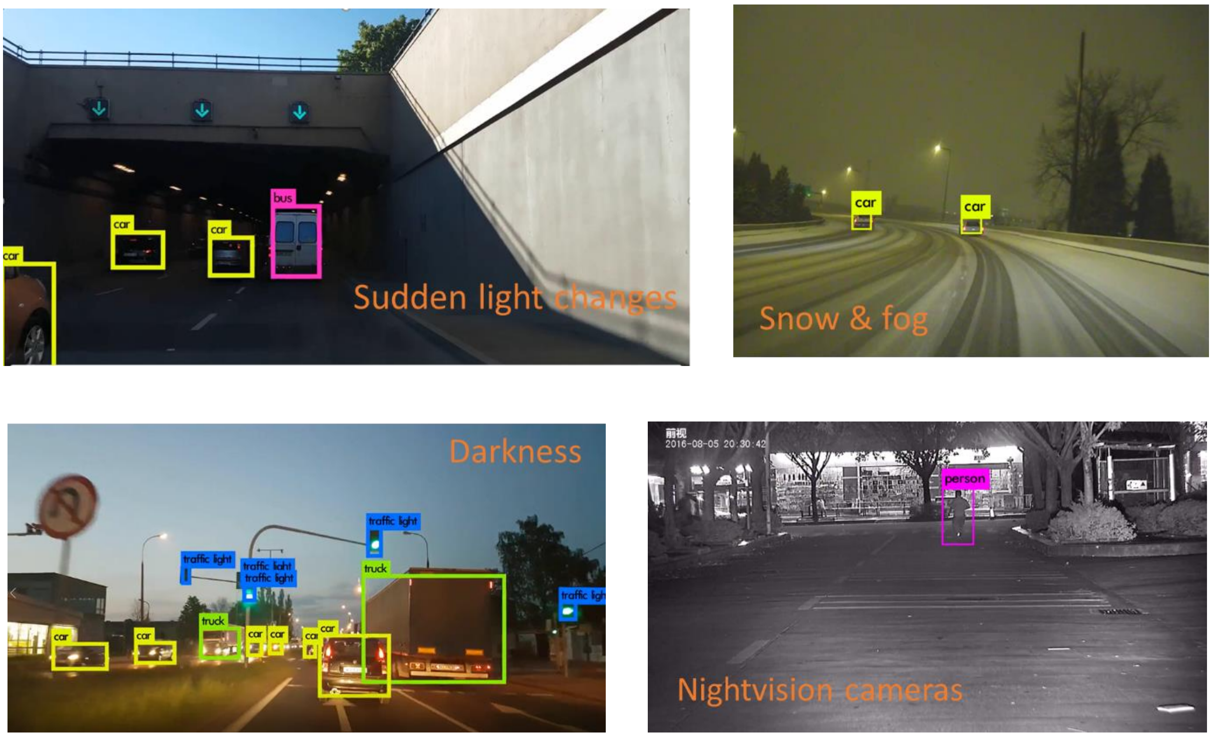


Figure 4.3 Illustrations of how our computer vision object detection algorithm is robust even with sudden light changes (going from a bright outdoor environment into a tunnel), under bad weather conditions which significantly impair visibility (snow and fog), at night when it is dark and vehicles use possibly bright headlights, and even with the type of low-resolution black and white images that can be obtained using nightvision cameras and which may be used in very dark situations. Video source: YouTube.

We have done extensive testing to test accuracy and robustness under various challenging light, weather, day and night and camera type conditions; to evaluate the suitability for further use on UDRIVE data and MeBeSafe data. *Figure 4.3* illustrates robustness under such difficult conditions.

4.3.4 Tracking road users detected outside the vehicle

Tracking is the process of following the individual detected road users *over time* (their *trajectories*); ideally even when there are large appearance changes and occlusions which make such tracking difficult. In particular, we need *multi-object* (or multi-target) tracking (e.g. Reid, 1979; Leal-Taixe, Milan, Reid, Roth, & Schindler, 2015). This process is done such that a spatio-temporal representation can be obtained of the trajectories of the multiple other road users around our driver—which describes the traffic situation around our driver. In other words, without such trajectories obtained by tracking, we cannot properly identify what is really going on around our driver, whether certain other road users are on a potential collision course with our driver, etc.

To do tracking in our application, we use a method inspired by a relatively new but existing method (Wojke, Bewly, & Paulus, 2017), but adapted now for the first time (as far as we know) to our particular object detection algorithm. This method is a so-called "tracking by detection" method: that is, a tracking method which relies heavily on accurate detections (which we have) from a state of the art object detection algorithm; and so-called "*appearance feature vectors*" derived from those detections.

Those appearance feature vectors act as a kind of "fingerprint". They should remain relatively constant for subsequent detections of the *same* object over time in *different* image frames (e.g. the same red Fiat); while distinguishing clearly between *different* instances of the *same object class* (e.g. a red Fiat and a grey Toyota). In this way, by storing past appearance vectors for different "active" tracks, and comparing newly detected appearance vectors to stored vectors, we can solve the so-called association problem and assign new detections to existing active (or new) tracks.

This can work even when objects temporarily disappear from view either partially or fully; or when their appearance changes drastically, for example when a car or bicycle comes closer or makes a turn, or a pedestrian changes its body pose. This is possible because the storage of past appearance vectors acts like a short-term memory (allowing the system to pick up a previously lost track when an object reappears with a familiar appearance vector); and new appearance vectors belonging to a single track (single object instance) can be added which reflect appearance changes over multiple image frames.

For example, in our application of tracking multiple road users, our object detection algorithms may detect multiple instances of cars in one image, as well as multiple cyclists, and multiple pedestrians. For each of the cars, an appearance vector is extracted (process to be described shortly). This appearance vector captures certain information which allows the system to distinguish one particular car detection (e.g. a red Fiat) from another (e.g. a grey Toyota). Over time, when new image frames are detected and analysed, multiple car detections and their appearance vectors are again analysed and if the red Fiat and grey Toyota appear again, they can be matched ("tracked") against their previous detections and locations. New car instances can appear as well, leading to new tracks; and old tracks can disappear after sufficient time (e.g. if the red Fiat is not seen for a substantial amount of time).

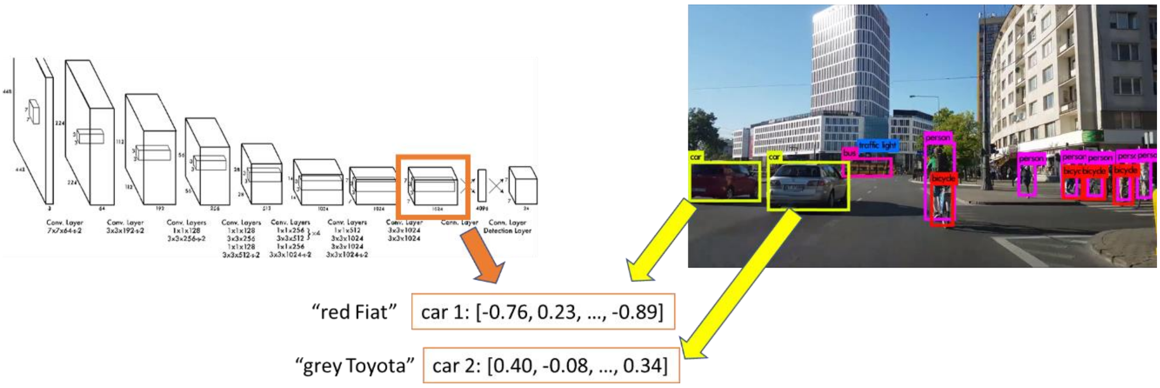


Figure 4.4 Schematic illustration of how our method extracts uniquely identifying "appearance feature vectors" associated with different instances of detections of the same object class "car" (right picture), by taking the activation vector of a deep layer in the deep neural network (left picture) which constitutes the object detection system. This allows other modules in the system to track, over time, those (and other) multiple targets.

Figure 4.4 illustrates the process of extracting appearance feature vectors from detections in images. In our case, our object detection algorithm (YOLOv2) allows for a very elegant, powerful, and fast production process of appearance feature vectors; by simply extracting the 1024-size activation vector of a deep layer (layer 29, to be precise) of the network corresponding to the object grid cell. This can be done in the same, single ("only look once") forward sweep. *Figure 4.4* illustrates this process as well.

Importantly, many previous methods have had difficulties (in terms of making it work well, or making it work fast enough) with this type of approach with deep learning systems. This is due to the fact that it is necessary but difficult to strike a balance in the appearance feature vector production process between, on the one hand, using 'higher-level', more semantic, categorical information that is relatively invariant to appearance changes; and, on the other hand, using 'lower-level' appearance information which identifies the particular instance of an object and distinguishes it from other objects of the same class ("red Fiat vs. grey Toyota") and which allows for precise localization (Danelljan, Häger, Khan, & Felsberg, 2015; Ma et al., 2015; Want, Ouyang, Wang, & Lu, 2015). Using YOLOv2 we are able to solve this dilemma, and find a good balance between the two, because in certain deeper ('higher-level') layers (including the layer 29 that we extract from) it already uses, in a similar vein to ResNet systems, short-cut connections from shallower ('lower-level') layers, and those deeper layers are in fact *trained* to find such a good balance.

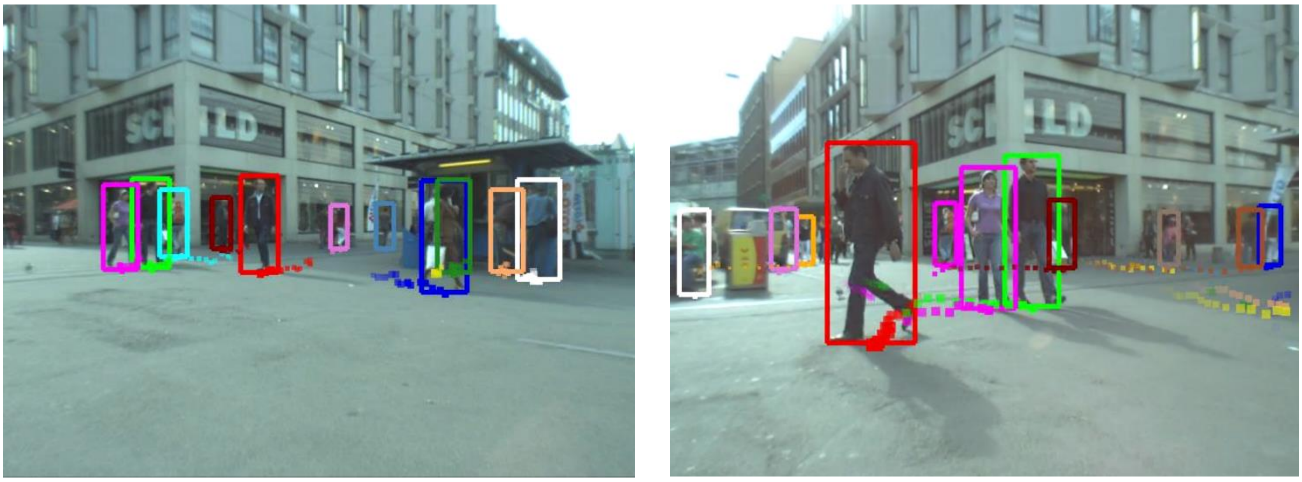


Figure 4.5 Illustration of tracking by our method of pedestrians. The left picture shows the situation and tracking at a certain time T, the right picture shows the situation, after both object motion and ego-motion, at a later time T+N.

Figure 4.5 illustrates the tracking performance of the combined detection and tracking system, using a third-party dataset (see Leal-Taixe et al., 2015) which focuses on tracking many instances of a single class of road user, namely pedestrians. In this figure, colours of the bounding boxes now indicate distinct *instances* of the object (pedestrian) being tracked. Note also the "trails" or "tails" of the same colour, connected to the boxes, which indicate the history of previous positions of the same object in the track. It can be observed (for example, for the man in dark clothes indicated by red bounding boxes, or the other man indicated by green boxes) that individual persons are tracked successfully over time, even under significant appearance changes and temporary occlusions, which are caused both by object motion (the pedestrians moving) and ego-motion (the camera moving).

4.3.5 Transform to real-world 2D/3D space trajectories, identification of traffic patterns and situations

For further trajectory analysis, identification and classification of traffic patterns and situations, the detections and tracks determined in the image space of the outward-facing video data is "transformed" to the real world 2D/3D space or domain, by using camera position and relative position and size of road users within the images. This process is illustrated in *Figure 4.6*. In a way, this produces a kind of "bird's eye view" of the situation.

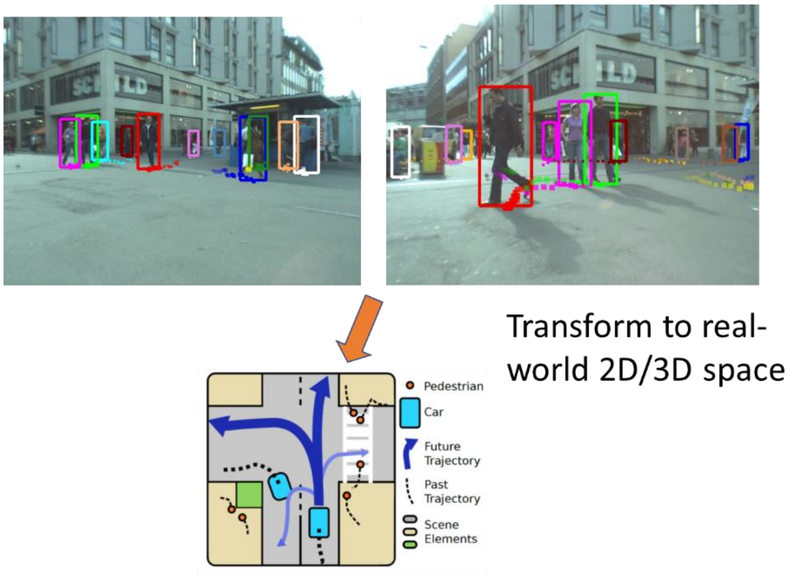


Figure 4.6 Illustration of the process of transforming detections and tracks obtained in the image domain to trajectories in a real-world 2D/3D representation which are meaningful with respect to a digital map.

In this process, we integrate with GPS and map data; so that we know where in the map (e.g. at what intersection) we are, and we obtain spatiotemporal trajectories that are mapped on the digital map. Within that domain we identify patterns, classes of situations; e.g. close following followed by harsh brake; left turn on an intersection with bicycle approaching followed by harsh brake; etc.

4.3.6 Analysing inward-facing camera data to complement the context

The situation or context analysis is made more complete by including information from inward-facing cameras. As with the outward-facing video data, we use modern AI-driven computer vision (Hansen & Ji, 2011; Czupryski & Strupczewski, 2014; Gudi, Tasli, den Uyl, & Maroulis, 2015; Baltrusaitis, Robinson, & Morency, 2013) to allow for a more automated analysis (and profiling) process than what is possible when having to rely on human eyes (as was discussed in the previous chapter on the analysis of UDRIVE data).

This technology allows us to automatically estimate, from video images showing the face of the driver, head pose and gaze direction, as well as “eyes open” versus “eyes closed”. Head pose and gaze estimation allows us to estimate, for each point in time for which there is driver video data, if the driver is looking at the road or not, and if so, provide an estimate of where he or she is looking. It also allows us to determine automatically whether the driver performed right or left shoulder checks before making a turning or overtaking manoeuvre.

Eyes open versus eyes closed detection provides another way to detect whether the driver is currently looking at the road and outside traffic or not; and importantly, by analysing eye closure duration and blink rate we are able to estimate with a reasonable degree of accuracy drowsiness and microsleeps, which are important causes of traffic accidents (Ng Boyle, Tippin, Paul & Rizzo, 2008; Horne & Reyner, 2001; Häkkänen, Summala, Partinen, Tiihonen & Silvo, 1999).

4.3.7 Facial landmark and action unit analysis technology

Specifically, we use deep learning-based facial landmark and action unit analysis algorithms (Gudi et al., 2015; Baltrusaitis et al., 2013; Jiang, Valstar, & Pantic, 2011) to determine the head pose, gaze, and eye variables described above. Despite its sophistication, the entire process takes computation time in the order of milliseconds on modern computers, allowing for fast and if necessary real-time processing. *Figure 4.7* illustrates the results of the analysis process. The technology works by first detecting the presence (or not) and location of a face within an image, in a manner quite similar to the object detection technology described in *Section 4.3.3*.

Next, facial “landmarks” are analysed based on image features and known face features, placing the eyes, nose, mouth, etc. on the correct positions; effectively overlaying the image with a flexible ‘mask’ describing the face in numerical terms. This allows for head pose orientation estimates and initial gaze estimates.

Figure 4.7 Illustration of how facial landmarks, effectively making up a 'mask', are placed on an image, allowing head pose and expression (action unit) estimates, and how further analysis of the eyes allows for refinement of gaze estimation. Video source: Cygnify Webcam.

Further analysis of the eyes in the image allows for more precise gaze estimates. "Action unit" analysis relies on analysis of relative positions and movements of detected facial landmarks to detect, for example, smiling, frowning, and also eye closures (blinks).

4.3.8 Challenging and real-world driver face video data

Figure 4.8 to *Figure 4.10* illustrate our experiments to analyse the usability for real-world driver data and robustness under challenging conditions.

Figure 4.8 illustrates robustness of the analysis on real-world driver video under difficult light conditions (large light variations, even within one video image, which is traditionally problematic for computer vision) and large head position and head pose variations. *Figure 4.9* shows robustness when sun glasses are worn; gaze detection cannot be as precise, but very reasonable estimates can still made based on landmark analysis and head pose. *Figure 4.10* shows the analysing ability of night vision video, which is useful in particular when analysing video images collected at night, when normal cameras do not have enough light to work with and inward-facing cameras based on infrared imaging can and are typically used. Situations of night time driving are particularly relevant for drowsiness and microsleep detection.

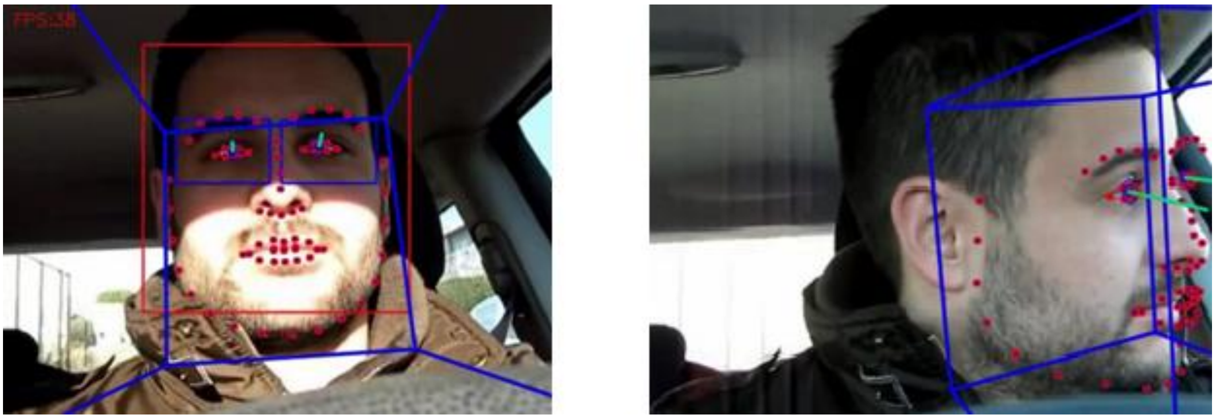


Figure 4.8 Illustration of robustness under challenging light conditions and large head position and pose variations. Video source: YouTube.

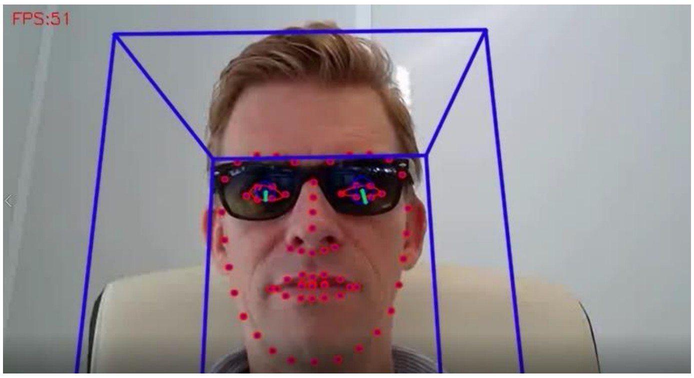


Figure 4.9 Illustration of robustness when sunglasses are worn. Video source: Cygnify webcam.

Figure 4.10 Illustration of the ability to work with night vision cameras. Video source: YouTube.

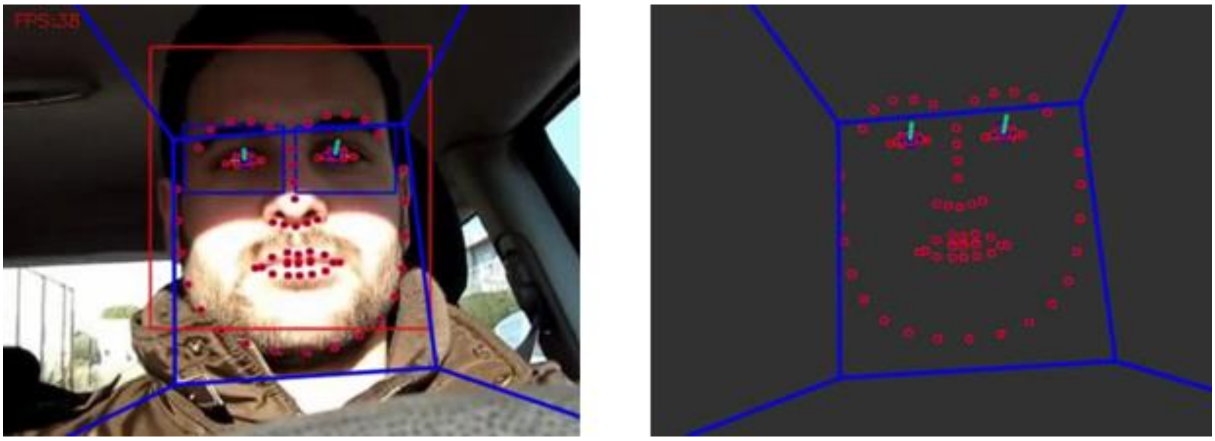


Figure 4.11 Alleviating privacy concerns to some extent: after processing of the image (left), one may discard the actual image data itself and only use, and store for later, the numerical facial landmark, gaze, and action unit data (right). Video source: YouTube.

4.3.9 Privacy

Figure 4.11 illustrates a way of alleviating privacy concerns to some extent. Understandably, many drivers are hesitant to have an inward-facing camera recording them all the time. This is a genuine concern and one that cannot be solved completely, by technology or otherwise.

However, technology may help to some extent. Unlike inward-facing camera systems where human operators (such as supervisors) inspect recorded videos, the technology allows for a set-up where immediately after processing the camera image in the vehicle, the video images are discarded, with only the numerical facial landmark, gaze, and action unit data remaining.

Furthermore, access to the video data could be limited (barring some exceptions, when there have been multiple dangerous situations, for example) to only the drivers themselves and to researchers or coaches.

In set-ups where post-hoc analysis and coaching is not even necessary but the aim is rather to provide real-time warnings for distraction and drowsiness, even those numerical data could be discarded only seconds after capturing.

4.3.10 Bringing it all together

We refer back to *Figure 4.1* to illustrate how the data from the various pieces of technology are brought together to obtain a more or less "complete picture" of the situation or context in which a particular type of behaviour occurs.

Using the outward-facing technology, we are able to estimate automatically trajectories of other road users in view of the driver. By integrating the trajectories of other road users with map information, information on GPS position and speed of the vehicle from IVMS sensors, we are able to place those trajectories on a particular location (e.g. around an intersection) relative to our driver, estimate potential collision courses, close following, etc.; i.e., a 'higher-level' understanding of the outside driving context. Furthermore, by integrating gaze estimation and eye closure information we are able to add even more context, namely on whether our driver has his or her eyes on the road, is possibly distracted, does a shoulder check, or is currently drowsy.

All together this offers very rich information about what is going on at a particular moment when, for example, a harsh braking event occurs, information which can potentially make feedback and coaching more focused than when such a rich analysis is not performed. It will be necessary, to "summarize" this rich information in such a way that coaching and feedback can easily make use of it (see the next sections in this chapter, as well as *Chapter 7*).

4.4 Application to naturalistic driving data and MeBeSafe: towards fine-grained automatic driver/situation profiles

4.4.1 Goals

The technology described in *Section 4.3* will be applied in this project to naturalistic driving data (specifically the UDRIVE data, see *Chapter 3*, and the MeBeSafe data on driving behaviour to be collected during the project), with the goals of:

- automating the analysis of that data to a larger extent than what has already been done (for UDRIVE, which relied on human analysis of video) and what otherwise can be done when using video within MeBeSafe;
- analysing that large amount of naturalistic data to determine typical patterns (classes) of situations which should be discerned for coaching in MeBeSafe;
- determining the added value of using outward- and inward-facing video data for situation analysis and coaching;
- helping to determine thresholds and trigger values for available input data (IVMS and video-based) which lead to fine-grained, numerical driver profile data able to distinguish meaningfully between different classes of situations and driving behaviours, and which provides useful input for the coaching scheme task of task 4.3 of Work Package 4.

4.4.2 UDRIVE data processing: hardware, software, algorithms, and processing toward trigger-ready numerical variables

Once the UDRIVE data is available locally at MeBeSafe partner SWOV, we will place a dedicated High Performance Computing (HPC) GPU-server there, locally, which is suitable for the type of deep learning-based algorithms described earlier. By placing the server locally with the database, we will both ensure the highest processing speed possible (keep in mind we are talking about a very large number of hours of driving data and video) and ensure that we respect privacy and security regulations (which require careful handling of sensitive data of drivers, including videos of drivers showing their faces etc., essentially ruling out cloud-based approaches).

Customized software interfaces will be built connecting the deep learning-based video analysis software to the database. Then, large portions of the UDRIVE video database will be processed using this software. The results will be further processed resulting in such a format that data on specific variables can be stored in the database together with the other numerical data – similarly as other numerical variables which are stored per timestamp. Like other variables they can be used as the basis for so-called trigger thresholds for identifying particular types of events, as was described and used in the previous chapter.

This means that subsequently one could, for example, query the database for events like this: "find all events where the drivers had their eyes off the road for more than 3 seconds", or "find all events where the driver performed a left shoulder check", etc. This will allow for additional questions to be answered, and facilitate and speed up analyses of events or trips where the relevant situation can only be derived from video. For example, one can then complete analyses where we look at all harsh braking events, and ask straightforwardly how many of those were preceded by the numerical variable stating "eyes off the road" or look at all turn manoeuvres at intersections, and ask how many of those were accompanied by an appropriate shoulder check.

However, since the automatic analysis process is not perfect, sanity checks with human eyes will still be necessary occasionally; but the automatic process can significantly facilitate and speed up naturalistic driving (including UDRIVE) analyses.

4.4.3 Identifying typical situations and alignment with the KPI variables

The approach described earlier can then be used to identify typical situations, i.e. general classes of situations or contexts which should be distinguished for our context-sensitive driver profiling. In other words, we want context-sensitive driver profiling to be fine-grained, but not too fine-grained: every particular traffic and driving scenario or event is unique in its details, but we want to summarize the behaviour and situations in such a way that we end up with a compact yet meaningful set of *profile output variables*.

For this we aim to bring this in line with, and stay close to, the *Key Performance Indicator (KPI) variables* identified within WP4:

- Harsh braking
- Harsh cornering

- Speeding
- Close following
- Distraction
- Drowsiness/Fatigue
- Lane departures

The first three can be derived from the in-vehicle sensor (IVMS) data on GPS positions, speeds and accelerations available from IVMS systems and smartphones (and storage thereof in the UDRIVE database and MeBeSafe databases). This was discussed already in *Chapter 2*. Close following, distraction, drowsiness/fatigue, and lane departures however, require camera data and the specific technology described in this chapter.

4.4.4 Close following

In the UDRIVE database, data from MobilEye devices is already available, which allows analysis on close following. However, we will add our own object and relative position data derived from our outward-facing video analysis software, from which close following (among other things) can be derived, because:

- it will allow for richer road user (and other relevant traffic situation) data to be extracted;
- it will have a more direct and neater relationship with the stored video data (note that the MobilEye camera was actually independent from the UDRIVE video capture cameras and had a different mounting position, viewing range and viewing angle);
- we aim to use the same video processing technology for analysis of the UDRIVE data as well as for analysis and profiling of MeBeSafe drivers participating in the study and equipped with cameras (but who will definitely not have MobilEye devices on board for our use).

Using the UDRIVE data, we will calibrate the relative positions data and close following trigger parameters. That is, we will determine which values of the relative position variables represent useful criterion values: values beyond which we should classify driver behaviour as "close following" for the purpose of MeBeSafe profiling and coaching. These will correspond to the trigger values to extract close following events from the UDRIVE data, and similarly function as trigger values to detect and log close following events in MeBeSafe participant driving, for those vehicles equipped with the necessary technology.

4.4.5 Distraction and drowsiness/fatigue

Distraction and drowsiness/fatigue are also KPI variables which require cameras, specifically inward-facing cameras in order to collect data on these behaviours. Distraction per se is actually only *one* of the various situations in which the driver has his or her eyes off the road and off relevant other road users that he or she should be paying attention to. Thus, it is better to adapt this KPI variable to what we may call the "**Eyes off target**" variable.

The "Eyes off target" KPI variable can then have two *subcategories*, each of which can be detected using our automated technology:

- o **Distraction** (which we operationalise as detecting that eyes are directed at a smartphone, radio, work papers, colleague; which is an imperfect operationalisation, but distraction due to, for instance, "daydreaming" cannot be detected easily);
- o **Failure to look properly** (eyes on the road but, for example, failure to do right shoulder check before right turn);

Drowsiness/microsleep is a separate main KPI category or variable, and will be operationalized by measuring when the blink rate is very high or the eyes are closed for a long duration.

4.4.6 Lane departures

Lane departures are a final KPI variable for which again cameras are necessary to determine them. The previous sections have not described the necessary technology for that, but we will similarly use AI-based computer vision technology to determine lanes and lane departures; algorithms exist which can perform this task.

4.5 A driver profiling approach for MeBeSafe that takes into account situations in a flexible way

4.5.1 Core profile output variables, computed as relative values to baselines

What we end up with as the proposed profiling approach, when making use of the available technology as described in this and the previous chapters, and which may vary somewhat between different MeBeSafe driver participants (it is likely that only a limited subset of the vehicles will be equipped with cameras), is the following. The MeBeSafe driver profiling approach will consist of profile output variables which directly link to KPI variables.

Numerical values for each of these seven “dimensions” or “categories” of the profile should be determined in a *relative* way, by relating the number of events as detected by trigger values to either the time that good behaviour was exhibited (e.g. speeding vs. no speeding) or to the average behaviour exhibited by the peer group (e.g. more close following or harsh braking than the average of the peer group), or both. In other words, the values are computed by relating individual driver behaviour and events to baselines.

4.5.2 Optional variables and refined variables or subcategories

Several of these variables (close following, eyes off target, drowsiness/fatigue, lane departures) are only possible to investigate with camera equipment (which may not be present on all or even most vehicles participating in the MeBeSafe coaching). Hence they must be *optional* variables, which may be omitted while still allowing for a meaningful profile, consisting of harsh braking, harsh cornering, and speeding.

At the other end of the spectrum, if more advanced (camera and video processing) technology is available, we can add more *refinement* to the seven main profile output variables; effectively adding *subcategories* to the seven main categories. For the eyes off target variable we already discussed this in a previous section (*Section 4.4.5*) with the subcategories consisting of distraction and failure to look properly.

For the other main categories this can be done as well, sometimes with and sometimes without the help of cameras and associated technology. For example, we can distinguish between speeding on urban roads vs. speeding on highways or provincial roads (cf. *Chapter 3*)). This does not require cameras but requires map matching of IVMS GPS data. More sophisticated and requiring the use of cameras, we can distinguish between harsh braking which can clearly be attributed to another road user violating traffic rules, in which case the harsh braking event is, in a way, "necessary" and "good"; and harsh braking which has no such outside cause, and which therefore may be interpreted as "unnecessary" and "bad". This would effectively split the single "harsh braking" variable or category into two.

Similarly, we can distinguish between harsh braking (or harsh cornering) when the traffic situation is "difficult" versus harsh braking (or harsh cornering) when the situation is not so difficult. Whether the traffic situation is difficult can be estimated (imperfectly, but to a reasonable degree) by counting the number of (various) other road users around the driver and the complexity of their trajectories, combined with road and intersection information. More coarsely this can be done without any camera data, by just using map data on roads (urban vs. highway) and intersections, rush hour time information, and by including mapped data on dangerous (high accident rate) map locations, which is available for some regions.

In this way, we have a flexible, open-ended, yet meaningful and easily interpretable driver profiling approach, which can work with the most advanced basic hardware and software set-up (including outward- and inward-facing cameras and their associated technology, as described in this chapter) as well as a more basic hardware and software set-up (cameras omitted, which is likely to be the case for many of our MeBeSafe driver participants); and in which we can choose to make refinements in an iterative, incremental way.

4.5.3 Concurrency analysis and indication of relationships between variables in the profile data to provide context information

The remarks concerning optional subcategories and refinement lead directly to the insight that even when we choose to work with these five core KPI and profile output variables, the variables are not (statistically and causally) independent but in fact are often correlated. It is very useful and meaningful to identify these relationships where possible, in particular for coaching purposes. Thus, it should be possible to express these relationships in the profiles, if one desires to contextualize the profile data as much as possible.

For example, it is possible (and even likely) that in situations (events) when there is close following, there is also (shortly thereafter) harsh braking, because if the car in front of you that you are tailgating is suddenly braking, you would brake harshly to prevent a collision. If this type of situation occurs frequently for a certain driver, this information should ideally become apparent and explicit in the profiling data which is used in coaching; and not just as having independently high values on both harsh braking and on close following, which may appear independent and coincidental at first sight.

This type of relationship could be determined on the level of events and can be analysed; where we can see that close following occurs together with a harsh braking event. That is, we can look at *concurrency* at the event level (which was called "co-occurrence" in the previous chapter and analysed there to some extent already). If there is such high concurrency between two variables at the event level, additional information should be included in the profile data which indicates that concurrency relationship (which may well indicate a causal relationship) between those two of the five core profile output variables (in our example: harsh braking and close following).

This type of analysis is especially relevant given the advanced situation analysis and context-sensitive profile information we intend to generate using the technology described in this chapter. After all, it is exactly that type of richer context information that we wish to add using the technology. Thus, we will do this type of concurrency analysis extensively on the UDRIVE data, using our technology, to determine realistic and useful trigger and concurrency correlation thresholds; and apply these during the driver pilot testing and coaching period.

In the next chapter the added value of measuring driver competences in addition to driver behaviour and driving context is explored.

5 Measuring driver competences

When it comes to the coaching of drivers, it can be useful to include a driver competence model to move beyond the limitations of merely modifying overt behaviour. A competence approach views behaviour as being caused by a driver's underlying competences, such as skills or knowledge, and shifts the focus from pure behavioural modifications to the improvement of less developed competences. This approach is commonly used in education where a teacher assesses student's competences in a specific field, measures them, and compares them to a norm model that allows to decide whether the student is at a eligible level or whether educational interventions are needed. Such an approach becomes necessary when the behaviour of interest is beyond a certain level of complexity: A student who cannot yet add or subtract numbers can make an infinite number of mistakes and it would be useless to just modify the behaviour, for example by rehearsing the specific results of calculations (e.g. 33 + 42 = 75). A more successful approach is to coach the student on the rules of addition. In this chapter we consider safe driving behaviour in a similar way by targeting "safe driving competences". Such an approach should widen the applicability of driver coaching targeted at behavioural modification and should increase acceptability of online coaching approaches. Similar trends have already occurred in the field of electronic tutoring (e.g. Albert & Schrepp, 1999).

The aim of this chapter is to describe a driver competence model and to offer suggestions on how these competences could be measured. The described work is intended to give directions for future research and development in order to assess "driver competences" in a more holistic way.

MeBeSafe deliverable 1.1 (Karlsson et al., 2017) describes in considerable detail the theoretical considerations of driver characteristics and an integration framework for nudging and coaching. This deliverable intends to build on deliverable 1.1 and to extend it in relevant aspects. We aim to focus on the practical aspects – what are safe driving competences? And how can they be measured?

5.1 Definition of Competence

This section aims to give a clear definition of the term competence in order to establish a common ground in the understanding of this term.

Competences allow an individual to master variable situations successfully and responsibly and can be seen as fundamentals for learning. Referring to Reber (1995), Stanton, Walker, Young, Kazi and Salmon (2007) see competence as a "collection of knowledge, skills and attitude", which results in "the ability to perform some task or to accomplish something" (p. 1210). Lindstrom-Forneri, Tuokko, Garrett and Molnar (2010) make a clear distinction between competence and performance. The authors define competence as "a latent construct that refers to what a driver is capable of given the dynamic individual-environment interaction." (p. 284) whereas driving performance "refers to the actual driving behaviours" (p. 284). Based on these definitions, two aspects are important: 1) competence can be seen as a triad of attitude, knowledge and skills and 2) competence and performance are two distinct constructs. These aspects are crucial for our formulation of a practical driver competence model, which will be described in the next section.

5.2 Driver Competence Model

This section will define a driver competence model which describes the relevant competences for successfully accomplishing the driving task. This section builds on the theoretical aspects of driver competences described in deliverable 1.1 and aims to extend it in relevant aspects.

5.2.1 Definition of driver task

The driver task has been described in detail by McKnight and Adams (1970): the authors identified around 1700 different tasks that a driver needs to complete. Michon (1985) summarized the driving task in three different levels: strategical, manoeuvring and operational level respectively. We want to add the description of Hollnagel, Nabo and Lau (2003), who summarized the driving task as "comprising of planning the drive, monitoring one's own car and other traffic, and controlling speed and direction (comprising steering, accelerating, braking)" (p. 87). The authors propose four different levels of control during the driving task: tracking, regulating, monitoring and targeting (see *Figure 5.1*).

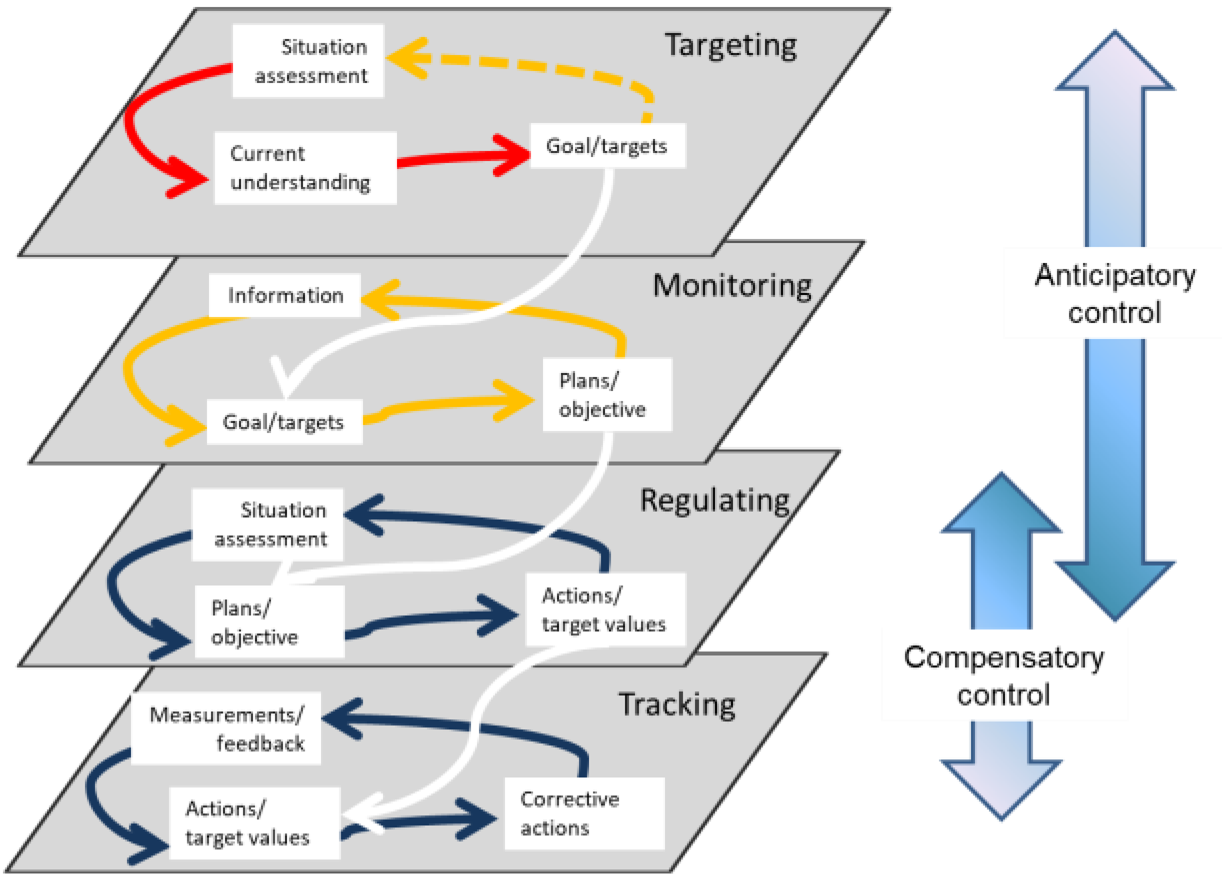


Figure 5.1 Driver in Control Model (Hollnagel & Nabo, 2003).

The first basic level of control is tracking, where the driver is required to maintain speed, distance from the car in front or behind and things alike. Hollnagel et al. (2003) describe tracking as a closed-loop control strategy, which, for experienced drivers, can be accomplished with little effort and without conscious processing. The level of regulating forms the goals and target values which will be forwarded to the tracking level. On this level, target speed, specific position and movement relative to other traffic elements are of interest. On the level of monitoring, the interaction between driver-car-environment (traffic situation, potential hazards, etc.) is constantly assessed. On this level, plans and objectives used by the subsequent levels are generated. The status is monitored constantly (car condition, fuel consumption, etc.), as well as are environmental cues, such as traffic signs, direction indicators, warnings, etc. The level of targeting is similar to Michon's (1985) strategical level. It is concerned with higher-level goals (target destination, for example) and is responsible for an infrequent monitoring-process of the higher-level goals which are stored in and retrieved from memory (as for example: recognising that one has driven in the wrong direction).

5.2.2 Formulation of a driver competence model

Derived from existing literature, we propose a basic driver competence model which concentrates on the interaction of the driver with the environment and the ability to adapt to the context, as this is the most prevalent aspect of safe driving behaviour.

Within this regard, we want to highlight the concept of Situational Awareness (Endsley, 1995), which describes three essential steps in order to perform tasks in a dynamic environment: 1) to perceive and select the relevant elements in the environment (what kind of intersection? How many cars are around me? What aspects of the environment – weather, traffic etc. – do I need to pay attention to?), 2) to comprehend the meaning of that information (what rules are applicable for this specific intersection? What does heavy rain, for example, mean for that situation?) and 3) to project the status of the situation into the near future in order to act accordingly (given these factors, can I cross the intersection safely?).

The driver in control model by Hollnagel et al. (2003) suggests that the safe handling of a driving task requires a broad set of competences, all based on the dynamic interaction of a driver with the environment. This in turn determines the driving performance (i.e. is the driver able to cross the intersection safely or has s/he interpreted aspects of the situation incorrectly, as for example the speed of an approaching car?).

The task of driving is accomplished most of the time via subconscious processing of the environment by the means of fast pattern recognition (see for example Bellet, Bailly-Asuni, Mayenobe, & Banet, 2009). This automatic processing frees up cognitive resources and allows for an initiation of fast actions. This automatic processing (the 'direct path') is not mentioned in Endsley's (1995) situational awareness model (see for example Bellet et al., 2009; Kallus, 2009). Bellet et al. (2009) describe this state as 'implicit awareness' about the situation and the activities taken while driving. If a situation becomes more complex, the driver may switch from automatic, subconscious monitoring to conscious processing ('explicit awareness', Bellet et al., 2009). A key element here is the level of uncertainty. Uncertainty is characterized by a lack of information, triggering an active search for more information, which is coupled to more conscious processing (see also: behavioural inhibition system, Gray & McNaughton, 2003).

Risk Assessment

One crucial aspect of the driving task is the appropriate risk assessment of the current situation. Although this aspect was briefly mentioned in deliverable 1.1, we want to further specify this topic as it is highly relevant for our goal to identify, measure and coach safe driving behaviour.

Risk can be defined as "the projected likelihood and severity of the consequence or outcome from an existing hazard or situation" (p. 27, International Civil Aviation Organization, 2013). Following this definition, for an effective risk-assessment, hazardous driving conditions need to be detected and their likelihood and severity of the consequences need to be assessed, based on continuous feedback from the traffic environment. The driving behaviour needs to be adapted accordingly. The notion of the specific amount of risk thereby leads the actions further taken such as adjustment of speed, etc. (Harré, 2000; Horswill & McKenna, 2004; Senserrick, 2006).

In regard to the detection of hazardous driving situations, a study from Borowsky, Shinar, and Oron-Gilad (2010) showed for example, that older and experienced drivers were able to anticipate potentially hazardous situations even if their probability was low and the hazard itself was not salient (e.g. anticipating potential hazards at an T intersection). Compared to that, younger drivers detected hazardous situations if their saliency was high and therefore imposed an obvious threat. This suggests that the ability to anticipate hazards with even low saliency is crucial for safe driving.

Harré (2000) formulated distinctive "risk states" which guide driving behaviour – if a driver perceives a situation as risky, s/he will drive more carefully; on the other hand, if in a situation risk is assessed as low, the driver might adopt a more risky driving behaviour. Harré (2000) proposes to distinguish between deliberate risk-taking vs. failure to detect a risky situation. This differentiated understanding of risk perception subsequently calls for different intervention strategies. The study describes these "risk states" in relation to adolescent drivers but we believe that this description may also generalize at least to some extent to adult drivers with a relatively higher level of expertise. Harré (2000) postulates five distinct risk states that describe the judgement of the driver:

1) **habitual, cautious driving**: the driver perceives a low crash-risk. Objectively a low crash-risk is also present
2) **active avoidance**: the person perceives a relatively high level of risk but objectively a low level of risk is present

The states of habitual, cautious driving and active avoidance (depicted in the upper quadrant of *Figure 5.2*) are desirable from a safety point of view. The states of reduced risk perception, acceptance of risk as a cost and risk seeking (as depicted in the lower quadrant of *Figure 5.2* "Risk states" after Harré (2000). Indication of subjective value of perceived crash-risk, negative (-) and positive (+).) are related to an objectively high crash-risk and are therefore not desirable from a safety point of view:

3) **reduced risk perception**: perceived low crash risk with an objectively high crash-risk. The proposed intervention according to Harré (2000) is driver

training in order to increase awareness of hazardous behaviours and conditions. Reduced risk perception can be seen as a temporary state, for example when a driver is distracted. A suitable intervention in this case would be to educate the driver about the effects of distracting conditions.

4) **acceptance of risk as a cost**: perceived high crash-risk with an objectively high crash-risk. Perceived risk (cost) is outweighed by perceived gain and risk is accepted in order to achieve a goal. Interventions could consist of identifying and refining the "goals" which lead to higher risk acceptance within the coaching session. For example one goal could be to arrive on time while at the same time the fulfilment of concurrent goals are expected by the driver, such as loading and unloading or filling out paper work.

5) **risk seeking:** one step further, objective crash-risk is even higher and the driver is actively seeking the risk. Harré (2000) sees this closely related to the concept of sensation seeking (cf. Zuckerman, 1971, 1979; see Harré, 2000). The aspect of being able to "perform skilful" driving behaviour seems to be one of the relevant factors (Harré, 2000). Harré (2000) sees this type of driving as "the most difficult state to shift" (p. 218).

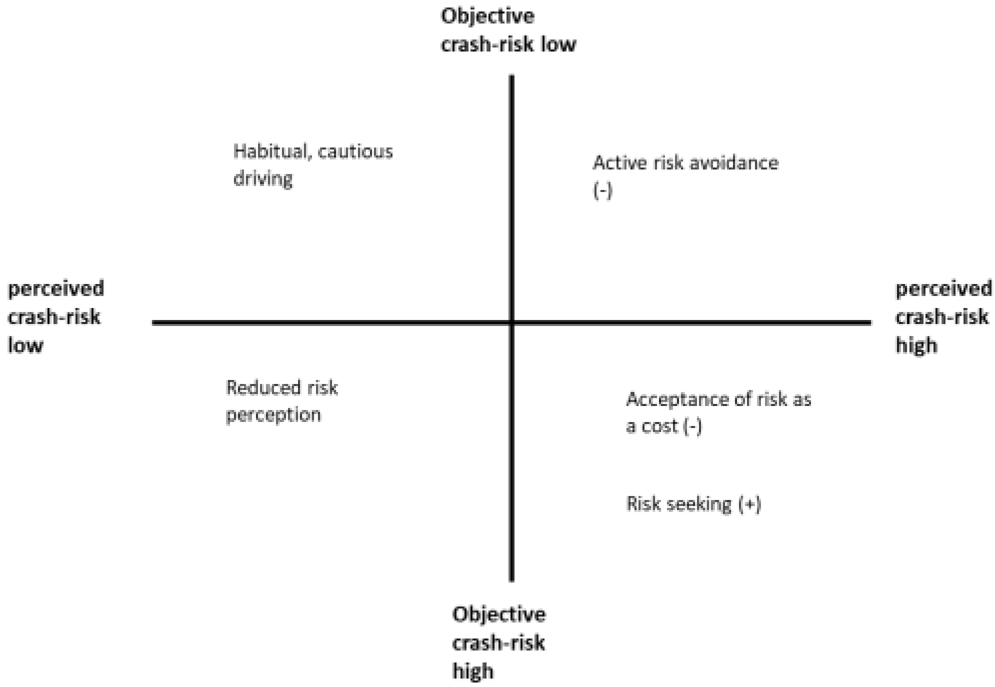


Figure 5.2 "Risk states" after Harré (2000). Indication of subjective value of perceived crash-risk, negative (-) and positive (+).

Horswill and McKenna (2004) argue that the perception of specific hazards itself is one specific skill that correlates with a driver's accident record. Senserrick (2006) provides a simple conceptual diagram (see *Figure 5.3*), depicting the sequence between hazard perception and crash avoidance actions and comparing three different "types" of drivers: a) experienced driver – unimpaired, b) experienced driver – distracted and c) inexperienced driver – unimpaired. The experienced driver scans the environment, detects a hazard and then several cognitive processes are engaged. The driver has to recognize the situation as a hazard which needs to be acted upon. S/he then needs to decide which response might be suitable and then to execute this specific action in order to avoid a crash. This overall process takes about two seconds to complete. A distracted but experienced driver (e.g. who is talking on the cell phone) detects the hazard maybe 0.5 seconds later than an undistracted driver and has therefore not enough time to respond accordingly. An unexperienced driver scans the environment not as efficiently as experienced drivers and therefore may also not have enough time to respond to the hazard accordingly. Senserrick (2006) therefore mentions two distinct factors: ineffective scanning (and therefore delayed hazard perception) and distraction, or attention allocation (regulation).

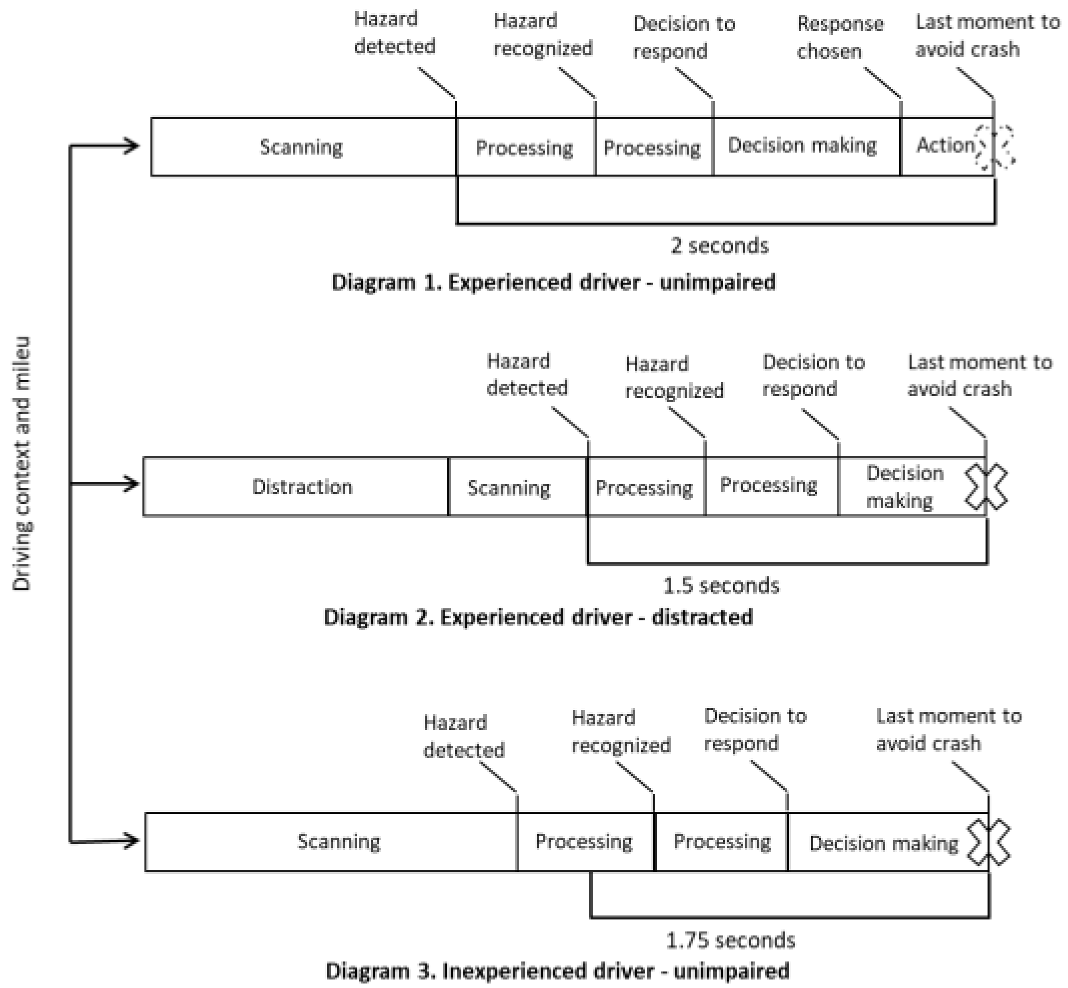


Figure 5.3 Conceptual diagram of hazard perception among three groups of drivers (after Senserrick, 2006).

Regulation (Attention)

The competence to regulate one's own attention is a crucial competence for driving. Inattention is defined as the "mismatch between the current allocation of resources and that demanded by activities critical for safe driving" (Engström et al., 2013, p.17). As described by Sollins, Chen, Reinerman-Jones and Tarr (2014), distractors are a major accident risk. The National Highway Traffic Safety Administration (2016) reported that in the year of 2014, 10% of fatal crashes and 18% of injury crashes are related to driver distractions within The United States. Distractors can be external (e.g.

passengers), self-generated (e.g. talking on the phone or eating while driving) or internal (e.g. rumination; Keating & Halpernfelsher, 2008).

Self-Appraisal

Self-Appraisal consists of the driver perceiving and validating his/her own driving capabilities. Depending on that assessment, drivers adjust their driving behaviour accordingly to the specific situation. As mentioned by for example Oxley, Fildes, Corben and Langford (2006), elderly drivers who are aware of their age-related limitations tend to adjust their driving behaviour for example by driving more slowly and carefully, avoiding difficult conditions, and reducing night driving. Interviews with stakeholders highlight factors such as 'normalization of risk' (as reported in deliverable 1.1): with increasing level of expertise, drivers tend to adjust their risk level as well such that speed for example is increased as relatively high level of expertise "allows" for that (see Karlsson et al., 2017).

As was also mentioned by Michon (1985), the competence of correct self-appraisal extends to the self-appraisal of one's own current state and one's own ability to drive in the presence of mediating factors which can have a detrimental influence on the driving performance, as for example fatigue and drowsiness, alcohol use, etc.

Driver competence model

Based on the literature described, we propose a basic driver competence model, which is depicted in *Figure 5.4*. Central here is the close interaction between driver, environment and vehicle (as the type of vehicle also influences driving actions). The driver her/himself has a set of competences, whereby the ability to adapt to changing driving situations can be seen as central. Following the definition of competence, we propose to specify competence in regard to these three aspects: attitude, knowledge and skills. This further specification will be done in the following section.

Additionally, factors, such as fatigue, stress etc. can moderate to which extent specific competences might be applied. For example, an in general safe and careful driver might show a very different driving style when under time pressure – the general competence of safe driving is significantly moderated by the factor time pressure.

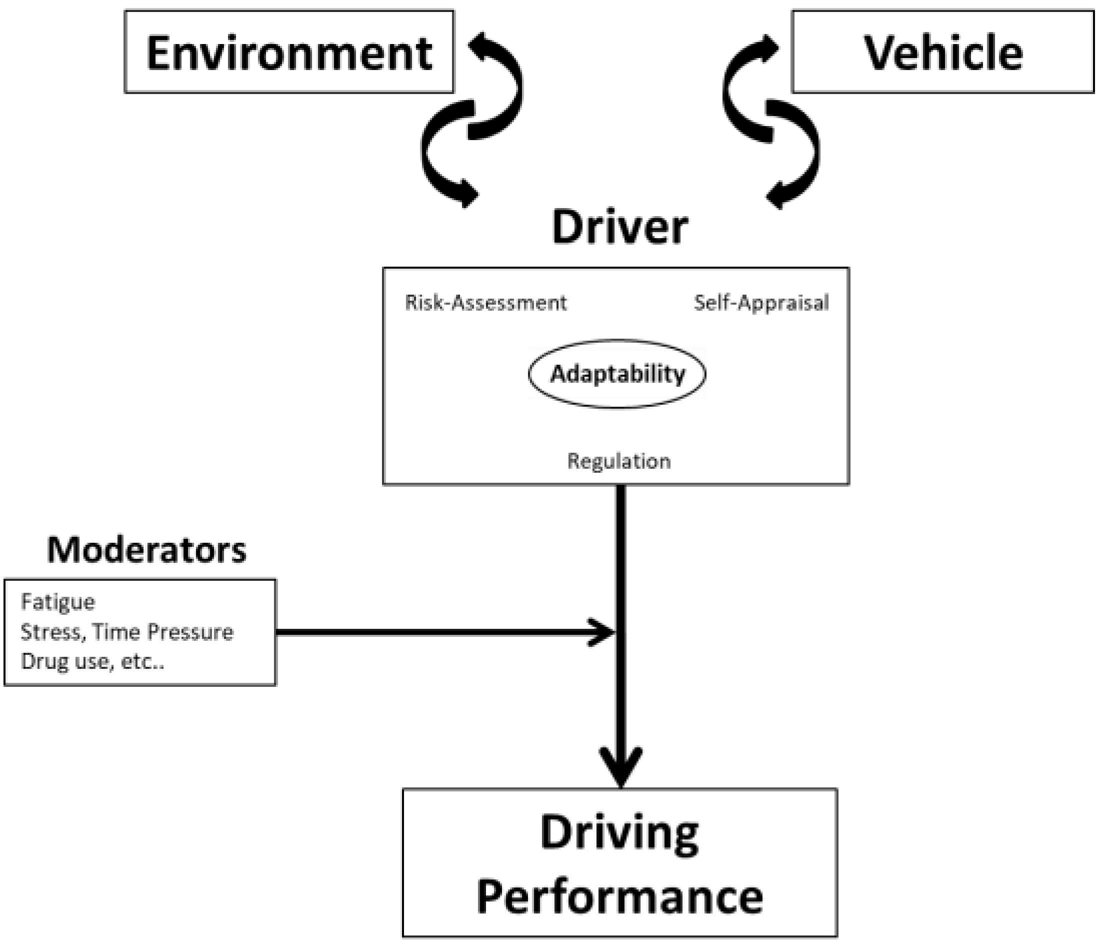


Figure 5.4 Basic driver competence model.

5.2.3 Specification of Driver Competences

Based on the competence model described in the previous section, this section will take the next step and define the specifics of driver competences concerning the three components of competence: attitude, knowledge and skills (see *Table 5.1*). With this approach, a very detailed picture of driving related competences is given. It combines observable behaviour (skills level) with more underlying components (attitude and knowledge).

Competence	Attitude	Knowledge	Skills
Risk-Assessment	o Beliefs/attitude towards risk related to driving (e.g. acceptance of risk, risk seeking, or cautious driver) o Beliefs/attitude about importance of safe driving actions (shoulder checks, etc.)	o Knowledge about hazardous/risky situations (based on past experience or theoretical knowledge to a lesser extend)	o Scanning patterns, monitoring strategy o Ability to detect hazardous driving situations o Compliance to safety margins; safe driving behaviour
Regulation	o Beliefs/attitude about importance of attention for the task at hand	o Knowledge about level of effort to apply concerning regulation o Knowledge about danger of distractors while driving	o Ability to allocate sufficient attention to driving task (active suppression of non-relevant information)
Self-Appraisal	o Adequate Beliefs/assessment about one's own skills and connection to risk assessment ('normalization of risk') o Beliefs about interplay between risk and current state of the driver	o Knowledge about one's own skills o Knowledge about influencing factors on driving performance (as for example fatigue)	o Ability to recognize fatigue or drowsiness o Ability to assess one's own skills adequately

Table 5.1 Definition of driver competences.

5.3 Profiling Requirements

In order to be able to assess the competences that a driver has, it is central to measure a broad set of variables. *Figure 5.5* gives an overview of how to derive these from observable driver behaviour (harsh braking, speeding, etc. – as discussed in the sections above) to the underlying competences. It is crucial to set the observable driving behaviour in context and in addition to also measure underlying attitudes, knowledge and skills.

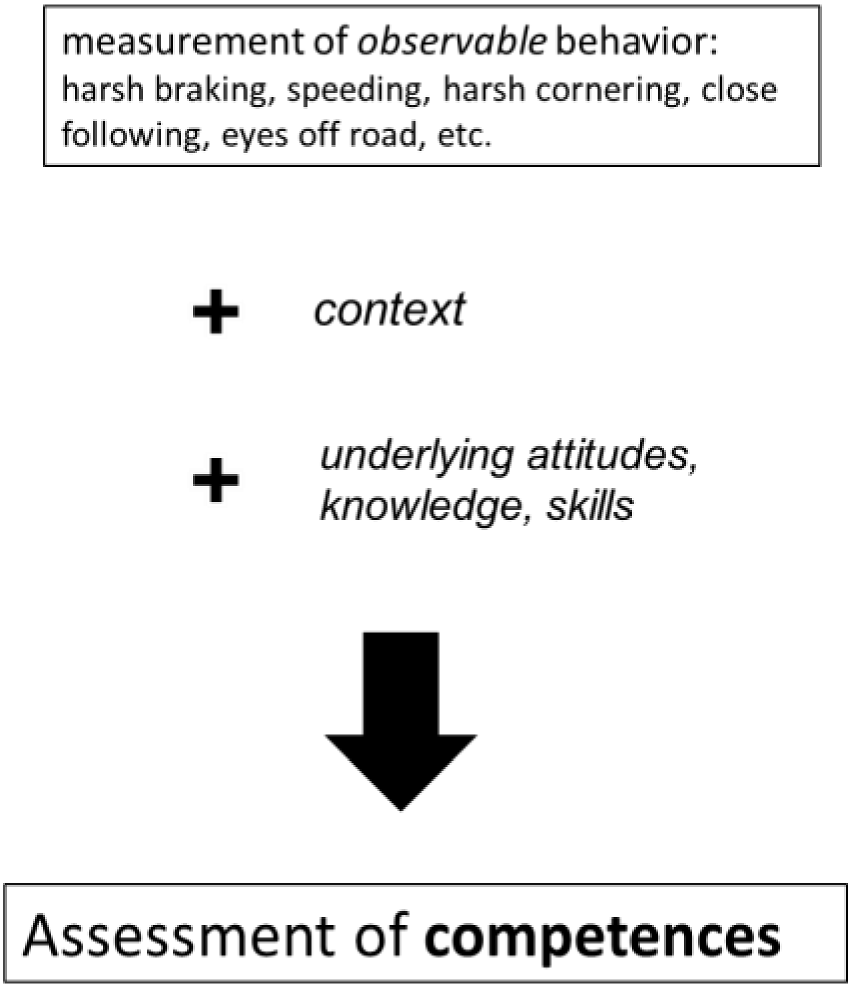


Figure 5.5 From observable behaviour to the assessment of driver competences.

For example, data could show that a driver was speeding at a specific point of time. This would be the observable behaviour, which can be measured via a cell phone or IVMS (see *Chapter 2*). The question would be, in which context the driver showed this kind of behaviour. Within this scenario, the driver was driving by a school at 50km/h while 30 km/h would be the appropriate speed. The driver was not slowing down while passing the school – the question would be, why s/he was not slowing down? There would be three possibilities: 1) the driver did not recognize that s/he was driving by a school (one possibility, because of distraction; skills level), or 2) the driver might not have the knowledge of the risk passing a school would impose, or lastly, 3) the driver might be well aware of the risk but accepts it anyway, because it serves a higher goal, as for example to arrive in time (attitude level).

For the measurement of the specific aspects of driver competences, measurements on varying levels can be conducted. This provides a possibility to structure the assessment of competences and subsequently the appropriate coaching strategy more accurately. One main advantage is that with this approach we do not rely solely on questionnaires, but rather combine a set of measurements, which allows assessing a broader picture of "driving competence". With the described model it is the intention to motivate future research and development in this direction. Although, it needs to be acknowledged that within the MeBeSafe project, the driver profiling will be concentrated on the assessment of observable behaviour (harsh breaking, speeding, etc.), as this is feasible in a naturalistic context.

	Measurement of		
Competence	**Attitude**	**Knowledge**	**Skills**
Risk-Assessment	○ survey on attitude towards risky situations (e.g. DSI, DAQ) ○ survey on beliefs about importance of safe driving actions (e.g. shoulder check)	○ survey on knowledge about situations with high risk	○ hazard perception test ○ monitoring strategy (eye-tracking)

Table 5.2 Specification measurement in regard to "risk-assessment".

A short questionnaire on attitude towards risky situations can provide information about the beliefs that a driver has regarding risk related to driving. Questions like:

- *"How acceptable is a certain level of risk for the driver?"*
- *"Which factors would influence acceptance of risk?"* – as for example time pressure: *"To what extend does time pressure influence the driver?"*

could guide the formulation of items. Also, existing questionnaires as for example the Driver Skill Inventory (DSI) (here: the balance within the self-assessment between manoeuvring and safety-skills as this might reflect the driver's attitude to safety; Lajunen, Corry, Summala, & Hartley, 1998; Lajunen & Summala, 1995) or the Driver Attitude Questionnaire (DAQ, (Parker, Stradling, & Manstead, 1996) could be suitable to assess a drivers attitude towards risk related to driving. It would be desirable to distinguish between the specific risk-states as stated by Harré (2000). Is the driver, for example, actively seeking risky situations? Or is the driver aware of the risk but accepts it anyway, maybe because it serves a "higher goal" – as for example arriving at the destination in time when under time pressure.

A short questionnaire on knowledge about risky situations might be suitable to assess whether there are possibly knowledge gaps in this regard. Coupled to a simulator session, this could be done, for example, by means of questions related to the situation (after the simulator session, the driver will be asked questions to the specific situations s/he encountered during the session). As for example, if the driver misses to reduce his/her speed while passing a school, it could be asked if s/he knows about the risk of that particular situation.

On the level of skills, we propose to introduce a short hazard perception test in order to test the driver's ability to detect possible hazardous situations. A hazard perception test could be video-based (the driver views short videos where s/he is asked to press a button when a hazard is detected) or simulator-based where hazardous situations are simulated and it is assessed how these situations are handled. In order to assess if a hazardous situation was detected (e.g. passing a school) it could be feasible to either ask the driver (questions related to the situation) or assess eye-tracking data.

	Measurement of		
Competence	**Attitude**	**Knowledge**	**Skills**
Regulation	○ survey: beliefs/attitude about importance of attention	○ survey: Knowledge about danger of distractors while driving	○ simulator: situations with distractors present (e.g. phone) → active suppression of non-relevant information

Table 5.3 Specification measurement in regard to "regulatory competence".

A short questionnaire on the beliefs about the importance of attention while driving could give valuable information on the general attitude towards this aspect of driving. Questions like

- *"How important is it for the driver to be undistracted?"*
- *"In as how far does distraction impose a safety risk for the driver?"*

could guide the item formulation process.

It could also be of interest to assess the level of knowledge a driver has about the dangers of distractors in general. One item example could be to ask the driver to estimate – given a specific speed - the distance driven without attending to the road while texting.

Within a simulator session, the level of skills in regard to the competence of regulation can be assessed:. A situation in which distractors are present (a call, navigating the navigation system, etc.) could be constructed. It could then be measured to what extent the driver engages in distracting activities.

	Measurement of		
Competence	**Attitude**	**Knowledge**	**Skills**
Self-Appraisal	o Survey: Beliefs about interplay between risk - current state of the driver – and driving skills ('normalization of risk')	o Survey: Knowledge about influencing factors (for example fatigue)	o Assessment of driver skills (e.g. subjective measurement: DSI) o Assessment of state (e.g. fatigue)

Table 5.4 Specification measurement in regard to "self-appraisal".

A short questionnaire can assess the driver's attitude concerning the interplay of risk - current state of the driver – and assessment of driving skills. Considering the self-assessment of one's own current state, it could be of interest to assess the number of times the driver has been driving when fatigued and whether this is considered as a safety risk. The measurement could also be extended to the assessment of driving skills and how this assessment influences the acceptance of risk ('normalization of risk' – with higher level of self-evaluated driving skills, drivers also tend to accept higher levels of risk). This could be measured, for example, by the means of subjective measurements (e.g. Driver Skill Inventory, DSI, Lajunen & Summala, 1995; Spolander, 1983).

A short questionnaire on the level of knowledge a driver has in regard to influencing factors, such as fatigue, drug use, etc. could indicate if knowledge gaps in this regard might exist.

Table 5.5 gives a short overview of the suggested instruments in order to assess the specific aspects of driver competences.

	survey, situative questions	visual observational data (eye-tracking, inward video)	objective tests (e.g. hazard perception test)	driving behavior	context information (outward video, GPS data)
A - risk assessment	x			x	x
K - risk assessment	x			x	x
S - risk assessment		x	x	x	x
A - regulation	x			x	x
K - regulation	x			x	x
S - regulation		x		x	x
A - self appraisal	x			x	x
K - self appraisal	x			x	x
S - self appraisal	x			x	x

A = attitude, K = knowledge, S = skill.

Table 5.5 Overview of instruments to assess aspects of driver competences.

As was discussed at the beginning of this section (see *Figure 5.5*), it is essential to assess a) what kind of behaviour was shown and b) in which context. However, as was already discussed in the previous chapters, it will be possible to assess the relevant data to some extent, as for example preliminary context information (see *Chapter 5*). *Table 5.5* also shows that the assessment of attitude and knowledge related aspects of the specific competences can mainly be assessed by means of surveys or situative questions and therefore, subjective data. As will be discussed in *Chapter 6*, the use of subjective data can be somewhat difficult in terms of validity. Although, within this approach, the combination of subjective data and observable behaviour can be of great value as it allows assessing a broader picture of driving related competences.

In the next chapter using questionnaires for driver profiling is explored.

6 Driver profiling based on questionnaires

In this chapter we examine the possibilities of using questionnaire data for driver profiling. The general idea of using questionnaires is that these could be used to give insight into why drivers behave in a risky manner and to make a distinction between risky and safe drivers. If we can categorise drivers or grade drivers in terms of their risk behaviour an option would be to evaluate the coaching programme for risky and safe drivers. Furthermore, questionnaires can add information about the drivers that cannot be measured by IVMS and the mobile phone. A disadvantage of questionnaires is however that self-reported behaviour could deviate from actual (driving) behaviour. Questionnaires that could be used in MeBeSafe are for example related to demographics, personality traits, attitude or self-reported driving behaviour.

Findings in this chapter are based on a literature review and an interview with a drivers coach at a haulier contracted by Shell. In this chapter we first discuss the demographics and the amount of driver experience that might have an influence on driving behaviour and could therefore be used for driver profiling. Then we discuss driver profiling based on personality traits and attitude related to driving behaviour. The chapter ends with a conclusion on the possibility of using questionnaires for driver profiling in the current MeBeSafe project.

6.1 Demographics and driving experience

Demographics and driving experience can be captured by asking the driver in a questionnaire or in an interview. In this literature review we explore what aspects influence HGV driver behaviour related to demographics and driving experience. In MeBeSafe deliverable 1.1 (Karlsson et al., 2017) several aspects were mentioned that could affect driver behaviour, namely age, fitness to drive, ability of information processing and action execution, gender, education and income, profession, expertise/driving exposure and cultural differences. Some of these characteristics also affect HGV driver behaviour, but not all are relevant. The gender of HGV drivers for instance will be mostly male; a distinction between male and female drivers would therefore not be feasible in WP4 of MeBeSafe. Age and experience are factors that could have an effect on driving behaviour for HGV drivers which could be looked at in the project.

A literature review by Duke, Guest, and Bogus (2010) shows that younger heavy vehicle drivers (27 years or younger) have an increased risk of crash involvement as do drivers aged 63-68 years. A study in Finland showed trailer-truck drivers younger than 30 years of age to have a 3.5 time higher risk of being responsible for an accident compared to drivers over the age of 50 years (Häkkänen & Summala, 2001). In contrary to the findings of Duke et al. (2010) a study by Guest, Bogges, and Duke. (2014) has indicated that professional drivers in Australia beyond the age of 65 are not at greater risk of traffic accidents. A study by Girotto (2016) indicates that less years of professional driving experience is associated with a higher involvement in accidents and near-miss accidents for Brazilian truck drivers, regardless of age, substance abuse, working conditions and behaviour in traffic.

Additionally, in an interview with a driver coach at a haulier it was explained that they viewed the amount of driving experience that drivers have at *their* company as an important predictor of how safe their drivers' behaviour is on the road. They stated that when drivers had been working for a longer period of time at their company, drivers drove safer, regardless of their age or the experience they had before joining the company. Also mentioned was that a difference between driving behaviour was noted between younger and older drivers; younger drivers tend to use the driver assistance systems more than older drivers. They also highlighted that the type of trips drivers undertake is an important factor to take into account. As some drivers drive shorter trips and will spend more time on urban roads, and others will drive more on the highway, furthermore workload might differ between trips. This again highlights the importance of taking context into account for the coaching of driving behaviour.

Concluding from this, there are indications that age and driving experience or work experience at a certain company could predict crash involvement or risky driving behaviour, though not all studies show the same findings. Similarly, the study described in *Chapter 3* of this deliverable did not find an effect of age on the number of harsh braking and speeding events. More research is needed to understand how these factors precisely shape risky driving behaviour of HGV drivers before using these factors for driver profiling in this MeBeSafe project.

6.2 Driving behaviour, personality traits and attitude

Besides looking at demographics and driving experience we looked at using questionnaires that measure personality traits, attitude and self-reported driving behaviour for driver profiling.

We therefore explored the use and validity of several scales that measure driving behaviour or related personality traits and attitudes. One questionnaire was of particular interest: the Manchester Driver Behaviour Questionnaire (DBQ) by Reason, Manstead, Stradling, Baxter and Campbell (1990). The DBQ is a widely accepted measure of aberrant driving behaviour and is often used to predict crash involvement. Nevertheless, there are drawbacks when it comes to its validity.

In 2003, Lajunen and Summala tested the scale's sensitivity to socially desirable responses in different settings, and concluded that the DBQ is relatively insusceptible to this bias. However, they did not test the DBQ's ability to predict actual crash involvement. af Wåhlberg, Dorn, and Kline (2011) did, and found that the DBQ can be used to predict self-reported, but not actual past crash involvement. These results are unsurprising, considering that research has shown that self-reported crash involvement data does not correspond with archived crash involvement data (Arthur, Bell, Edwards, Day, Tubre, & Tubre, 2005). A more recent study found that the DBQ, in combination with the Driver Skill Inventory (DSI; Lajunen & Summala, 1995), can predict traffic offenses but, again, not actual crash involvement (Martinussen, Møller, Prato, & Haustein, 2016). Despite concerns on the validity of the DBQ Martinussen, Møller and Prato (2014) state that the DBQ is a valid measure to divide drivers into sub-groups. Wishart, Freeman and Davey (2006) suggest that the DBQ might be used to develop tailored interventions. However, the results of both studies are limited, since they solely rely on self-reported data.

How problematic the use of self-reports actually is, is shown in a study by af Wåhlberg, Dorn, and Kline (2011), who found social desirability effects on self-reported traffic accidents. In addition, research by af Wåhlberg (2010) has shown that the association between several leading questionnaires on driver characteristics and self-reported crash involvement, as well as penalty points, is clearly weakened when social desirability effects are controlled for. This was found for all investigated scales: the violation scale of the DBQ, the Driving Anger Scale (DAS; Deffenbacher, Oetting, & Lynch, 1994), the short version of the Sensation Seeking Scale (SSS; Slater, 2003) and the aggression scale of the Driver Behaviour Inventory (DBI; Gulian, Glendon, Matthews, Davies, & Debney, 1988). Furthermore, a meta-analysis study by af Wåhlberg, Barraclough & Freeman (2015) on the DBQ indicated that drivers whom report a high number of violations and/or crashes might do so because they spend more time on the road, not because they violate more often. A study by Freeman, Barraclough, Wishard, and Rowland(2014) looked into the validity of the DBQ based on data collected from three Australian fleet samples. Regression analysis revealed that road exposure was the best predictor of (self-reported) crash involvement, rather than the factors measured by the DBQ. A study by Precht et al. (2017) looked into the main factors that contribute to driving errors and traffic violations for car drivers, using data collected in a large naturalistic driving study in the United States (SHRP 2). The added value of naturalistic driving data is that questionnaire data can be related to actual driving behaviour. Their study found that scores on a sensation seeking scale did not predict violations or errors committed by car drivers as identified on video, but also age, experience and the number of self-reported accidents were not related to the number of violations and errors. Anger, passenger presence and self-reported violations were the main factors associated with committed violations. Surprise, high-risk visually distracting secondary tasks and passing through an interchange were the main factors associated with committed errors. It should be noted though that their study was based on driving behaviour of 38 car drivers, i.e. the sample is relatively small and factors associated with violations and errors may be different for truck drivers.

Saberg, Piccinini, and Engström (2015) conclude in a literature review on driving styles and road safety that caution is needed for generalising self-reported data to actual driving behaviour. For self-reported speed behaviour correlations of 0.6 have been reported with actual driving behaviour, though for other behaviours correlations are often weak. Regarding driver profiling, Saberg, Piccinini, and Engström (2015) state that different driving styles might predict crash involvement, but that more research that involves actual crash involvement is needed. According to the same reference, studies so far have used various definitions and have assessed different aspects of driving styles. Before assessments of driving styles can effectively be used for driver profiling or the development of training programs, a common theoretical framework should first be developed that captures different aspects of driving styles. The authors provide the same argument for demographic and sociocultural factors like age and driving experience. More research is needed to better understand how we can use these individual driver characteristics for driver profiling.

Lastly, it should be noted that the existing questionnaires are research tools and are not specifically developed for a coaching context. Considering this and all presented information about the validity of existing driving behaviour questionnaires, especially the most commonly used DBQ, we conclude that driver profiling based on the discussed questionnaires that measure personality traits and attitude is not useful in the present context.

6.3 Implications based on the literature review

In this chapter the possibilities of driver profiling based on demographic characteristics, driving experience, personality traits and attitude were investigated.

In terms of demographic characteristics age, the amount of driving experience drivers have and the number of years working at the present company could be important predictors for risky behaviour. Studies show that crash involvement is higher for drivers that are younger and drivers that have less driving experience, so they might drive riskier. These factors could therefore be used for driver profiling. However, more research is recommended to understand how these factors precisely shape risky driving behaviour before doing so.

Questionnaires that aim to capture driving behaviour or related personality traits and attitudes, like the DBQ, are not considered useful for driver profiling since the validity of these measures is low. Also, most research has focused on self-reported offences, self-reported accidents and archived accidents. Caution is needed when generalising self-reported behaviour to actual driving behaviour. More research is needed looking into actual driving behaviour and driver characteristics using more objective methods in naturalistic settings.

To conclude this chapter, more research is needed into driver profiling based on driver characteristics before using driver profiling in the current MeBeSafe project. Nevertheless, questionnaires and other subjective methods of measuring driver characteristics and for example competences can provide insights into why a driver acts in a risky manner while driving. Measuring driver characteristics with questionnaires or interviews could therefore still be of value.

7 Conclusion and the representation of driver profiles

The objective of the work described in deliverable 4.1 was to investigate what data is needed for coaching of heavy good vehicle drivers, how we can collect these data, what variables are relevant for driver profiling and how we can use these variables for driver profiling. In this chapter, an overview is presented on the possibilities of how driving behaviour, the environment and driver characteristics can be measured, and how driver profiles based on driving behaviour and the environment can be represented. The results can be used as input for the design of coaching schemes and the app that will measure and provide feedback on driving behaviour (Task 4.3) and eventually for the field evaluation test in WP5.

We start the conclusion with a summary of what driving behaviour variables we wish to improve with coaching and what variables can have an impact on driving behaviour, followed by how we can measure these variables. We conclude with a suggestion of how to represent driver profiles based on driving behaviour and context.

7.1 What driving behaviour variables do we wish to improve by coaching?

Risky driving behaviour can lead to crashes but by coaching drivers on their driving behaviour the aim is to reduce risky driving, reduce crashes and consequently increase traffic safety. The following variables are relevant coaching in this MeBeSafe project: harsh braking, speeding, distraction, drowsiness/fatigue, close following, harsh cornering, lane departure and optionally fuel consumption. These variables have been labelled the Key Performance Indicators (KPIs).

Measuring driving behaviour in itself is often not sufficient, as behaviour should be put into context whenever possible. The characteristics of the driving environment can influence how a driver behaves. For example, in an urban environment drivers will most likely have to brake harshly more often compared to when driving on a highway. Drivers that drive more in urban environments will therefore show more harsh braking behaviour compared to drivers that drive more often in rural environments.

Consequently, a difference between drivers could be the result of differences in the environment they are driving in, and not so much their actual driving behaviour. To generate a fair representation of driving behaviour it is important to consider the characteristics of the trips and situations the drivers are in. Furthermore, to understand *why* a driver is behaving in a certain way we should look at the characteristics of a driver. The behaviour of a driver can be influenced by characteristics like personality, attitude, age, experience and competences. In this deliverable, we examined how we can measure driving behaviour, the driving environment, and driver characteristics, which can all be relevant for interpreting the behaviour of a driver and therefore for coaching.

7.2 How can we measure driving behaviour, environment and driver characteristics?

The possibility to measure any set of variables depends on the tools at hand. *Figure 7.1* gives an overview of what driving behaviour variables (KPIs), driver and trip characteristics that can be measured by IVMS, mobile phone, questionnaires and cameras respectively. Listed are also the advantages and disadvantages per type of measurement.

	Variables	IVMS	Mobile phone	Cameras	Questionnaires
Driving behaviour (KPI)	Harsh braking				
	Speeding	Possibly, combined with map data	Possibly, combined with map data		
	Eyes off target (distraction)			Inward-facing	
	Eye closure rate (drowsiness)			Inward-facing	
	Close following			Outward-facing	
	Harsh cornering				
	Lane departures			Outward-facing	
	Fuel consumption				
Driver and context characteristics	Location, route, date and time of a trip				Limited to general behaviour
	Urban/rural, intersection	Possibly, combined with map data	Possibly, combined with map data	Outward-facing	
	Conditions (weather, congestion, etc.)		Possibly, combined with map data	Outward-facing	
	Competences				
	Age, experience				
	Advantages	Accurate information.	All drivers can use it (no need to install expensive equipment).	Additional information on the traffic situation and driver behaviour.	Possible to measure background variables.
	Disadvantages	Not all vehicles are equipped with (the same) systems.	We need to ensure drivers turn the device on.	Privacy issues for inward-facing cameras. Possibly not allowed in some countries. Costs for installing and cameras.	Validity issues.

Figure 7.1 Variables that can be measured with IVMS, mobile phone, cameras and questionnaires. A green rectangle indicates that measuring is possible, yellow indicates that measuring is possible but that there are drawbacks, and red indicates that measuring is not possible. Listed also are the advantages and disadvantages of the type of measurement.

The options that were looked at for measuring basic driving behaviour regarding position, speed, braking, cornering, and the optional variable fuel consumption were ***IVMS*** and ***mobile phone***. Results show that roughly the same, but not all, KPI variables can be measured by these devices. Although IVMS can measure certain variables more precisely than the mobile phone, a great disadvantage of IVMS is the fact that not all vehicles have the same system installed. This is a problem, because the software that is used to read the data from the IVMS needs to be adapted for every system, making it a costly and time consuming endeavour to adapt software for each new system. Also, data like speed limits, weather conditions, and type of road are relatively easily combined with data collected by a mobile phone, but not as easily with an IVMS system. Therefore, the mobile phone is a better option to use for measuring driving behaviour in the current project.

Nonetheless, a lot of variables cannot be measured by mobile phone or IVMS (*Figure 7.1*). Some of these variables are related to driving behaviour (distraction, drowsiness, close following and lane departures), as well as additional environment related variables that provide information on the trip and situation the driver is in. ***Cameras*** can capture additional driving behaviour as well as information on the environment. Cameras would therefore be of great value for collecting additional information for coaching and feedback provided by the app. Outward-facing cameras would give more information on the conditions a driver is in, for example traffic density can be measured and videos of relevant situations can be saved and shown to a driver. Inward-facing cameras could provide information on distraction and drowsiness, both important factors related to traffic safety.

There are however also some important disadvantages with using cameras that can influence the project and the development of coaching schemes and app. Presumably not all drivers will want to have cameras installed (especially inward-facing) in their vehicle. During pilot tests we will further examine the readiness and acceptance by HGV drivers to have cameras installed. In addition, the technology is relatively expensive. For the MeBeSafe project this means that not all trucks can be equipped with cameras. Furthermore, for the usage and implementation of coaching schemes and app after the MeBeSafe project, reliance on cameras could possibly make these coaching measures too expensive.

The number of variables relevant for the current project that can be measured accurately with ***questionnaires*** is relatively limited, and several studies point out that the validity of questionnaires measuring personality and driving behaviour is low. However, gaining insight into why a driver behaves in a certain manner is only possible by looking into competences, personality and attitudes, which are difficult to measure other than by questionnaires, interviews or tests (simulator, hazard perception tests). With questionnaires, the attitude towards coaching could also be measured. Since more research is needed on using basic driver characteristics (age, gender, self-reported behaviour etc.) for driver profiling and it is questionable at this point whether clear relationships can be determined, the focus for now is on profiling based on driver behaviour and the characteristics of the environment. Questionnaires can be used mainly as extra input for initial data on demographics and for assessing competence and attitude.

7.3 Representing driver profiles based on driving behaviour

We propose a representation of driver profiles based on driving behaviour in a "Traffic Safety Wheel" (see *Figure 7.2*), wherein we can:

- Represent individual driver behaviour as variables relative to our main KPI variables;
- Compare driver behaviour to fleet behaviour (see *Figure 7.3*);
- Visualise the amount of overlap in driving behaviour (see *Figure 7.4*) and;
- Make a distinction between behaviour in different conditions (see *Figure 7.5*).

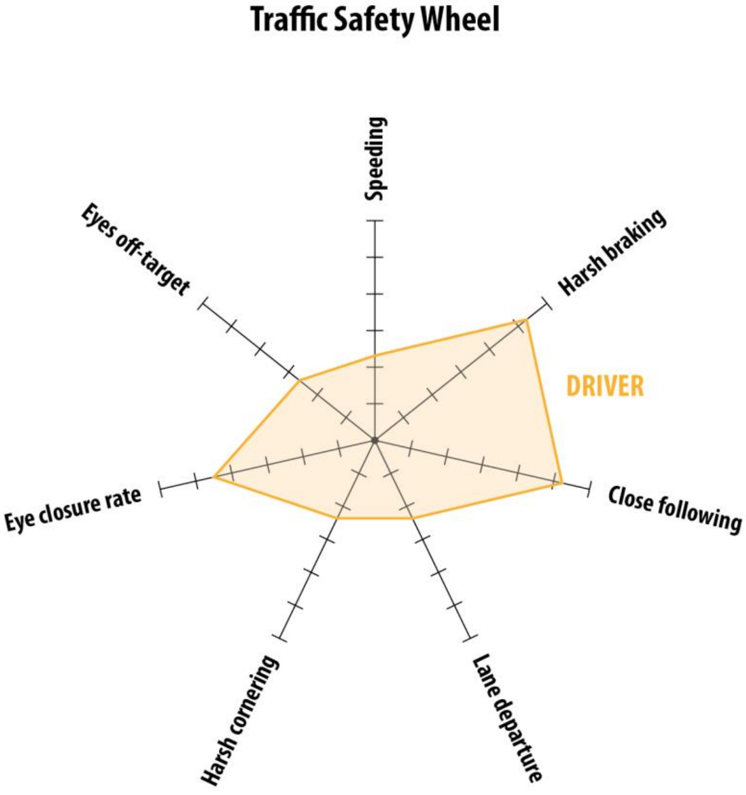


Figure 7.2 Traffic Safety Wheel wherein driving behaviour of one driver is visualised on seven different axes.

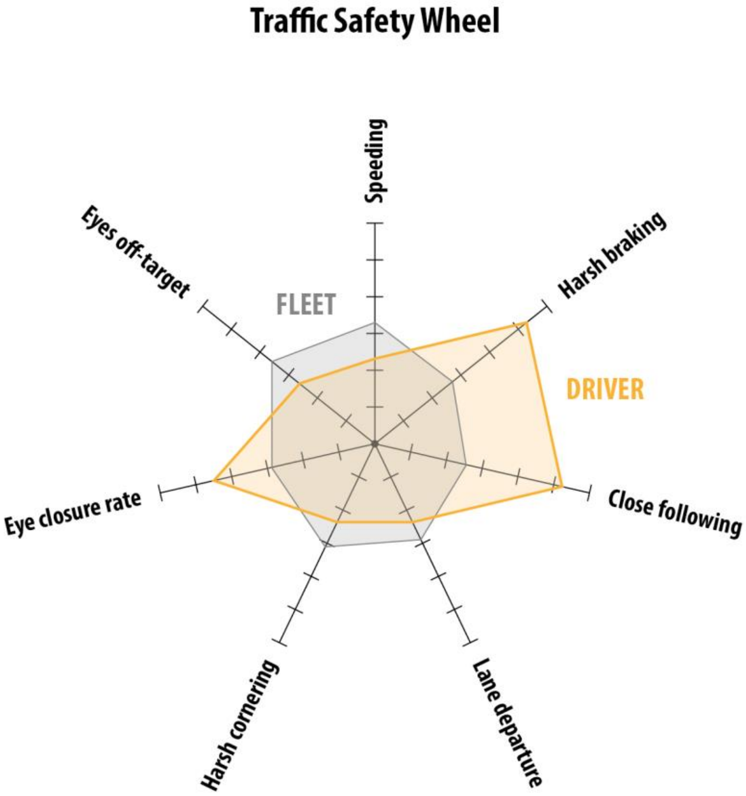
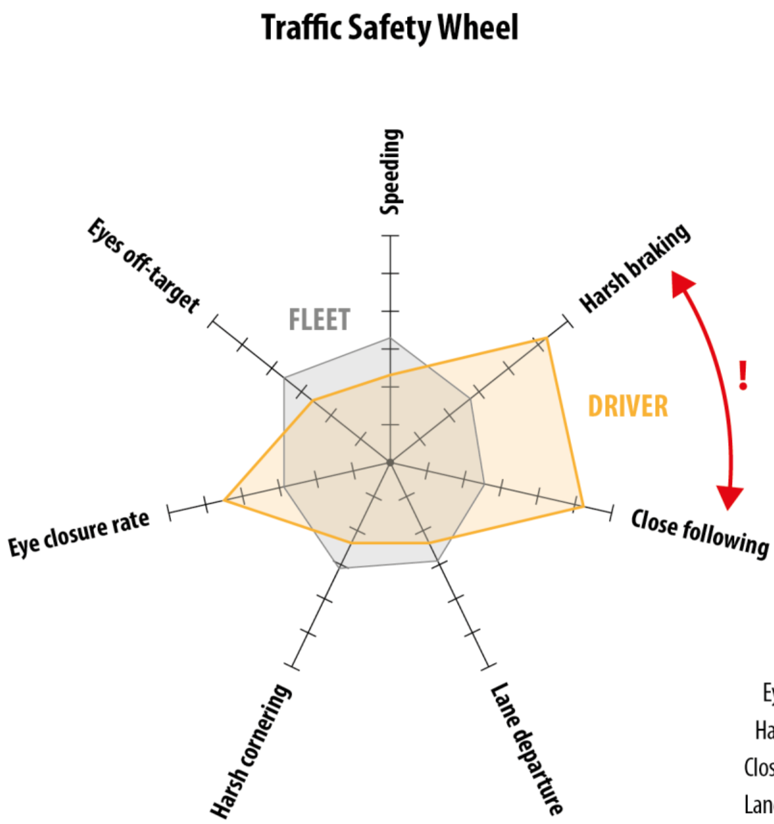


Figure 7.3 Traffic Safety Wheel wherein driving behaviour of one driver can be compared to average driving behaviour of the fleet.

Figure 7.2 maps the behaviour of a specific driver on several KPI-axes. The shape that results from this mapping gives an overview of how well a driver scores on each of the KPI variables. A larger surface in the safety wheel implies that a driver shows more risky behaviour compared to a smaller surface in a safety wheel. Variables that need more attention can be highlighted and axes can be left out in the figure. For instance, when an inward-facing camera hasn't been used, eyes off road/target and eye closure rate (drowsiness/fatigue) can be left out.

Figure 7.3 shows the profiles of a driver as well as his/her fleet. The visualization of the profiles facilitates an evaluation of driver performance in relation to the fleet: it can be easily seen on what variables a driver scores better/worse than the average truck driver in a fleet.

What can also be visualized in the safety wheel are high degrees of "overlap" between different KPI variables, as visualized by the exclamation mark in *Figure 7.4*. This can be determined at the level of individual events, i.e. short periods of time in which particular types of relevant driving behaviour occur at a particular place and time, and which are measured by the data acquisition system. In the figure, an example is given of harsh braking and close following. When a harsh braking event and a close following event overlap, they are labelled as concurrent. A high event concurrency between two variables means they are possibly causally related, because many of the corresponding events occur at the same time; and this is naturally useful and possibly important information for coaching. In the above example, this means that when we coach risky close following behaviour, we will likely also influence harsh braking behaviour. When drivers take more distance they will have more time to brake and therefore have to brake less harshly. Event concurrency can be calculated for each pair-wise combination of KPIs, as shown in the matrix depicted in *Figure 7.4*. Coaching can be supported by, for example, selecting the combination with the highest event concurrency, or by selecting combinations with a concurrency rate above a to-be-chosen threshold.



Event concurrency matrix (%)

	Speeding	Eyes off-road	Harsh braking	Close following	Lane departure	Eye closure rate	Harsh cornering
Speeding	1	.1	.0	.2	.1	.2	.1
Eyes off-road		1	.2	.3	.4	.2	.4
Harsh braking			1	.6	.1	.0	.2
Close following				1	.2	.3	.0
Lane departure					1	.1	.3
Eye closure rate						1	.2
Harsh cornering							1

Figure 7.4 Traffic Safety Wheel wherein the overlap between two driving behaviour variables is visualised (concurrency).

Another important feature of the safety wheel is that we are able to visualize the effect different conditions and environment factors have on driving behaviour. This is important, because driver behaviour can differ across circumstances. For example, by distinguishing between behaviour in the city and behaviour on the highway, as is shown in *Figure 7.5*, specific behaviour can be targeted for coaching under specific circumstances.

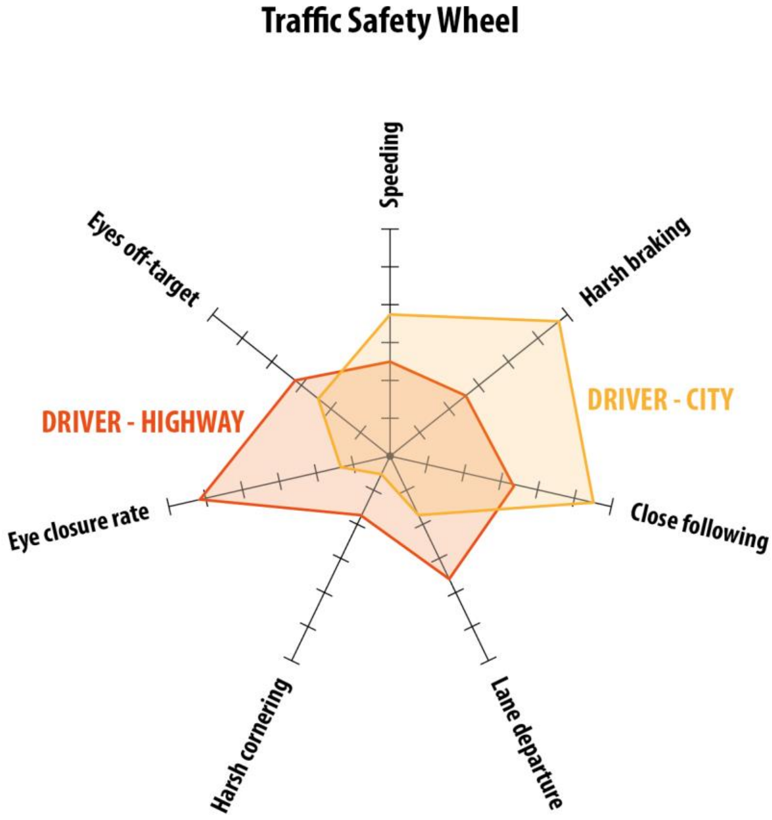


Figure 7.5 Traffic Safety Wheel wherein driver behaviour is visualised in different contexts; behaviour in the city and behaviour on the highway.

We suggest a three-step approach to use driver profiles in support of coaching:

1) Map the driver profile on the safety wheel and look at the event concurrencies. If there is a combination of KPIs with a high event concurrency rate, these KPIs can be the first focus for coaching. The reason for this is that by improving one KPI, the other KPI(s) will likely improve as well;
2) Determine under what circumstances risky driving behaviour is most prevalent. For example, coaching may be most relevant at highways, and within highways focusing only on junctions, etc. This can be visualised in the safety wheel;
3) If possible, investigate the underlying causes of risky driving behaviour by looking at the competences of a driver or other driver characteristics. This can be done using questionnaires or interviews.

7.4 Concluding remarks

In this deliverable we have focused on capturing driving behaviour and the driving context in driver profiles. With regard to technology, our recommendation is to measure driving behaviour and context with a mobile phone, augmented with inward- and outward-facing cameras where possible. The disadvantages of using cameras need to be kept in mind though; the installation of cameras brings more expenses compared to using only mobile phones and there are privacy and acceptance issues for drivers and due to legislation in certain countries.

In terms of driver profiling, we have proposed the Traffic Safety Wheel, a representation of driver profiles wherein we can compare driver behaviour with fleet behaviour across varying driving contexts. The safety wheel as presented could serve as a foundation for driver profiles, but it is not yet intended to use directly as visualization for drivers – even though it can be the basis for it. For the latter purpose, the safety wheel should first be tested by HGV drivers and usability experts on visual appeal and ease of use, possibly followed by a redesign. Furthermore, in the app used on the mobile phone and for coaching there could be an additional focus on emphasizing positive driving behaviour, next to risky behaviour. Positive driving behaviour is not yet most naturally represented in the Traffic Safety Wheel. But could be visualised for example by using a "positive" green colour as background colour, when variable values are close to the centre of the Traffic Safety Wheel more of the green colour is shown. Another option would be to use a kind of "bull's eye" visualization. The Traffic Safety Wheel needs more development before it can be used in the MeBeSafe project.

Based on the results described in this deliverable further decisions can be made on what driver profiles can be used and how data should be collected for Work Package 4.

8 References

A

Albert, D., & Schrepp, M. (1999). Structure and design of an intelligent tutorial system based on skill assignments. Knowledge Spaces-Theories, Empirical Research, and Applications, 179-196.

Arthur, W., Bell, S. T., Edwards, B. D., Day, E. A., Tubre, T.C., & Tubre, A.H. (2005). Convergence of Self-Report and Archival Crash Involvement Data: A Two-Year Longitudinal Follow-Up. *Human Factors*, *47*(2). 303-313.

Arnett, J. (1994). Sensation seeking: A new conceptualization and a new scale. *Personality and Individual Differences*, *16*(2), 289–296.

B

Baltrusaitis, T., Robinson, P., & Morency L.P. (2013, December). Constrained Local Neural Fields for Robust Facial Landmark Detection in the Wild. Paper presented at *IEEE International Conference on Computer Vision Workshop (ICCVW).*

Bellet, T., Bailly-Asuni, B., Mayenobe, P., & Banet, A. (2009). A theoretical and methodological framework for studying and modelling drivers' mental representations. *Safety Science*, *47*(9), 1205–1221. https://doi.org/10.1016/j.ssci.2009.03.014

Borowsky, A., Shinar, D., & Oron-Gilad, T. (2010). Age, skill, and hazard perception in driving. *Accident Analysis & Prevention*, *42*(4), 1240–1249. https://doi.org/10.1016/j.aap.2010.02.001

C

Czupryski, B., & Strupczewski, A. (2014). High accuracy head pose tracking survey. In: lzak, D., Schaefer, G., Vuong, S., & Kim, Y.S. (eds) Active Media Technology. Lecture Notes in Computer Science, vol 8610, 407–420. Springer International Publishing.

D

Danelljan, M., Häger, G., Khan, F.S., & Felsberg, M. (2015, December). Convolutional Features for Correlation Filter Based Visual Tracking. Paper presented at *IEEE International Conference on Computer Vision Workshop (ICCVW).*

Dingus, T.A., Guo, F., Lee, S., Antin, J.F., et al. (2016). *Driver crash risk factors and prevalence evaluation using naturalistic driving data*. In: Proceedings of the National Academy of Sciences of the United States of America PNAS.

Dotzauer, M., Stemmler, E., Utesch, F., ..., & Carsten, O. (2017). UDRIVE deliverable 42.1 Risk factors, crash causation and everyday driving of the EU FP7 Project UDRIVE (www.udrive.eu).

Duke, J., Guest, M., & Boggess, M. (2009). Age-related safety in professional heavy vehicle drivers: A literature review. *Accident Analysis and Prevention*, *42*, 352-371.

E

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, *37*(1), 32-64. https://doi.org/10.1518/001872095779049543

Engström, J., Monk, C. A., Hanowski, R. J., Horrey, W. J., Lee, J. D., McGehee, D. V., ... Yang, C. Y. D. (2013). *A conceptual framework and taxonomy for understanding and categorizing driver inattention* (Project report). Brussels, Belgium: European Commission.

F

Federal Motor Carrier Safety Administration (FMCSA), 2006. Report to Congress on the Large Truck Causation Study, Washington DC.

Freeman, J. Barraclough, P., Wishard, D., & Rowland, B. (2014). Using the Driver Behaviour Questionnaire to predict crashes and demerit point loss: Does it get better with larger sample sizes? *Centre for* Accident *Research and Road Safety – Queensland (CARRS-Q),. Queensland University of Technology.*

French, D.J., West, R.J., Elander, J., & Wilding, J.M. (1993). Decision-making style, driving style, and selfreported involvement in road traffic accidents. *Ergonomics*, *36*, 627–644.

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis & Prevention*, *37*(3), 461-472.

G

Girshick, R. (2015, December). Fast R-CNN. Paper presented at *IEEE International Conference on Computer Vision (ICCV).*

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014, June). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. Paper presented at *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR).*

Gudi, A., Tasli, H.E., den Uyl, T.M., & Maroulis, A. (2015, May). Deep Learning based FACS Action Unit Occurrence and Intensity Estimation. Paper presented at *IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG).*

Gray, J. A., & McNaughton, N. (2003). *The neuropsychology of anciety: An enquiry into the function of the septo-hippocampal system*. Oxford: University press.

Gudi, A., Tasli, H.E., den Uyl, T.M., & Maroulis, A. (2015, May). Deep Learning based FACS Action Unit Occurrence and Intensity Estimation. Paper presented at *IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG).*

Guest, M., Boggess, M.M, & Duke, J.M. (2014). Age related annual crash incidence rate ratios in professional drivers of heavy goods vehicles. *Transportation Research Part A, 65*, 1-8.

H

Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006). *Multivariate Data Analysis (6th ed.).* New Jersey: Pearson Educational, Inc.

Häkkänen, H., & Summala, H. (2001). Fatal traffic accidents among trailer truck drivers and accident causes as viewed by other truck drivers. *Accident Analysis and Prevention, 33* (2), 187–196.

Häkkänen, H., Summala, H., Partinen, M., Tiihonen, M., & Silvo, J. (1999). Blink Duration as an Indicator of Driver Sleepiness in Professional Bus Drivers. *SLEEP*, *22*(6), 798-802.

Hanowski, R.J., Perez, M.A., & Dingus, T.A. (2005). Driver distraction in long-haul truck drivers. *Transportation Research Part F, 8*, 441-458.

Hansen, D.W., & Ji, Q. (2010) In the eye of the beholder: a survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(3), 478-500.

Harré, N. (2000). Risk Evaluation, Driving, and Adolescents: A Typology. *Developmental Review*, *20*(2), 206–226. https://doi.org/10.1006/drev.1999.0498

Hollnagel, E., Nabo, A., & Lau, I. V. (2003). A Systemic Model for Driver-in-Control. In *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 86–91). Park City, Utha.

Horne, J., & Reyner, L. (2001). Sleep-related vehicle accidents: some guides for road safety policies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *4*(1), 63-74.

Horswill, M. S., & McKenna, F. P. (2004). Driver's hazard perception ability: situation awareness on the road. In S. Banbury & S. Tremblay (Eds.), *A Cognitive Approach to situation Awareness.* Aldershot, UK: Ashgate.

I

International Civil Aviation Organization (Ed.). (2013). *Safety management manual (SMM)* (3. ed). Montreal: ICAO.

J

Jiang, B., Valstar, M.F., & Pantic, M. (2011, March). Action unit detection using sparse appearance descriptors in space-time video volumes. Paper presented at *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011).*

K

Kallus, K. W. (2009). Situationsbewusstsein und Antizipative Prozesse. *Zeitschrift Fuer Arbeitswissenschaft: Zentralblatt Fuer Arbeitswissenschaft Und Soziale Betriebspraxis*, (63), 17–22.

Karlsson, M., Brouwers, R., Fahrenkrog, F.,...., & Ziegler, J. (2017). *Integrated Framework. MeBeSafe deliverable 1.1.*

Keating, D., & Halpernfelsher, B. (2008). Adolescent DriversA Developmental Perspective on Risk, Proficiency, and Safety. *American Journal of Preventive Medicine*, *35*(3), S272–S277. https://doi.org/10.1016/j.amepre.2008.06.026

L

Lajunen, T., Corry, A., Summala, H., & Hartley, L. (1998). Cross-cultural differences in Drivers' self-assessments of their perceptual-motor and safety skills: Australians and Finns. *Personality and Individual Differences*, *24*(4), 539–550. https://doi.org/10.1016/S0191-8869(97)00202-X

Lajunen, T., Parker, D., & Summala, H. (2004). The Manchester Driver Behavior Questionnaire: A crosscultural study. *Accident Analysis, and Prevention*, *36*, 231-238.

Lajunen, T., & Summala, H. (1995). Driving experience, personality, and skill and safety-motive dimensions in drivers' self-assessments. *Personality and Individual Differences*, *19*(3), 307–318. https://doi.org/10.1016/0191-8869(95)00068-H

Lajunen, T., & Summala, H. (2003). Can we trust self-reports of driving? Effects of impression management on driver behaviour questionnaire responses. *Transportation Research Part F: Traffic Psychology and Behaviour*, *22*(6), 97-107.

Leal-Taixe, L, Milan, A., Reid, I., Roth, S., & Schindler, K. (2015). MOTChallenge 2015: Towards a benchmark for multi-target tracking. arXiv:1504.01942

Lin, T.Y., Maire, M., Belognie, S., Hays, J., Perona, D., Ramanan, D., Dollár, P., & Zitnick, L. (2014) Microsoft COCO: Common Objects in Context. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8693. Springer, Cham.

Lindstrom-Forneri, W., Tuokko, H. A., Garrett, D., & Molnar, F. (2010). Driving as an Everyday Competence: A Model of Driving Competence and Behavior. *Clinical Gerontologist*, *33*(4), 283–297. https://doi.org/10.1080/07317115.2010.502106

M

Ma, C., Huang, J.B., Yang, X., & Yang, M.H. (2015, December). Hierarchical convolutional features for visual tracking. Paper presented at *IEEE International Conference on Computer Vision (ICCV).*

Malta, L., Ljung Aust, M., Faber, F., Metz, B., Saint Pierre, G., Benmimoun, M., Schäfer, R. (2012). *EuroFOT D6.4 Final results: Impacts on traffic safety.*

Martinussen, L. M., Møller, M., & Prato, C. G. (2014). Assessing the relationship between the Driver Behavior Questionnaire and the Driver Skill Inventory: Revealing sub-groups of drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *26*, 82-91.

Martinussen, L. M., Møller, M., Prato, C. G., & Haustein, S. (2016). How indicative is a self-reported driving behaviour profile of police registered traffic law offences. *Accident Analysis and Prevention*, *99*, 1-5.

McKnight, A., J., & Adams, B. B. (1970). *Driver Education Task Analysis, Volume II: Task Analysis Methods. Final Report.* Alexandria, VA: Human Resources Research Organization.

Michon, J. A. (1985). A critical view of driver behaviour models: What do we know, what should we do. In *Human behaviour and traffic safety* (pp. 485–520).

N

National Highway Traffic Safety Administration. (2016). *Distracted Driving 2014* (Summary of Statistical Findings No. DOT HS 812 260). Washington: U.S. Department of Transportation.

van Nes, N., Christoph, M., Jansen, R., Schagen, I., Verschragen, E., Wesseling, S., … Eenink, R. (2018). How you drive. Overview of main results of the UDRIVE project. Summary of the EU FP7 Project UDRIVE. (www.udrive.eu).

Ng Boyle, L., Tippin, J., Paul, A., & Rizzo M. (2008). Driver Performance in the Moments Surrounding a Microsleep. *Transportation Research Part F: Traffic Psychology and Behaviour*, *11* (2): 126–136.

O

Olsen, R.L., Hanowski, R.J., Hickman, J.S., & Bocanegra, J.L. (2009). *Driver distraction in commercial vechicle operations (report no FMCSA-RRR-09-042).*

Oxley, J., Fildes, B., Corben, B., & Langford, J. (2006). Intersection design for older drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *9*(5), 335–346. https://doi.org/10.1016/j.trf.2006.06.005

Özkan, T. and Lajunen, T. (2005). Multidimensional Traffic Locus of Control Scale (T-LOC): factor structure and relationship to risky driving. *Personality and Individual Differences, 38*(3), 533–545.

P

Parker, D., Stradling, S.G., & Manstead, A. (1996). Modifying beliefs and attitudes to exceeding the speed limit: An intervention study based on the theory of planned behaviour. *Journal of Applied Social Psychology*, *26*(1), 1-19. https://doi.org/10.1111/j.1559-1816.1996.tb01835.x

Precht, L., Keinath, A., & Krems, J.F. (2017). Identifying the main factors contributing to driving errors and traffic violations – Results from naturalistic driving data. *Transportation Research Part F*, *49*, 49-92.

R

Reason, J., Manstead, A., Stradling, S., Baxter, J., & Campbell, K. (1990). Errors on the roads: a real distinction? *Ergonomics, 33,* 1315-1332.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016, June). You Only Look Once: Unified, Real-Time Object Detection. Paper presented at *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR).*

Redmon, J., & Farhadi, A. (2017, July). YOLO9000: Better, Faster, Stronger. Paper presented at *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR).*

Reid, D. (1979). An algorithm for tracking multiple targets. *IEEE Transactions Automatic Control*, *24*(6), 843–854.

Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(6), 1137-1149.

S

Saberg, F., Selpi, Piccinini, G. F., & Engström, J. (2015). A review of Research on Driving Styles and Road Safety. *Human Factors*, *57*(7), 1248-1275.

Senserrick, T. M. (2006). Reducing young driver road trauma: guidance and optimism for the future. *Injury Prevention*, *12*(suppl_1), i56–i60. https://doi.org/10.1136/ip.2006.012773

Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & Lecun, Y. (2014). Overfeat: Integrated recognition, localization and detection using convolutional networks. In *International Conference on Learning Representations (ICLR2014), CBLS, April 2014* (http://openreview.net/document/d332e77d-459a-4af8-b3ed-55ba, http://arxiv.org/abs/1312.6229).

Sollins, B., Chen, D.-W., Reinerman-Jones, L., & Tarr, R. (2014). Truck Driving Distractions: Impact on Performance and Physiological Response. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 2171–2175. https://doi.org/10.1177/1541931214581456

Stanton, N. A., Walker, G. H., Young, M. S., Kazi, T., & Salmon, P. M. (2007). Changing drivers' minds: the evaluation of an advanced driver coaching system. *Ergonomics*, *50*(8), 1209–1234. https://doi.org/10.1080/00140130701322592

Summala, H. (1997). *Hierarchical model of behavioural adaptation and traffic accidents.* In: Rothengatter, T. & Carbonell Vaya, E. (Eds.) Traffic and Transport Psychology: Theory and Application. Elsevier Science, Oxford, 41-52.

SWOV (2016). *Speed and speed management*. SWOV Fact sheet, November 2016 , The Hague.

Volvo Trucks (2017). Volvo Trucks safety report 2017. Gothenburg.

T

Ting, P.-H., Hwang, J.-R., Doong, J.-L., & Jeng, M.-C. (2008). Driver fatigue and highway driving: A simulator study. *Physiology & Behavior*, *94*(3), 448-453.

W

af Wåhlberg, A. E. (2010). Social desirability effects in driver behavior inventories. *Journal of Safety Research*, *41*(2), 99-106.

af Wåhlberg, A. E., Dorn, L. and Kline, T. (2011). 'The Manchester Driver Behaviour Questionnaire as apredictor of road traffic accidents'. *Theoretical Issues in Ergonomics Science*, *12*(1), 66 – 86.

af Wåhlberg, A.E., Barraclough, P., & Freeman, J. (2015). The Driver Behaviour Questionnaire as accident predictor; A methodological re-meta-analysis. *Journal of Safety Research, 55*, 185-212.

Want, L., Ouyang, W., Wang, X., & Lu, H. (2015, December). Visual tracking with Fully Convolutional Networks. Paper presented at *IEEE International Conference on Computer Vision (ICCV).*

Wojke, N., Bewley, A. & Paulus, D. (2017). Simple Online and Realtime Tracking with a Deep Association Metric. University of Koblenz-Landau, Queensland University of Technology. arXiv:1703.07402, 2017.