Road Safety Data, Collection, Transfer and Analysis

Naturalistic Driving for monitoring safety performance indicators and exposure: considerations for implementation

Deliverable 6.4

Please refer to this report as follows: Van Schagen, I. and Reed, S. (2012) Naturalistic Driving for monitoring safety performance indicators and exposure: considerations for implementation. Deliverable 6.4 of the EC FP7 project DaCoTA.

Grant agreement No TREN / FP7 / TR / 233659 /"DaCoTA"
Theme: Sustainable Surface Transport: Collaborative project
Project Coordinator:
Professor Pete Thomas, Vehicle Safety Research Centre, ESRI
Loughborough University, Ashby Road, Loughborough, LE11 3TU, UK

Project Start date: 01/01/2010 **Duration 36 months**

Organisation name of lead contractor for this Deliverable: SWOV Institute for Road Safety Research, The Netherlands			
Report Author(s): Ingrid van Schagen (SWOV) and Steven Reed (Transport Safety Research Centre (TSRC) at Loughborough University)			
Due date of Deliverable	**31/12/2012**	**Submission date:**	**31/12/2012**

Project co-funded by the European Commission within the Seventh Framework Programme	
Dissemination Level:	
PU	**Public**

Project co-financed by the European Commission Directorate General for Mobility and Transport

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

CAN:	Controller Area Network
DAS:	Data Acquisition System
EEG:	Electro-Encephalography
ERSO:	European Road Safety Observatory
EC:	European Commission
EU:	European Union
FOT:	Field Operational Test
GPRS:	General Packet Radio Service
GPS:	Global Positioning System
IMU:	Inertial Measurement Unit
ND:	Naturalistic Driving
OBD:	On-Board Diagnostics
RED:	Risk Exposure Data
RFID:	Radio-frequency identification
SD:	Standard Deviation
SHRP2:	Strategic Highway Research Program 2
SPI:	Safety Performance Indicator
USA:	United States of America

EXECUTIVE SUMMARY

DaCoTA was a Collaborative Project under the European Seventh Framework Programme that aimed to develop tools and methodologies to support road safety policy and further extend and enhance the European Road Safety Observatory (ERSO). One of the Work Packages in DaCoTA, WP6, focused on the usefulness and feasibility of applying the Naturalistic Driving method for collecting comparable information about the road safety level in EU Member States and its development over time. The current Deliverable was prepared in this framework and gives an overview of the aspects to be taken into account when implementing ND research for monitoring purposes.

Naturalistic Driving (ND)
Naturalistic Driving (ND) can be defined as “A study undertaken to provide insight into driver behaviour during every day trips by recording details of the driver, the vehicle and the surroundings through unobtrusive data gathering equipment and without experimental control”. Typically, in an ND study passenger cars, preferably the subjects' own cars, are equipped with several small cameras and sensors. These devices continuously and inconspicuously register vehicle manoeuvres (like speed, acceleration/deceleration, direction, location), driver behaviour (like eye, head and hand manoeuvres), and external conditions (like road, traffic and weather characteristics).

ND for monitoring purposes
ND data can, among other things, be used to establish how often drivers routinely are exposed to or engaged in certain situations/behaviours that are known to increase the risk of a crash. This includes monitoring safety-relevant behaviour (Safety Performance Indicators - SPIs) and mobility (Risk Exposure Data – RED). An important reason for monitoring road safety and comparing road safety levels and their developments over time in different countries is benchmarking. It allows countries to determine their relative position in comparison to other selected countries, to understand differences and find ways and get motivated to improve their position. Obviously, monitoring road safety also allow countries to evaluate their own road safety policy and road safety targets. ND is considered a promising approach for collecting reliable and comparable information about various RED and SPIs, as well as several relevant context variables. The main advantage of the ND approach for monitoring purposes as compared to the more traditional SPI data collection methods, such as road-side surveys and questionnaires, is that ND ensures continuous, automatic and standardized data collection. Provided that similar data acquisition systems and methods are applied in all participating cars, this approach substantially increases international comparability and level of detail. Though the current Deliverable is purely focused on road safety and exposure data, the collected data will also be useful for other transport areas, in particular eco-driving and traffic management.

Three data collection scenarios
Depending on the variable of interest, ND data collection needs different technologies ranging from simple and relatively cheap data acquisition systems to more sophisticated systems with several sensors as well as several videos covering the inside of the car and various directions outside the car. By combining the RED and SPIs of interest and the technological requirements for collecting that type of data, we distinguish three scenarios to collect meaningful data within reasonable limits of cost and complexity.

It is recommended to start off with Scenario 1: a low-cost simple, off-the-shelf simple data acquisition system (e.g. an OBD GPS tracker or a Smart Phone) and a limited number of additional sensors, measuring:

- Vehicle kilometres
- Person kilometres
- Number of trips
- Time in traffic
- Speed
- Seat belt use
- Light use

In a later stage, additional SPIs and network characteristics could be added successively (Scenario 2), including:

- Time headway
- Acceleration
- Lane departures
- Inappropriate speed
- Signal use
- Junction type

SPIs that would need continuous external and/or internal video recordings do not seem to be feasible in the short term, because this results in huge amounts of data and very high costs for the related data transfer and data coding. That means that the SPIs like fatigue, inattention, distraction and the (proper) use of child restraints can currently not be monitored by means of ND research. Technical developments may allow reconsideration of this conclusion in due time.

Furthermore, it is recommended to equip a limited number of cars also with an event-triggered video in order to monitor numbers of near crashes as yet another relevant SPI (Scenario 3). As a very useful side product, this effort will provide data that can be used to further specify and refine the quantitative and qualitative relationship between near crashes and real crashes.

Study design and organisational issues

In principle, the techniques and procedures for ND data collection, data transfer, data storage and data analysis are available and not too complicated. In order to get reliable information, a fairly large sample is needed. The exact size of the sample depends on the variation in behaviour in the population and the required level of precision of the results. Assuming that the sample is drawn in a cleverly stratified way, resulting in a number of mutually exclusive and homogeneous subgroups (e.g. based on gender and age), a sample of around 10,000 drivers per country seems to be required for RED such as the annual amount of vehicle kilometres. This number is usually independent of the size of the population of car drivers in a country. Only if the sample size is larger than 10% of the population, a correction is applicable.

It is recommended to collect data throughout the year on a continuous basis and to follow each individual in the sample for one year, applying a rotating scheme of 50% per 6 months. Analyses are best performed at a national level, applying a series of definitions on variables and disaggregation levels and following fixed analysis protocols. It could be considered to identify a limited number of 'core' variables (SPIs/RED) to be analysed at a central/ERSO level to ensure exact comparability.

Participation in ND research is per definition on a voluntary basis and experiences in the USA and Europe have shown that it is requires special attention to find sufficient suitable participants, especially if there are strict sample stratification requirements. In addition, there are legal and ethical issues involved in ND research, in particular in the area of privacy and data protection.

Exploring Scenario 4
In parallel to the implementation of the previous three scenarios, it is recommended to start exploring the possibility of a Scenario 4 now, i.e. a scenario where relevant data is extracted directly from the vehicle via CAN-bus, OBD, and other trip and travel data collected automatically by the vehicle. In theory, that way a lot of relevant information is already available with no or little additional costs; in practice, however, the information is not generally accessible nor comparable between car makes and models. So, this is a scenario that cannot be realised overnight. One of the first steps, in consultation with the car manufacturers, is an elaboration of the requirements for this data: what is available, what is needed, what is technically feasible.

A central role for the EC
Since harmonisation and international comparability of data are the key reason for this effort, there is a central role for the European Commission in initiating this task and taking the lead from here, most likely within the ERSO framework. A stepwise approach is recommended, including successively:

1. Creating support and finding budget by presenting the case to the relevant road safety bodies at European and Member State level, explaining the need for harmonised, comparable international data, the ND approach, and its added value.
2. Preparing a detailed description of / handbook for all practical implementation aspects, including the functional specifications of data collection equipment, participant selection, data transfer and storage, as well as definitions of variables, disaggregation levels and analyses.
3. Identifying the relevant national organisations which will be responsible for national data collection and pre-analyses, and fine-tuning data collection procedures (including legal aspects) and variable definitions in consultation with them.
4. Developing and equipping a database at EU level and defining the required (pre-analysed aggregated) data to be provided by the Member States as well as the procedures and time schedule, in consultation with the relevant national organisations.
5. Setting up European-wide communication strategies to guarantee maximum dissemination and use of the collected data.
6. Setting up one year national pilots in at least four Member States, well spread of Europe (North, West, South, East).
7. Adapting procedures and definitions, based on the pilot experiences.
8. Successive implementation of Scenario 1 in additional Member States.

Parallel to steps 6 and 7, Scenario 2 (additional SPIs/RED) and 3 (monitoring near-crashes) can be elaborated, piloted and implemented, applying a similar stepwise process.

From the very beginning, the EC is advised to initiate discussions with the car manufacturers, using existing discussion platforms, with the aim to explore longer term possibilities of Scenario 4, i.e. the scenario where relevant data is extracted directly from the vehicle.

1. INTRODUCTION

1.1. The DaCoTA project

DaCoTA was a Collaborative Project under the Seventh Framework Programme, co-funded by the European Commission DG Mobility and Transport. It ran from January 2010 until December 2012. DaCoTA stands for Road Safety **Da**ta **Co**llection, **T**ransfer and **A**nalysis and it developed tools and methodologies to support road safety policy and further extend and enhance the European Road Safety Observatory (ERSO)[1] developed within the preceding SafetyNet project[2].

ERSO was created with the aim of being the primary focus for road safety data and knowledge for policy makers at European and national/regional level, acknowledging that data and knowledge are essential for making well-founded, evidence-based policy decisions. Monitoring developments over time in Europe as a whole and comparing the state of affairs and developments between Member States is one approach for identifying opportunities for further road safety improvements.

1.2. WP6: Naturalistic Driving

One of the Work Packages in DaCoTA, WP6, focused on the usefulness and feasibility of applying the Naturalistic Driving method for collecting comparable information about the road safety level in EU Member States and its development over time.

Naturalistic Driving (ND) can be defined as "A study undertaken to provide insight into driver behaviour during every day trips by recording details of the driver, the vehicle and the surroundings through unobtrusive data gathering equipment and without experimental control" (Van Schagen et al., 2011). Typically, in an ND study passenger cars, preferably the subjects' own cars, are equipped with several small cameras and sensors. These devices continuously and inconspicuously register vehicle manoeuvres (like speed, acceleration/deceleration, direction, location), driver behaviour (like eye, head and hand manoeuvres), and external conditions (like road, traffic and weather characteristics).

ND research can be used in different ways. So far, ND studies have been mainly applied to get an in-depth understanding of road user behaviour in interaction with the vehicle, the road and other road users and the relationship with (near) crashes as well as the risk of specific road user behaviours, e.g. mobile phone use. Well-known examples are the 100-car study in the USA (Dingus et al., 2002), and the current large-scale ND study in the SHRP2 framework as its follow-up. In Europe, the feasibility of ND research for providing road safety data has been studied in the PROLOGUE project[3]. This has resulted in the recently started UDRIVE project, the smaller-scale equivalent of the SHRP2 study in the USA[4].

The ND principles are also regularly applied in Field Operational Tests (FOTs). An FOT is "a study undertaken to evaluate a function, or functions, under normal operating conditions in environments typically encountered by the host vehicle(s) using quasi-experimental methods" (FESTA, 2011). Or, in other words, an FOT applies the ND methodology to evaluate new vehicle technologies in real life

---

[1] ERSO: http://ec.europa.eu/transport/road_safety/index_en.htm
[2] SafetyNet: http://ec.europa.eu/transport/wcm/road_safety/erso/safetynet/content/safetynet.htm
[3] PROLOGUE: http://www.prologue-eu.eu/
[4] UDRIVE: www.udrive.eu

conditions to see how drivers use the technology and if and how it affects their driving behaviour.

DaCoTA focused on a third possibility for using ND research: establishing how often drivers routinely are exposed to or engaged in certain situations/behaviours that are known to increase the risk of a crash. This includes monitoring safety-relevant behaviour (Safety Performance Indicators - SPIs) and mobility (Risk Exposure Data – RED). The main aim is to enable comparisons between countries and to assess developments over time. This would be a new area of application for ND research with specific requirements for technology, methodology and organisation. WP6 of DaCoTA aimed to define these requirements and to identify the possibilities and limitations of this specific application area as well as the practical consequences for the implementation of a European-wide ND monitoring study. DaCoTA will not focus on establishing the crash risk of engaging in certain behaviours, instead the focus will be on determining how often these situations and behaviours occur during everyday trips.

1.3. This Deliverable

The current Deliverable gives an overview of the aspects to be taken into account when implementing ND research for monitoring purposes. It elaborates on, and partly summarizes, the findings of the preceding tasks in this Work Package:

- An inventory of relevant variables to monitor road safety, the required measurement tools and their feasibility (D6.1: Talbot et al., 2010).
- An overview of requirements for study design and methodology (D6.2A: Bonnard et al., 2012) with special attention to the issue of sampling techniques (D6.2B: Commandeur, 2012).
- Two small-scale feasibility studies (D6.3: Pilgerstorfer et al., 2011).

Chapter 2 of the current Deliverable first briefly elaborates on the value of monitoring road safety by means of SPIs and RED and the added value of the ND approach as compared to current methods. Secondly this Chapter presents an overview of the theoretically relevant variables (SPIs and RED).

Based on general technological, organisational and cost considerations, Chapter 3 presents a three scenario approach that with increasing complexity and costs will result in an increasing amount of information.

Chapter 4 and Chapter 5 present and discuss various practical issues when implementing ND for monitoring purposes. Chapter 4 focuses on the critical issues in study design, including issues of length of data collection and data collection period(s), sample size, and sample representativeness. Chapter 5 discusses issues in relation to data transfer and storage, data analysis, participant recruitment, as well as the legal and ethical issues that have to be taken care of. Chapter 6 briefly explores the options of a Scenario 4 approach where information is extracted directly from all vehicles rather than adding equipment to a sample of vehicles. Chapter 7 finally, summarizes the main findings and recommendations for future ND data collection for monitoring purposes.

2. MONITORING SAFETY PERFORMANCE INDICATORS AND EXPOSURE

An important reason for monitoring road safety and comparing road safety levels and their developments over time in different countries is benchmarking. It allows countries to determine their relative position in comparison to other selected countries, to understand differences and find ways and get motivated to improve their position. Obviously, monitoring road safety also allow countries to evaluate their own road safety policy and road safety targets. But why monitoring safety performance indicators (SPIs) and risk exposure data (RED) and what, then, is the added value of naturalistic driving research?

2.1. SPI's: cross-national monitoring of road safety

In principle, the most direct way for monitoring road safety developments is to look at the number of road traffic fatalities and injuries. However, there are several reasons why this indicator is only of limited value for comparing the safety level between countries. First, the number of crashes and casualties is dependent on the mobility in a country. In theory, cross-national comparisons could then be based on risk, i.e. the number of crashes/casualties divided by exposure, usually the number of kilometres driven. However, reliable and comparable exposure measures in EU countries are lacking. Second, direct comparisons are impossible because there are substantial differences between countries in definitions of crashes and casualties, in particular of road injuries. Third, there are indications that registration levels generally decrease in Member States, but at different speed. Again this makes it impossible to make useful comparisons between countries. And last, but not least, when it comes to understanding the reasons for different road safety levels, fatality and injury numbers only tell part of the story.

This all means that we would need to find other relevant variables that allow for meaningful comparisons of safety levels. The road safety pyramid in *Figure 2.1* shows that safety performance indicators (SPIs) can be a good alternative for this.

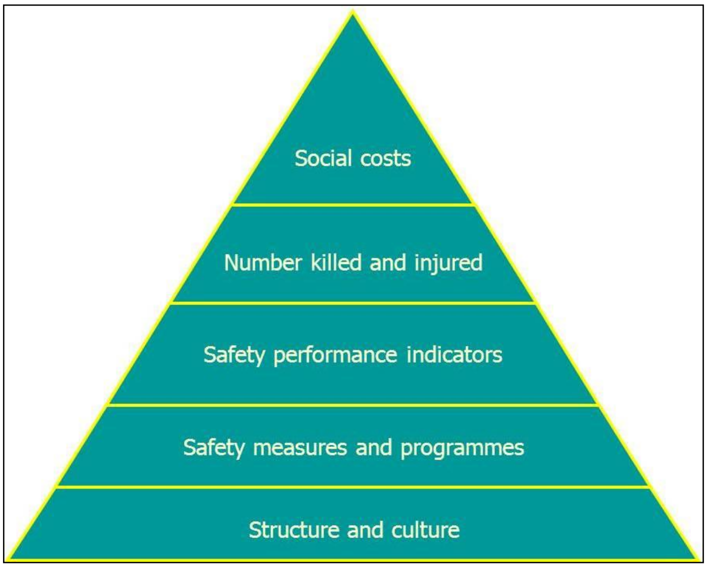

Figure 2.1 *The SUNflower road safety pyramid (From: Koornstra et al., 2002; based on LTSA, 2000).*

The pyramid shows that the road safety level of a country is related to, at the bottom of the pyramid, the structure and culture of that country, which in turn, relate to the safety measures and programmes being taken.

Subsequently, these measures and programmes generally focus at and affect the relevant SPIs such as the safety quality of the road infrastructure, the vehicle fleet, and road user behaviour. Also the number of near crashes can be considered as an SPI. Finally, in their turn, the SPIs are related to the number of killed and injured road users and subsequently to the social costs of road safety.

Hence, measuring SPIs could be an alternative indication of the road safey level of a country; and when performed in a comparable way, both over time and between countries, a useful tool for cross-national monitoring.

2.2. RED: cross-national monitoring of exposure

When monitoring developments over time in terms of SPIs or when comparing the status of SPIs in different countries, it is crucial to have reliable information about the developments and cross-national differences in mobility or exposure to risk. This makes it possible to relate different road safety aspects to one another. If, for example, changes have been found in the number of speed violations on motorways or the number of cars with daylight running lights on rural roads, it is important to link these findings to the car mobility. Are the identified changes due to a decrease or increase in the number of cars on the relevant type of road or is there indeed a change in behaviour. For a correct interpretation of this type of findings, comparable risk exposure data (RED) is needed.

2.3. The added value of ND

As indicated there are several reasons why, for international comparisons, measuring safety performance indicators (SPIs) can be a useful addition or even an alternative for recording numbers of road crashes or road casualties. In addition, as explained in Section 2.2, risk exposure data (RED) is needed for useful interpretation of cross-national differences of both crash and injury data and SPIs. A prerequisite is, of course, that the SPIs and RED are defined and measured in a comparable way. An additional prerequisite for the observed SPIs is that they are relevant for safety, i.e. that they have a direct and proven causal relationship with safety outcomes (ETSC, 2001; Hakkert, Gitelman and Vis, 2007). Currently, many countries have no or limited SPI data and RED, and the data that is available is often not comparable.

Naturalistic Driving is considered a promising approach for collecting reliable and comparable information about at least some RED and SPIs. The main advantage as compared to the more traditional SPI data collection methods, such as road-side surveys and questionnaires, is that ND research ensures continuous, automatic and standardized data collection. Provided that similar data acquisition systems and methods are applied in all participating cars, this approach substantially increases international comparability and level of detail.

The example of the SPI Speed can illustrate this. Speed has a proven relationship with road safety (Aarts and Van Schagen, 2006; Elvik, 2009) and as such can be considered as a useful SPI. Many countries collect speed data, but there is a wide variety of methods: all types of roads or just on some types of roads, continuous or for limited periods, in speed classes or specific, for all vehicle types or distinguishing between vehicle types. Already a fairly simple data acquisition system mounted in a car can continuously monitor speeds as well as, assuming that it is linked to GPS, for various road types and transfer this to a central database. If applied in all EU countries, and applied to a sufficiently large sample (see Chapter 4) this approach provides reliable, comparable and detailed information about, for example, average speeds and differences in speed between vehicles on different road types.

Another example, in the area of RED, is the number of kilometres driven by a person. Generally, this type of information is collected through interviews or questionnaires, based on a person's recollection or willingness to report in detail. Again, a fairly simple data acquisition system applied to a sufficiently large sample, would be able to objectively monitor mobility, even per road or per type of road if connected to a GPS. This would, for example, provide much more insight in the risk of a crash on different types of roads and the developments over time. Currently, this type of information is lacking in almost all EU Member States.

2.4. Theoretically relevant RED and SPIs

Deliverable 6.1 of DaCoTA (Talbot et al., 2010) provides an overview of the theoretically relevant RED and SPI, largely based on the work in the SafetyNet project (e.g. Hakkert, Gitelman and Vis, 2007; Hakkert and Gitelman, 2007) supplemented with some literature research and an expert survey on additional relevant SPIs. Given the nature of ND research, only those RED and SPIs that have to do with the behaviour of individual drivers are relevant. Furthermore, as pointed out in SafetyNet, several background or context variables need to be collected and registered to allow for valid comparisons across countries.

Table 2.1 gives and overview of the most relevant RED, SPIs and context variables that, in principle, can be collected by ND as identified by Talbot and her colleagues.

Table 2.1 *Overview of relevant RED, SPIs and context variables for ND data collection (Source: Talbot et al., 2010)*

Risk Exposure Data (RED)	Safety Performance Indicators (SPI)	Context variables
- Vehicle kilometres - Fuel consumption - Person kilometres - Number of trips - Time in traffic	- Alcohol and drugs - Speed - Protective systems - Daytime Running Lights - Fatigue - Distraction/inattention - Gap acceptance/headway - Near crashes - Crash causation - Safety systems	- Driver variables: e.g. age, gender - Vehicle variables: e.g. vehicle age, make, model - Network variables: e.g. road type, area type, speed limit - Other contextual variables (transient): year, month, day, hour

Most of the variables are self-explaining. Some words, however, need to be said about the SPI "near crashes", since this is not a single behaviour, but the outcome of a series of behaviours and events. The reason for considering near crashes as an SPI is that they are far more frequent than real crashes while they are generally assumed to be related to real crashes. That would mean that the number (and situation) of near crashes would be a good indicator of the number (and situation) of actual crashes. It must be noted, however, that the exact definition of a near crash is not yet clear and the relationship between near crashes and real crashes may not be as straightforward as suggested. This could be an issue for current ND research projects such as UDRIVE[5] and SHRP2[6].

Crash causation was one of the variables identified as interesting and relevant by experts. Whereas indeed crash causation factors are very relevant for road safety research and might even help to identify new SPIs, accident causation in itself is not an SPI, i.e. a variable that can be used as a proxy for the road safety level in a country. Therefore, crash causation is not further discussed in the current report.

2.5. In conclusion

Cross-national comparisons of road safety levels allow countries to determine their relative position in comparison to other selected countries, and to identify ways to improve their position. Monitoring road safety through safety performance indicators (SPIs) can be a useful addition or even an alternative for recording numbers of road crashes or road casualties. Risk exposure data (RED) is needed for useful interpretation of cross-national differences.

ND research is of course not the panacea for all SPI and RED data collection. Not all SPIs and RED that are theoretically relevant are suitable for being monitored by means of ND and some need a more advanced and more costly data acquisition system than others. This will be discussed in more detail in Chapter 3. Similarly, it is very important that there are sufficient drivers participating in the study and that they can be considered representative for the overall population in terms of, for example, age, gender and region. This will be discussed in more detail in Chapter 4. But, given the overall characteristics of ND research, it can be concluded that for some SPIs and RED this approach could bring international data collection at a much higher level.

[5] www.udrive.eu

[6] http://www.trb StrategicHighwayResearchProgram2SHRP2/Pages/The-SHRP-2-Naturalistic-Driving-Study-472.aspx

3. OPTIONS FOR ND DATA COLLECTION

3.1. Two basic scenarios

Depending on the variable of interest, ND data collection needs to use different technologies ranging from simple and relatively cheap data acquisition systems (e.g. a Smart Phone-like application) to more sophisticated systems with several sensors as well as several videos covering the inside of the car and various directions outside the car.

When considering the options for data collection equipment from both a technological and organisational point of view, the main characteristics of ND research for monitoring purposes must be taken into account. These are (further elaborated in Chapter 4):

- large number of cars in each of the participating countries;
- continuous or at least long periods of data collection;
- unobtrusive data collection not requiring input of participants; and,
- given the huge amount of generated data, simple and largely automatic data processing and analysis

By combining the RED and SPIs of interest (*Table 2.1*) and the technological requirements for collecting that type of data, Talbot et al. (2010) distinguish two scenarios to collect meaningful data within reasonable limits of cost and complexity.

<u>Scenario 1</u> would use the most basic data acquisition system comprising of a GPS logger, and using existing, off-the-shelf technology. One possibility would be using Smart Phone, another would be and OBD-GPS tracker. For processing the data in a meaningful way, Scenario 1 would at least include map matching facilities and a digital speed limit map. As specified by Talbot and colleagues, Scenario 1 would allow collecting most of the relevant RED:

- Vehicle Km
- Person Km
- Number of Trips
- Time in Traffic

With respect to identified relevant SPIs, Scenario 1 allows collecting data on:

- Excessive speed (i.e. exceeding the posted speed limit)

The latter would assume map matching tools, and availability of and access to a digital map with speed limits.

<u>Scenario 2</u> would use the Scenario 1 system with some relatively simple additional sensors or information sources for monitoring some extra variables. Given current problems with access to vehicle brand specific Controller Area Network (CAN) data, this is not considered as a feasible additional data source.

In addition to the Scenario 1 data, the following topics mentioned in Table 2.1 could be monitored with Scenario 2 DAS (between brackets the required addition):

RED:

- Fuel consumption (sensor)

SPIs:

- Alcohol (sensor, but not reliably)
- Protective systems: seat belt use (sensor)
- Daytime Running Lights: light use in general (sensor)
- Gap acceptance/headway: time and/or distance headway (radar)
- Inappropriate speed, inappropriate for the conditions, based on indirect measures like windscreen wiper use (sensor), outside temperature (sensor) and time/distance headway (radar).

Additional variables, not mentioned in *Table 2.1* but relatively easy to monitor in Scenario 2, are:

- Acceleration (e.g. by an Inertial Measurement Unit (IMU) sensor)
- Lane departures - in combination with headway as indication of overtaking behaviour (sensor)
- Signal use (sensor)

It must be noted, however, that it may not always be best to just answer the need for a sensor by adding a sensor. In some applications this approach works very well, but it is sometimes best to look at the bigger picture and balance the pros and cons of a more integrated approach, where installation of sensors is combined.

Talbot et al. also include in-vehicle technology/safety systems in use (e.g. ABS and ESC) as a variable to be studied in Scenario 2. However, the main input for this would be CAN data. As already indicated, getting access to this type of CAN data of all car makes and models is not feasible in the short term. Therefore, we suggest for the time being to skip this variable from the Scenario 2 approach. Obviously whether a car does have such safety system can be deduced from make, model and age.

3.2. A third scenario

It can be concluded that all relevant RED can be measured by either a Scenario 1 or a Scenario 2 data acquisition system. This is not the case for the SPIs. The following relevant SPIs require a more advanced data acquisition system:

- Drugs use, requiring manual administration of drug detection equipment.
- Child restraint use, requiring video of passenger front and back seats.
- Fatigue, requiring continuous monitoring of brain activity (electro-encephalography - EEG) or video monitoring of eyes/head.
- Distraction/inattention, requiring continuous video monitoring of head and hands of driver.
- Near crashes, requiring (event-triggered) video recordings of driver and external conditions.

Per definition, drugs use cannot be applied in an unobtrusive, naturalistic way since it requires manually operated drugs testers. Similarly, current techniques do not allow for unobtrusive measurements of fatigue by means of EEG. Even though there has been a long research tradition aiming to identify behavioural indications of fatigue (e.g. steering wheel movements or correction), so far these indications are insufficiently validated to use as a reliable SPI for fatigue[7].

Unobtrusively monitoring inattention, distraction and fatigue would be possible by having one or more targeted video cameras in the car as was shown in several studies in the USA (e.g. Hanowski et al., 2009; Klauer et al., 2006). In theory, video cameras can also be used to identify drug use by looking at pupil size. However, for getting meaningful information continuous recording would be required. When doing that at large-scale, as required for monitoring purposes, this would result in immense, hardly imaginable amounts of data. As an indication, just 1 hour of video requires 1 to 3 GB of data storage, depending upon the resolution of the video (Talbot et al., 2010). As a consequence, this would require very frequent active rather than automatic data transfer with corresponding burden for the participant, huge data storage capacity and very time consuming data coding efforts. In short, continuous video recording is currently considered to be too impractical and too expensive to implement on a large scale. Technical developments may allow reconsideration of this conclusion in due time.

Near crashes could be monitored on a video event-triggered base. In case of an event, as automatically identified by means of harsh braking, accelerating, or steering movements (thresholds to be defined), the video data of the period just before and after the event is saved; otherwise it is overwritten. In combination with RED, near crashes can be a useful SPI to compare countries and developments over time.

The video data is necessary to minimise false alarms and false positives and give a more reliable picture of the number of real near crashes. In addition, the video data gives background information about the seriousness of the near-crash and the road and traffic circumstances. As a minimum, this requires a video camera with a fairly wide-angled view in front of the vehicle. The video data can be used for research purposes as well, studying the circumstances of near crashes.

Therefore, in the Scenario 3 data acquisition system, it is proposed to consider an additional event-triggered forward view video with the aim of monitoring the number of near crashes. It must be noted, however, that an event-triggered video does not give the full picture; it does not provide information about near crashes that were not preceded by an action (braking, steering correction) of the driver, but where a crash was avoided because of an action of another road user.

[7] ND research could be a useful approach for validating these behavioural indications of fatigue, but this is outside the scope of the DaCoTA project.

3.3. Monitoring the context variables

Chapter 2 identified several context variables about driver, vehicle, network and time/date that need to be collected. Driver and vehicle variables data is collected once-only at participant level and do not lead to specific requirements of the data acquisition system. One important aspect here is that the driver of a participating vehicle is recognised as the participant and is not a member of the family or a friend driving the car. Driver identification is required in each DAS scenario, also in the basic one. As pointed out by Talbot et al., driver identification can be best achieved through the use of video. Only a few seconds of video of the driver as the vehicle sets off are sufficient and advances in machine vision technology mean that we will not have to rely on manual analysis to identify drivers. However, for now, this option would still be too expensive for Scenario 1 data collection. For Scenario 1, a more realistic alternative is to provide each driver with a magnetic swipe card to use at the beginning of each journey or an RFID tag (Radio-frequency identification) with a receiver in the vehicle that could register the driver ID. As Talbot et al. point out, neither are 100% reliable as they depend upon the drivers carrying a device with them and/or using it every time they drive. In addition, using the swipe card reminds the driver of his/her participation in each trip, which undermines the unobtrusive character of the ND approach. Just as an indication, small-scaled trials showed that typically in around 15% of the trips someone else than the participant drove the car Pilgerstorfer et al., 2011).

The time and date context variables can be automatically logged and saved without additional effort, in combination with the GPS logger in the Scenario 1 DAS. In order to be able to say something about excessive speed, Scenario 1 would also need the possibility to link GPS information to a speed limit database.

Monitoring the relevant network variables is more complicated. General characteristics such as road type (urban, rural), road classification (motorway, arterial roads, collector roads, etc.) and intersection type (junction, roundabout, etc.) are best retrieved based on GPS and subsequent map matching and a road information database. As such, this type of more detailed information would be available with a Scenario 2 DAS. It should be kept in mind, however, that currently, countries may differ substantially in the type and definition of road classifications. Also, the information in the road information database, and the availability and reliability of the information differ between countries. For comparisons between countries, probably only some major road categories can be distinguished.

In some countries, more detailed information about road geometry (e.g. number of lanes, lane width), road signs and road marking may be available in a road information database. Otherwise it would require continuous external video, which is considered to be unsuitable for large-scale monitoring purposes, or specialist sensors. The latter could be added to Scenario 2, once sufficiently reliable sensors are available.

3.4. In summary: DAS scenarios 1, 2 and 3

Table 3.1 provides a summary overview of the data that can be collected by means of a Scenario 1, a Scenario 2, or a Scenario 3 data acquisition system. Though alcohol consumption was mentioned as an SPI that could be measured with a Scenario 2 data acquisition system, it is not included in the summary overview. There are sensors that unobtrusively measure the presence of alcohol in the air of a vehicle, however, it will not be clear whether it was the driver, the passenger or both who had drunk the alcohol, making it unsuitable as an SPI.

The Scenario 1 and Scenario 3 DAS are both all-or-nothing scenarios. When choosing for either one of these two scenarios, all variables in that scenario are included. The Scenario 2 DAS, on the other hand, can be best considered as an 'eclectic' scenario, i.e. adding just the currently 'easy' and 'cheap' sensors and information sources, and adding variables once the required sensors and additional information sources become better suitable in terms of quality and costs in the various EU Member States and the variable is considered important.

Table 3.1 *RED, SPIs and context variables to be collected with Scenario 1 (Smart Phone-like application), Scenario 2 (additional sensors/information sources) or Scenario 3 (additional forward view event-triggered video) ND data acquisition system (DAS).*

	Scenario 1 GPS logger	Scenario 2 plus additional sensors/data sources	Scenario 3 plus event-triggered video
Risk Exposure Data (RED)	- Vehicle Km - Person Km - Time in Traffic - Number of Trips	- Fuel consumption	
Safety Performance indicators (SPIs)	- (Excessive) speed	- Inappropriate speed - Seat belt use - Light use - Headway - Acceleration - Lane departure - Signal use	- Near crashes
Context variables	- Driver identification - Time/date - Location - Speed limits	- Network characteristics	

3.5. Variable definitions and disaggregation levels

Even when it has been decided which variables to measure, several additional decisions have to be made as how exactly to define and operationalize them and what level of disaggregation to allow for. Especially for ND research for cross-national monitoring it is essential that definitions and aggregation levels are clearly defined and applied.

As part of DaCoTA, Bonnard et al. (2012) have elaborated this type of issues and we refer to that report for details. Just as an example, to illustrate the type of decisions to be made at that level, some elaboration in relation to RED are described here. Bonnard et al. further elaborated the Scenario 1 RED indicators (see *Table 3.1*) as follows:

- Vehicle Km: Mean kilometres driven by participating passenger cars during one year in a country

- Person Km: Mean kilometres driven by participating drivers at the wheel of their car during one year in the country
- Time in traffic: Mean time spent by the participating drivers at the wheel of their car during one year in the country
- Number of trips: Mean number of trips made by the participating drivers at the wheel of their car during one year in the country
- Characteristics of trips made by the participating drivers at the wheel of their cars during one year in the country: mean and standard deviation of distance, duration and speed.

It must be kept in mind, as correctly pointed out by Bonnard et al., that the ND method does not measure all exposure in a country, but just focuses on:

- Trips with passenger cars, excluding trips with other modes of road transport.
- Trips of drivers of passenger cars, excluding people who do not drive.
- Trips as driver of the participating (instrumented) car, excluding trips as passenger or as driver of another car.

When zooming in on the RED in relation to trip, a next question pops up, namely what exactly must be considered a trip. Bonnard et al. propose to define a trip as the period between switching on and switching off the vehicle's engine. This is not a watertight definition, since drivers may switch off the engine when waiting for a bridge or a tunnel, or they may have a short break for coffee or a stop for filling the fuel tank. These would then be considered as two separate trips, whereas in fact it is just one. A solution for this may be to consider two 'trips' as one if the time between the end of the first and the beginning of the second is *short*. Obviously, then, it has to be defined what is considered to be short. The other way around, the definition based on switching on and off the engine may result in aggregating two or more trips with different aims and characteristics, e.g. dropping a child at school before going to work. So far, there is no straightforward solution for this bias.

For monitoring purposes, continuing with the RED example, it will be interesting to disaggregate the data to allow for identifying differences or different developments in exposure between e.g. day and night, different days of the week, at different road classes, for different age groups, etc. (See *Table 3.2* for some examples). Also this type of variables must be exactly defined and applied in the same way in the various countries.

Table 3.2 *Possible variables for disaggregation and their categories for RED (Source: Bonnard et al., 2012)*

Driving situation characteristics	• Road type (e.g. urban, outside urban area, motorway) • Hour and period of the day (e.g. dawn, daytime, dusk, night-time) • Day and period of the week (e.g. week, week-end) • Month and period of the year (spring, summer, autumn, winter) • Weather condition (e.g. clement, adverse) • Presence or not of passengers
Trips characteristics	• Duration of the trip (e.g. inferior to 20 minutes, between 20 minutes and 60 minutes, superior to 60 minutes) • Local mobility or far distance mobility (e.g. trips included or not in a 80 km area around the participant's home) • Regularity of the trip (e.g. done more than 10 times a year)
Vehicle characteristics	• Vehicle type = passenger car • Vehicle age • Vehicle engine size
Driver characteristics	• Age • Gender • Driving experience • Occupation • Home location • Country

3.6. Recommendations

As a first step towards an EU cross-national SPI and RED data collection effort it is recommended to start off with a low-cost, simple Scenario 1 DAS plus a limited number of additional variables from the Scenario 2, notably

- Seat belt use (simple sensor + logging)
- Light use (simple sensor + logging)
- Acceleration (IMU sensor)
- Network variables: distinguishing between a limited number of well-defined road classes

One of the first variables to add once the equipment becomes less expensive is headway. To measure this, specific equipment is required since the distance and time gap between the lead and following vehicle continually changes and requires continuous recalculation. Measuring headway can be done by either radar or machine vision, but both are currently considered to be too expensive for large-scale application.

Furthermore, it is recommended to equip a limited number of vehicles with a Scenario 3 DAS, to develop the technical conditions for monitoring near crashes as well as the knowledge about the correct definition (in terms of indicators and thresholds) for near crashes and their relationship with real crashes.

4. STUDY DESIGN FOR ND MONITORING

This Chapter will discuss some important issues related to the design of an ND study that aims to collect data for monitoring purposes and identifying differences between countries and developments over time. The next sections, successively, present considerations about the sample size, sample representativeness and length and period(s) of data collection, mainly based on the work of Commandeur (2012).

4.1. Sampling

As a starting point, it must be kept in mind that the ND monitoring effort discussed in WP6 of DaCoTA is meant to give a reliable picture of the status of the defined variables in the passenger car drivers in each of the EU Member States. Since it is impossible to study all car drivers, a sample must be drawn. Therefore, the sampling frame, i.e. the source from which a sample is drawn, can be defined as: all passenger car drivers in a country.

Subsequently, the sample can be drawn in a (simple or systematic) random way or in a stratified random way. A random sample means that all car drivers in a country have an equal chance to find themselves in the sample. A stratified sample means that the car driver population is first divided into mutually exclusive and homogeneous subgroups or strata (e.g. based on gender and/or age); subsequently within each subgroup or stratum a random sample is drawn. Stratified sampling is recommended when it is possible to define subgroups that can be expected to be more homogeneous in respect to the variables of interest and as such differ from other subgroups or strata. This would increase the precision of the estimates of a variable for the total population.

For example, for RED and SPIs it is known that there are structural differences between men and women, between different age groups, and between drivers of a diesel car or petrol car. Therefore, stratification is recommended for at least:

- Gender (man / woman)
- Age (e.g. 17/18-24; 25-54; 55-69; 70+)
- Drivers of petrol vs. diesel cars

In other words, it has to be ensured that, within the sample, the number of men/women, the age division and type of fuel is representative for the total population of car drivers within a country. This also offers the opportunity of information about these specific subgroups, provided the numbers per group are sufficiently large. In general, when information about these or other subgroups is required, it is recommended to use these subgroups as strata in a stratified random sampling design since this yields more precise estimates than when the subgroups are composed afterwards and cut through other strata.

An additional variable for stratification that may need to be considered is:

- Region of residence of the driver

In some countries there are structural differences in road behaviour and mobility patterns in different parts of the countries, for example because of differences in environment (flat vs. mountainous or hilly; dense or less dense motorway network) or in attitudes.

With respect to the car, it may be useful to stratify based on:

- Year of construction of the car
- Lease or private car

4.2. Non-response

A statistical problem with ND research is that participation is voluntary. That means that complete random assignment to the sample is per definition impossible. It cannot be excluded that those who want to participate differ in some relevant ways from those who do not want to participate.

Whenever possible, this likely selection bias as a result of non-response should be corrected by poststratification based on:

1) Demographic information of the car driver population;
2) Technical characteristics of the passenger car population; and/or
3) Odometer readings of passenger cars as registered during (periodic) motor vehicle inspection as an indication of the distance travelled by a vehicle.

When specific cars or car types cannot participate in the study because technical restrictions prevent the installation of the chosen data acquisition system, then these cars should be treated the same as non-response.

4.3. Sample size

A crucial step in the definition of the study design is the sample size: how many drivers/passenger cars need to be in the sample to get a reliable estimate of the actual situation of the RED or SPI at hand? The optimal sample size depends on three factors:

- The amount of homogeneity/dispersion of the RED/SPI in the population;
- The required precision of the estimate;
- The required probability of obtaining this required precision.

The following three examples give an indication of the required sample sizes, given that a simple random sampling strategy is applied. The three examples show that the required sample sizes vary quite substantially. Therefore, if different RED/SPIs need to be monitored, this requires sample size estimations for each of these variables. Subsequently, the largest estimated sample size should be used in order to guarantee the required precision for all RED and SPI’s. Since random samples will have the largest standard errors, sample sizes need to be relatively large. Other sampling techniques like stratified random sampling usually require smaller sample sizes for the same amount of precision, since the standard errors are generally smaller. However, these also require more information about relevant population characteristics, i.e. estimated of the standard errors in each of the subgroups.

Example 1: RED vehicle kilometres

The required sample size when aiming for an estimate of the annual mileage of passenger cars in a country and using a simple random sample is shown in *Table 5.1*. First an assumption has to be made on the average mileage per car driver. In this example, an average mileage of 15,000 kilometres per year was assumed. *Table 5.1* shows the required sample size when the standard deviation (SD), i.e. the amount of variation or dispersion from the average, in the population is 5,000, 10,000 and 15,000 kilometres. Furthermore, it shows the required sample size for three levels of precision: ±10%, ±5%, and ±1%. The applied confidence level is 95%, i.e.

the probability that a specific random sample does not provide the required precision is 1 in 20.

In general, the lower the standard deviation/the less variation in the population, and the less precision is required, the smaller the sample size can be. The implication of the chosen precision level is that only differences between two time periods or between countries larger than twice this precision level will be detected with the corresponding sample size. From *Table 5.1* it can be concluded that, for example:

- With a population standard deviation (SD) of 10,000 and a sample of around 700 cars, differences in the actual annual mileage up to 10% (plus or minus 5%) will remain undetected.
- With a population standard deviation (SD) of 10,000, and a sample of around 17,000 cars only differences up to 2% will remain undetected.
- If, however, the population standard deviation (SD) is 15,000, a sample of over 38,000 would be needed to reach this level of precision of ±1%.

Table 5.1. *Sample sizes required for the estimation of total number of vehicle kilometres driven by cars in a country using a simple random sample with precision levels of ±10%, ±5%, and ±1%, population standard deviations of SD = 5,000, SD = 10,000, and SD = 15,000, and a confidence level of 95% (based on formula 2.65 in Commandeur, 2012).*

SD = 5,000			SD = 10,000			SD = 15,000		
±10%	±5%	±1%	±10%	±5%	±1%	±10%	±5%	±1%
43	171	4,269	171	683	17,074	385	1,537	38,416

Example 2: SPI, average speed on a 50 km/h road section

As a second example, we look at the SPI speed. Suppose we want to know the average speed of passenger cars on a road section using a simple random sample. Given that the average speed in the population on that road is 50 km/h with a standard deviation of 7 km/h, the minimum sample sizes to obtain an estimate of the average speed with a precision level of ±10% (i.e. estimations will be in the range of 45 to 55 km/h), ±5% (i.e. estimations will be in the range of 47.5 to 52.5 km/h), and ±1% (i.e. estimations will be in the range of 49.5 to 50.5 km/h) will need to be around 8, 30 or 750 respectively (*Table 5.2*).

Table 5.2. *Sample sizes required for the estimation of average speed on a 50 km/h road section using a simple random sample with precision levels of ±10%, ±5%, and ±1%, population standard deviation of SD = 7km/h, and a confidence level of 95% (based on formula 2.57 in Commandeur, 2012).*

SD = 7		
±10%	±5%	±1%
8	31	753

Example 3: SPI, seat belt use

As a last example, we suppose we want to estimate the percentage of seat belt use in the population of passenger cars of a country using a simple random sample (*Table 5.3*).

First, based on available previous data we assume that this percentage will be 90%. Then the minimum sample sizes to obtain a precision of ±10% (estimations will be between 80 and 100%), ±5% (estimations will be between 85 and 95%), and ±1% estimations will be between 89.9 and 90.1%) are around, respectively, 35, 140 and 3,500.

Second, we cannot make an assumption about the actual percentage of seat belt use in the population. In that case, it is safest to assume a percentage of 50%. As shown in Table 5.3, in that case the required sample size needs to be substantially larger to produce the chosen precision levels of respectively ±10%, ±5%, and ±1%.

Table 5.3. *Sample sizes required for the estimation of the percentage seat belt use by car drivers in a country assuming a 90% wearing rate or wearing rate unknown (assumed 50%) using a simple random sample with precision levels of ±10%, ±5%, and ±1% and a confidence level of 95% (based on formula 2.69 in Commandeur, 2012).*

Assuming seat belt use 90%			Assumed seat belt use 50%		
±10%	±5%	±1%	±10%	±5%	±1%
35	139	3,458	97	385	9,604

The indications of the required sample sizes in the three above examples assumed simple random sampling. As indicated, a stratified sample with subgroups with known more homogeneous characteristics may reduce the required sample size considerably. Depending again on the required precision of the estimate, the estimated average value in the population and the (smaller) standard deviation in each of the subgroups, required sample sizes may be up to 70% smaller (see example 3.8 of Commandeur (2012, page 45-46) for the detailed assumptions and formulas.

Furthermore, the continuous nature of the measurements obtained in a naturalistic driving study implies that the ratio and/or regression estimators discussed by Commandeur (2012) are natural and well-suited candidates for statistically improving the precision of the estimates. This is because the sample observations obtained for a previous time point or time period can be used for the estimates in the next time point or time period. However, both ratio and regression estimators do require knowledge or an estimate of the value in the population; hence, they can only be applied for some SPIs.

4.4. Data collection period and sample rotation

The ultimate aim of ND for monitoring is to get a good estimate of the annual average of a particular variable and compare this between countries and over the years. For getting a reliable annual average, given the various seasonal fluctuations, for example due to holiday periods and weather conditions, it is considered best to observe the RED and SPI’s throughout the year. An alternative would be to collect data for more limited time at fixed periods throughout the year, e.g. one week every month or two weeks every three months, and extend the results to estimate an annual value. However, the result will be less reliable and, moreover, this approach requires substantially more effort in installing and uninstalling the equipment.

Assuming a one year continuous data collection effort for several consecutive years, the best sampling strategy for measuring changes over time is to use a rotating sample. A rotating sample means that after some fixed period of time one part of the sample is replaced, while the other part remains in the sample. A possible rotation scheme is to replace half of the original sample after half a year and observe these replacements for one full year. The other half of the sample is observed the whole first year, and then replaced with a new sample, et cetera. This way none of the sampled cars are in the sample for more than one year, while still being rotated on a fifty percent basis. It is also possible to apply a smaller rotation scheme, e.g. replacing one third, a quarter or even one fifth.

In general, when the main aim is to identify changes over time, it would be ideal to try to keep the original sample constant over the years. If it is also the aim to get an indication of the current average value in a country (as compared to another), sample rotation is required. In that case, Cochrane (1977, cited by Commandeur, 2012) considers retention of half, three quarter or four fifth of the sample from one period to the next as a good practical policy. Moreover, in relation to Naturalistic Driving, it must be noted that the larger the percentage of drivers/cars replaced and the shorter the rotation period, the more effort in recruiting participants (See Chapter 5), the more effort in replacing equipment, and the more chance for technical breakdowns due to these replacements.

4.5. Recommendations

In summary, regarding study design it is recommended to:

- Use a stratified sample of the passenger car drivers in a country based on gender, age, petrol vs. diesel cars, and region of residence.
- Define the sample size based on:
    - the simple random sampling formulas presented by Commandeur (2012);
    - aiming for a precision level of ±1%;
    - Subtracting around 50% of the resulting required sample size when using a clever, well-founded way of stratification.
- Collect data throughout the year on a continuous basis.
- Follow each individual in the sample for one year and apply a rotating scheme of 50% per 6 months.
- A sample of 10,000 drivers per country seems to be the minimum for RED such as the annual amount of vehicle kilometres. This number is usually independent of the size of the population of car drivers in a country. Only if the sample size is larger than 10% of the population, a correction is applicable.

5. ORGANISATION AND TECHNIQUES

Chapter 5 discusses the general principles and current state of the art of various organisational issues in relation to participant recruitment, data acquisition, data transfer and storage, data analysis, as well as the legal and ethical issues that have to be taken into account. We do not consider it necessary to re-cover each and every facet of a naturalistic driving trial in great detail here or to go to the level of functional specifications. There is a lot of knowledge already available through other projects so this section will begin from where one of the most recent projects, PROLOGUE, left off.

5.1. Participant recruitment

Participation in ND research is per definition on a voluntary basis and experiences in the USA and Europe have shown that it may take time to find sufficient and suitable participants. Like for other studies, the process begins with setting out driver criteria; this is very dependent on study design but will be aimed at getting a reliable sample of the driver population in question. This selection has historically been based on gender, age, driving experience and annual mileage although other criteria can be added to further define the driver population (see also Chapter 4 about stratification).

Previous experience has shown that compromises normally have to be made in respect to the criteria rather than the selection i.e. some softening of the criteria is necessary to get the required numbers rather than having an unlimited choice of drivers and selecting the ‘best’.

Another bias will be generated if the vehicle type is one of the selecting criteria. For example if certain make and model of vehicle are selected (for instrumentation simplicity) then the recruitment is even more restricted and even more bias of driver types will be expected to recruit the required numbers.

In summary, if recruiting by driver criteria, then you will normally get a random selection of vehicle types with a fairly tight control over biases in the driver sample. If however the recruitment is predominantly based on vehicle type then the effect may well be a more serious bias of the driver types and as such much more relaxation of the criteria will be needed.

Generally there has been little use of driver style measure or sensation seeking in the recruitment phase although it has been done in specific trials. This information is normally revealed through background questionnaires.

Recruitment can and has been done through contacting fleet drivers, motoring organisations, clubs and companies. This process does generally lead to a higher response rate as there is an element of compulsion in some sectors however this very fact introduces bias into samples as the types of driver who may use driving organisations/clubs are perhaps not reflective of the general driver population; the same can be argued for the other groups.

Some form of reimbursement or incentive is incredibly important in the recruitment and, perhaps more importantly, the retention of drivers. There are no set guidelines specifically for naturalistic studies but a number of initiatives have been used:

Fuel cards: payment can be provided in the form of fuel cards; this system ensures that payment is distributed evenly and that continuation in the study will result in continued payments. This technique also circumvents some taxation issues with payments to participants.

Incremental payments: Issuing payments only on set dates distributed throughout the study or issuing payments after certain milestones (e.g. questionnaire completion) can ensure continued participation; drop outs can also be paid only up to the point they left the study.

Competitions: In some studies, normally of short durations, a prize fund is offered of greater value than an individual participants' payment would be; this technique can work if everyone wants the prize but can distance some participants resulting in early drop out – the prize is normally only available to participants completing the whole duration.

Free access: In some instances where a test device is used (e.g. Navigation device or other technology) then free access to this device and the device as a 'gift' can be enough to secure participant involvement. Caution needs to be taken if the test device is not significantly desirable as this can lead to early drop out.

The recruitment and management of participants has been considered a full time job in many studies. The demand posed by phone calls and e-mails from participants or collecting and storing data negates all other roles for the duration of the trial.

Having someone who is responsible for all participants in a country can also be a good way of retaining participation. A friendly contact that is familiar with each participant's requirements is much preferable to an ad-hoc approach with a number of staff.

A help desk is a useful approach as this appears to give some comfort to the participants – most studies offer a 24/7 service although the actual contact is very dependent on the method used with phone calls/texts being staffed more extensively whereas e-mail is predominantly restricted to office hours. Experiences from the UK suggest that contact with the participant initiated from the centre is much more valuable in detecting issues. This contact may be face to face or by phone, however, the effect is the same; participants seem much more open to respond to questions than to initiate contact themselves.

5.2. Data acquisition

As mentioned in Chapter 3, for costs reasons data collection has to be realised by a rather low-tech, simple device. One solution seems to be a Smart Phone-type of device with a GPS function. The advantage of using a smart phone based logger is that the device is more likely to be fitted (or at least carried) in the participant's vehicle[8].

Conversely, any device which requires the participant to turn it on (or to start the logging) will suffer with respect to data quality and completeness. UK TeleFOT data where an auxiliary, automated logger was used alongside a manually activated device shows a significant loss in data between the continuous logging and the user

[8] as a side effect of this it can be possible to deduce the travel mode from a continuous GPS log whether carried in a car, while walking or using public transport (this technology as already available through smart phone apps'). All of these modes exhibit distinctive GPS 'signatures' and as such a much more detailed overview of all travel modes is possible.

operated. This disadvantage carries over to powering the device where battery life can abruptly end logging if not topped up by a vehicle charger.

In addition, a reliable indication of acceleration and deceleration force requires that the device has a fixed position and is oriented correctly, so a deceleration records as a deceleration and not as a cornering force. It is necessary to understand that the accelerometer in these devices will be built down to a cost and as such will not compare directly with more expensive scientific equipment.

5.3. Data transfer and storage

There are a number of proven methods for the transfer of data from the participant's vehicle to a central or local data server. There is no 'correct' method with the decision based predominantly on ease of use and the type of data.

Early studies in the USA and more recently in Europe have made good use of removable media; essentially a hard disk in the vehicle, to record data. These disks are then posted or picked up manually from participants and uploaded to a local/central server. The disks are relatively stable (for one use at least) and can record large quantities of data. There are, however, a number of limitations to this technique including the need for regular contact with the participant and a smaller test region to cover. These limitations especially apply when having a large sample, covering all regions of a country, as will be the case when using ND for monitoring purposes.

When using simple data collection methods such as smart phones or GPS loggers, and particularly if video is not present, then the data can also be transferred wirelessly. Depending on the sample rate and quantity of data this can be nearly continuous via general packet radio service (GPRS) or in some form of data dump.

For somewhat less simple data, an intermediate step that can prove effective if the fleet of vehicles is relatively well controlled in location, is the use of a beacon: data (both digital and video) will be sent wirelessly once the vehicle is returned to a location in close proximity to a beacon which can transfer data much faster than a commercial GPRS network for example.

5.4. Data processing and analysis procedures

In the same way as the data transfer and storage, there is no prescriptive way of analysing data from vehicles. The techniques used will depend very much on the format in which it is presented, any privacy or ethical concerns and on how the data is processed.

In principle, as indicated by Bonnard et al. (2012), there are two solutions for processing the data and come to national SPI and RED results within the framework of ERSO:

1. Each country is in charge of the calculation of their respective RED and SPI indicators, only the indicators necessary for the ERSO are shared.

2. Setting-up of a joint database for all members of the ERSO and calculating the national SPIs and RED at a central level.

For practical reasons, the first option, where data is processed at country level, is to be recommended. This option provides a better opportunity to protect the privacy of participants and to take account of local characteristics, e.g. related to the sample or regional differences.

When storing and processing data at national level, it must be ensured that this is realised in as much the same way as possible, so that the results are comparable. However, this will always increase the chance of inconsistencies between countries, since it is very difficult, may be even impossible, to guarantee that all the data is filtered, processed or calculated in the same way. Hence, it could be considered to move to central analysis for so called ‘core’ data. This data set can be seen as a minimum data specification which is needed from every country.

If (event-triggered) video or specialist sensor data is involved then this will be almost certainly analysed by the respective partner who collected it, as has been done in recent European projects. In that case, close cooperation will be required to ensure data comparability when the data is merged. To circumvent this issue, it may be considered to nominate an analysis centre to cover this piece of work. This is a technique to be used in the current UDRIVE project to ensure that all the data is analysed uniformly. Care has to be taken, however, to avoid overwhelming an analysis centre if not managed correctly.

5.5. Legal and ethical issues

There is a considerable amount of legal and ethical issues involved in ND research that need to be taken into account. Bonnard et al. (2012) summarized the main issues and ways to solve them based on the FESTA Handbook on Field Operational tests (FESTA, 2011) as well as recent European ND initiatives (e.g. PROLOGUE, INTERACTION, UDRIVE).

In summary, the two most important *European Directives* are:

- Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data; and
- Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector.

Just to mention a few of the important issues in these two Directives for participants in an ND study:

- Personal data may be processed only if the participant has unambiguously given his/her consent;
- The participant should have the right to object, on legitimate grounds, to the processing of data relating to him/her;
- It must be ensured that personal data is accessed by authorised persons only;
- It must be ensured that a security policy on the processing of personal data is implemented.

National laws have often extended the requirements of these Directives with their own laws, describing national restrictions and obligations for collecting and storing personal data. Some European countries also require approval of an ethical committee, either at national level or at university or institute level. Hence, not only the European requirements but also the national data protection laws and related legal requirements and (often time-consuming) procedures should be examined closely.

Some additional legal issues that have to be dealt with include:

- Ensure that the participants hold a valid driving permit for the duration of their participation.

- Ensure that the participating car is sufficiently insured and that participation does not invalidate the insurance.
- Ensure that the vehicle type approval is not invalidated due to the instrumentation of the vehicle.

In addition, in case of external video views, provisions must be made to prevent recordings in areas where cameras are not allowed, e.g. military bases, some international border crossings, or similar facilities.

As indicated, the participants must give their unambiguous consent that their data is processed. This so-called *informed consent* that each participant has to sign must also specify and explain issues like (Bonnard et al., 2012):

- Costs; who is responsible for certain costs (e.g. vehicle maintenance, damage to vehicle, insurance excess, traffic penalties)?
- Benefits; what is the allowance the participant will receive and are there possible other benefits (e.g. use of instrumented vehicle, fuel cost reimbursements)?
- Risks; is the participant exposed to increased risks (of involvement in crashes or of theft or burglary of the vehicle or ND-devices) by participating and if so, what has been done to minimise the risks?
- Withdrawal; is the participant free to withdraw his/her participation at any moment and how will this affect the agreed participant allowance?
- Confidentiality of recorded data; how is the participant's privacy protected? What will and what will not be done with the data gathered? Which parties will have access to the recorded data (e.g. police, the judiciary)? Who owns the data (during and after data collection)?
- Who is allowed to drive the vehicle, how will be dealt with data recorded of non-participating drivers (in case non-participating drivers are allowed to drive)?

As an example, the SHRP2 (USA) informed consent for participating drivers and for non-participants that may drive the equipped car, can be consulted at:

http://projects.battelle.org/drivingstudy/IRB_0434_Rv%208.1_SHRP2_PrimaryDriver.pdf

http://www.shrp2nds.us/PDFs/Indiana_Forms/IN_Secondary_Driver_Consent_1yr.pdf)

With respect to *data protection*, data storage and data transfer need to be properly secured:

- Data that is temporarily stored in the vehicle should be secured (or encrypted) to avoid unauthorized access, for example in case of burglary.
- Data transfer from the vehicle to the (intermediate) database should be secured. In case of wireless data transfer, encryption methods are needed to prevent unauthorized access by the network operators.
- In case of intermediate data storage, this data should be deleted after transfer to the final data storage, in such a way that data recovery is impossible.
- Also the final data storage should be secured and data access should be regulated. All 'users' of the data should sign a confidentiality agreement.
- Data that could lead to identification of the participant should never be released to other parties then described in the informed consent without prior agreement of the participant.

Furthermore, in case of an external video view, identification of other road users by unauthorized parties should be avoided by, for example, blurring faces or licence plates.

5.6. Recommendations

In principle, the techniques and procedures for ND data collection, data transfer, data storage and data analysis are available and not too complicated, given that the data collection effort that is required for monitoring SPIs and RED excludes continuous video registration, and as such, is not that extensive. Wireless data transfer via general packet radio service (GPRS) or in some form of data dump or via a beacon can be applied.

It is recommended to do the analyses at a national level, applying a series of definitions on variables and disaggregation levels and following fixed analysis protocols, to increase comparability. It could be considered to identify a limited number of ‘core’ variables (SPIs/RED) to be analysed at a central/ERSO level to exact comparability.

A word of caution is justified when considering participant recruitment and legal/ethical issues. Participation in ND research is per definition on a voluntary basis and experiences in the USA and Europe have shown that it is a time-consuming effort to find sufficient suitable participants, especially if there are strict sample stratification requirements (e.g. related to age, gender, region, type of vehicle), as identified in the previous Chapter. In addition, there is a considerable amount of legal and ethical issues involved in ND research that need to be taken into account, in particular in the area of privacy and data protection. It is recommended to develop a clear protocol to fulfil the EU regulations, based on the FESTA handbook, and to develop a checklist to ensure that national regulations are met. The latter may be facilitated by involving a national lawyer.

6. TOWARDS A SCENARIO 4?

So far, three scenarios were presented. These have in common that SPI and RED data is collected through equipment and sensors added to the vehicle. This has proven to be a feasible approach that can provide useful information. However, given that fairly large samples are needed, it is also a rather costly and labour intensive approach. In addition, the reliability of the data is dependent on the recruitment of a representative sample of the population, an effort that has proven to be not that easy at all.

In theory, there is a fourth scenario; a scenario that is not dependent on equipping cars nor on voluntary participants, but a scenario that extracts data directly from all cars based on CAN-data, OBD, and other trip and travel data collected automatically by the vehicle (e.g. trip recorder, event recorder, E-call-related data). This approach would result in more reliable data because it would include the complete passenger car population as well as other motorised vehicles and would not be dependent on participation of volunteer car drivers. In addition, this scenario would make data collection considerably cheaper.

This option, however, is not something that can be realised overnight. One important aspect is that, currently, car manufacturers apply their own technical specifications for most of the CAN and OBD data and they are not very keen on sharing these with other car manufacturers or other external parties. This means that this type of data is not widely accessible nor comparable between car makes and models.

Given the theoretically promising characteristics of this approach, it is time now to explore the feasibility and future options as well as the roles of the various parties involved.

As a first step, the requirements for this data need to be elaborated:

- Which data is needed as an absolute minimum;
- Which additional data is desirable;
- What is the minimum level of disaggregation;
- What is the desirable level of disaggregation.

This is a process that needs to take place in consultation with the car manufacturers themselves, in order to ensure that the requirements are specified in such a way that they are technically feasible. Moreover, timely involvement may help to realise their commitment and a positive attitude.

Furthermore, the European Commission can play an important role as well by promoting or maybe even regulating harmonisation of, and free access to the relevant data of the different European car makes and models. An important issue is that the access to and the use of the data do not conflict with European or national privacy legislation.

Since, eventually, also non-European car makes and models would need to be included, this effort would also affect car manufacturers outside Europe, e.g. because it might result in specific requirements for non-European cars that are imported to the EU.

An important other aspect related to this approach would be the public support for transferring all sorts of privacy-sensitive data from their car to a central database. Even though subsequent data aggregation and data storage can be arranged so that information cannot be traced back to individual vehicles/cars, it is not unlikely that a majority will develop strong anti-‘Big Brother’ sentiments. If this means that people have to give their informed consent for logging the data of their car, there is again the issue of a representative sample.

As indicated, it will take a long time to make this approach work. But in the end, an approach that directly extracts the relevant information from the vehicle, seems to be a more solid and sustainable approach than monitoring through ND research. Therefore, it is recommended to start discussions now, trying as a first step to break the taboo of sharing some information between different car manufacturers.

7. CONCLUSIONS & RECOMMENDATIONS

This Deliverable presented some considerations and recommendations for applying the Naturalistic Driving (ND) approach to collect information about safety performance indicators (SPIs) and risk exposure data (RED) with the aim to compare differences between EU Member States and developments over time.

The main conclusions is that, in principle, the ND approach has substantial added value compared to more traditional data collection methods like crash registration and surveys, because ND ensures continuous, automatic and standardized data collection. This is true for both SPIs and RED. A prerequisite is that similar data acquisition systems and methods/definitions are applied. These systems as well as technology for data transfer and data storage is available and has proven to be operational. Though the current Deliverable is purely focused on road safety and exposure data, the collected data will also be useful for other transport areas, in particular eco-driving and traffic management.

In order to get reliable information, a fairly large sample is needed. The exact size of the sample depends on the variation in behaviour in the population and the required level of precision of the results. Assuming that the sample is drawn in a cleverly stratified way, resulting in a number of mutually exclusive and homogeneous subgroups (e.g. based on gender and age), a sample of 10,000 drivers per country seems to be the absolute minimum for RED such as the annual amount of vehicle kilometres. Experiences in the USA show that it may require substantial effort to get sufficient participants with the required characteristics to allow for good stratification.

With regard to data collection, based on cost considerations, three scenarios are distinguished. It is recommended to start off with Scenario 1: a low-cost simple, off-the-shelf simple data acquisition system (e.g. an OBD GPS tracker or a Smart Phone) and a limited number of additional sensors, measuring:

- Vehicle kilometres
- Person kilometres
- Number of trips
- Time in traffic
- Speed
- Seat belt use
- Light use

In addition, the data acquisition system would need to register continuously the time, the date, and the location (GPS). In combination with a map matching tool, and an indication of road class and the speed limit, this would allow for comparisons of the mentioned RED and SPI for different road classes and would give an indication of the occurrence of excessive speed. For cross-national comparisons it is important to define a (limited) number of comparable road classes. Furthermore, as a relatively simple driver identification method, it is recommended to use a magnetic swipe card or an RFID tag.

In a later stage, additional SPIs and network characteristics could be added successively (Scenario 2), including:

- Time headway
- Acceleration
- Lane departures

- Inappropriate speed
- Signal use
- Junction type

A few SPIs are very relevant from a safety point of view, but with current techniques cannot be measured reliably in an unobtrusive way. This applies, in particular, to alcohol and drugs use.

In addition, SPIs that would need continuous external and/or internal video recordings do not seem to be feasible in the short term, because this results in huge amounts of data and extreme high costs for the related data transfer and data coding. That means that the SPIs fatigue, inattention, distraction and the (proper) use of child restraints cannot be monitored by means of ND research.

Furthermore, as Scenario 3, it is recommended to equip a limited number of cars also with an event-triggered video in order to monitor numbers of near crashes as yet another relevant SPI. As a very useful side product, this effort will provide data that can be used to further specify and refine the quantitative and qualitative relationship between near crashes and real crashes.

For all three scenarios very strict European and national legislation applies in relation to data protection and privacy, among others requiring all participants to sign an informed consent. Also non-participants who sometimes drive the car as well as participants need to be informed.

Though it is impossible to give a reliable estimate of the costs involved, the costs can be expected to be fairly high. Just assuming a simple OBD GPS tracker of €100 and a participant incentive at the value of €400 would add up to an annual 5 million euro per country assuming the recommended sample size of 10,000 drivers. And this amount does not include the costs of man power related to participant recruitment and contact, and the organisation and management of the data collection, transfer, storage and analysis.

In short:

- ND research can provide very useful information about several very relevant SPIs and RED for cross-national comparisons and comparisons over time.
- Technology for data collection, data transfer and data collection is available and has proven to be operational, at least on small and medium scale.
- Bottlenecks in the successful implementation of ND research for monitoring may be:
    - Recruitment of sufficient participants
    - Harmonization of definitions of variables, disaggregation levels and analyses
    - Operation costs

Hence, in parallel , it is recommended to start exploring the possibility of a scenario 4 now, i.e. a scenario where relevant data is extracted directly from the vehicle via CAN-bus, OBD, and other trip and travel data collected automatically by the vehicle. In theory, that way a lot of relevant information is already available with no or little additional costs; in practice, however, the information is not generally accessible nor comparable between car makes and models. This will take a long time to realise, but first steps can be made now. One of the first steps, in consultation with the car manufacturers, is an elaboration of the requirements for this data: what is available, what is needed, what is technically feasible. The European Commission can play an important role as well by promoting or maybe even regulating harmonisation of, and free access to the relevant data of the different car makes and models.

Whatever data is collected, whatever data acquisition system is applied, the ND approach for monitoring, as discussed in this report, is largely oriented towards passenger cars and their drivers; as a consequence, the resulting information about SPIs and RED is restricted to that user group.

The ND methodology can also be applied to other motorised vehicles, such as trucks and motor cycles, but that will involve several additional organisational and technical requirements and related efforts. Developments are going on in relation to enabling ‘Naturalistic Cycling’ which would allow collection of SPIs and RED for cyclists, but current technology is not sufficiently robust and stable to apply on a large scale. A naturalistic approach for collecting safety and exposure data for pedestrians is not yet available. This all means that getting an overall view of the safety related behaviour and the exposure to risk of *all* road users, would require more additional methods including the more traditional surveys, trip diaries, and observations.

8. A CENTRAL ROLE FOR EUROPE

Despite various bottlenecks and challenges, the potential of ND research for monitoring purposes is sufficiently large to start off, as suggested before, with the implementation of Scenario 1. Since harmonisation and international comparability of data are the key reason for this effort, there is a central role for the European Commission in initiating this task and taking the lead from here, most likely within the ERSO framework. A stepwise approach is recommended, including successively:

1. Creating support and finding budget by presenting the case to the relevant road safety bodies at European and Member State level, explaining the need for harmonised, comparable international data, the ND approach, and its added value.
2. Preparing a detailed description of / handbook for all practical implementation aspects, including the functional specifications of data collection equipment, participant selection, data transfer and storage, as well as definitions of variables, disaggregation levels and analyses.
3. Identifying the relevant national organisations which will be responsible for national data collection and pre-analyses, and fine-tuning data collection procedures (including legal aspects) and variable definitions in consultation with them.
4. Developing and equipping a database at EU level and defining the required (pre-analysed aggregated) data to be provided by the Member States as well as the procedures and time schedule, in consultation with the relevant national organisations.
5. Setting up European-wide communication strategies to guarantee maximum dissemination and use of the collected data.
6. Setting up one year national pilots in at least four Member States, well spread of Europe (North, West, South, East).
7. Adapting procedures and definitions, based on the pilot experiences.
8. Successive implementation of Scenario 1 in additional Member States.

Parallel to steps 6 and 7, Scenario 2 (additional SPIs/RED) and 3 (monitoring near-crashes) can be elaborated, piloted and implemented, applying a similar stepwise process.

From the very beginning, the EC is advised to initiate discussions with the car manufacturers, using existing discussion platforms, with the aim to explore longer term possibilities of Scenario 4, i.e. the scenario where relevant data is extracted directly from the vehicle.

Finally, in order to elaborate these steps and to assist the EC in performing these steps, it is advised to compose a consortium of organisations. Possibly, this can be part of the future research agenda that is currently being prepared by the PROS consortium.

REFERENCES

Aarts, L. & Van Schagen, I. (2006). Driving speed and the risk of road crashes; A review. Accident Analysis and Prevention, 38 (2), 215-224.

Bonnard, A., Brusque, C., Hugot, M., Commandeur, J., and Christoph, M. (2012). Study design of Naturalistic Driving observations within ERSO – Development of innovative indicators for exposure and safety performance measures, Deliverable 6.2A of the EC FP7 project DaCoTA.

Commandeur, J.J.F. (2012). Sampling techniques and naturalistic driving study. Deliverable 6.2B of the EC FP7 project DaCoTA.

Dingus, T. A., Klauer, S.G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., and Knipling, R.R. (2002). Phase II – Results of the 100-Car Field Experiment. NHTSA report DOT HS 809 593. Washington, US.

Elvik, R. (2009). The Power Model of the relationship between speed and road safety: update and new analyses. TØI Report 1034/2009. Institute of Transport Economics TØI, Oslo.

ETSC (2001). Transport safety performance indicators. Brussels, European Transport Safety Council ETSC.

FESTA Consortium (2011). FESTA Handbook Version 4. D6.4 of the Field opErational teSt supporT Action (FESTA).

Hakkert, A.S and Gitelman, V. (Eds.) (2007). Road Safety Performance Indicators: Manual. Deliverable D3.8 of the EU FP6 project SafetyNet.

Hakkert, A.S., Gitelman, V., and Vis, M.A. (Eds.) (2007) Road Safety Performance Indicators: Theory. Deliverable D3.6 of the EU FP6 project SafetyNet.

Hanowski, R. J., Hickman, J., Olson, R. L., Bocanegra, J. (2009). Evaluating the 2003 revised hours of service regulations for truck drivers: The impact of time-on-task on critical incident risk. Accident Analysis and Prevention, 41, 268-275.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., Ramsey, D. J. (2006). The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study.Report No: DOT HS 810 594. National Highway Traffic Safety Administration (NHTSA), Washington D.C.

Koornstra, M.J. Lynam, D. Nilsson, G. Noordzij, P.C. Petterson, H.-E. Wegman, F.C.M. & Wouters, P.I.J. (2002.) SUNflower: a comparative study of the development of road safety in Sweden, the United Kingdom, and the Netherlands. Leidschendam, SWOV; Crowthorne, TRL / Linköping, VTI.

LTSA (2000). Road safety strategy 2010. A consultation document. National Road Safety Committee. Land Transport Safety Authority, Wellington, New Zealand

Pilgerstorfer, M., Runda, K., Brandstätter, C., Christoph, M., Hakkert, S., Ishaq, R., Toledo, T., Gatscha, M. (2011). Small Scale Naturalistic Driving Pilot, Deliverable 6.3 of the EC FP7 project DaCoTA.

Talbot, R., Meesmann, U., Boets, S. and Welsh, R (2010). Naturalistic Driving Observations within ERSO, Deliverable 6.1 of the EC FP7 project DaCoTA.

Van Schagen, I., Welsh, R., Backer-Grøndahl, A., Hoedemaeker, M., Lotan, T., Morris, A., Sagberg, F., Winkelbauer, M. (2011). Towards a large-scale European Naturalistic Driving study: main findings of PROLOGUE. PROLOGUE Deliverable D4.2. SWOV Institute for Road Safety Research, Leidschendam, The Netherlands.