UDRIVE: the European naturalistic driving study

Rob Eenink\textsuperscript{a}, Yvonne Barnard\textsuperscript{b,*}, Martin Baumann\textsuperscript{c}, Xavier Augros\textsuperscript{d}, Fabian Utesch\textsuperscript{c}

\textsuperscript{a} SWOV Institute for Road Safety Research, Leidschendam, The Netherlands
\textsuperscript{b} ERTICO - ITS Europe, Brussels, Belgium
\textsuperscript{c} German Aerospace Center (DLR), Braunschweig, Germany
\textsuperscript{d} Volvo Group Trucks Technology, Saint Priest, France

Abstract

UDRIVE is the first large-scale European Naturalistic Driving Study on cars, trucks and powered two wheelers. The acronym stands for “European naturalistic Driving and Riding for Infrastructure & Vehicle safety and Environment”. Naturalistic driving 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. Data collection will take place in seven EU Member States. Road User Behaviour will be studied with a focus on both safety and environment. The UDRIVE project follows the steps of the FESTA-V methodology, which was originally designed for Field Operational Tests. Defining research questions forms the basis of the study design and the specification of the recording equipment. Both will be described. Although the project has yet to start collecting data from drivers, we consider the process of designing the study as a major result which may help other initiatives to set up similar studies.

Keywords: naturalistic driving study; road user behaviour; road safety; eco-driving

Résumé

UDrive est la première étude de conduite naturaliste européenne à grande échelle portant sur les voitures, les camions et les deux-roues motorisés. UDrive est l'acronyme de « European naturalistic Driving and Riding for Infrastructure & Vehicle safety and Environment » qui signifie « conduite naturaliste européenne pour l'environnement et la sécurité des infrastructures et des véhicules ». La conduite naturaliste peut être définie comme une étude visant à fournir un aperçu du comportement chauffeur au cours de ses trajets quotidiens à partir d’une collecte de données provenant du véhicule, de son environnement et du conducteur lui-même, collecte effectuée grâce à un équipement de collecte discret et sans contrôle expérimental. La collecte des données aura lieu dans sept états membres de l’UE (Allemagne, Autriche, Espagne, France, Pays-Bas, Pologne et Royaume-Uni). Le comportement des usagers de la route sera étudié avec un accent sur la sécurité et l'environnement. Le projet UDrive respecte les étapes de la méthodologie FESTA-V, qui a été initialement conçue pour les essais opérationnels. Définir les questions de recherche est fondamentale pour la conception et les spécifications du matériel d'enregistrement. Les deux seront décrits. Bien que la collecte de données en tant que telle n’ait pas encore commencé, nous considérons le processus de conception de l'étude comme le premier résultat important qui peut aider d'autres initiatives à mettre en place des études similaires.

Mots-clé: étude de conduite naturaliste, comportement des usagers de la route, sécurité routière, éco-conduite

* Tel: +32 (0)2 400 07 12, Fax: +32 (0)2 400 07 01.
E-mail address: y.barnard@mail.ertico.com.
Nomenclature

DAS | Data Acquisition System
ND | Naturalistic Driving
NDS | Naturalistic Driving Study
SCE | Safety Critical Event
UDRIVE | eUropean naturalistic Driving and Riding for Infrastructure and Vehicle safety and Environment

1. Introduction

1.1. Human behaviour in road transport

Road transport is indispensable for the exchange of goods and persons within the European Union and between neighbouring countries. Unfortunately, road transport also has several major negative consequences, in particular those related to crash-related fatalities and injuries, and to harmful emissions and the use of non-renewable energy.

Human behaviour is, directly or indirectly, an important determinant of these negative consequences. Consequently, an in-depth understanding of road user behaviour is needed to identify the main causes of, and the most promising approaches to mitigate the negative consequences. For example, how do road users behave in different conditions; how and when do normal traffic conditions or ordinary behaviour evolve into critical events or (near-)crashes; which factors affect driving style and related vehicle emissions, etc.?

Although there is a wealth of research on driver behaviour obtained in laboratory and simulator studies, there is a lack of understanding of what happens on the road in every day traffic situations. Naturalistic Driving (ND) studies are considered most appropriate as this method provides insight into the actual real-world behaviour of road users, unaffected by experimental conditions and related biases.

1.2. Naturalistic Driving

An ND study can be defined as a study undertaken to provide insight into driver behaviour during everyday trips by observing in detail 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 vehicles - passenger cars, trucks or motorcycles - are equipped with several small cameras and sensors. For several months to several years, these devices inconspicuously record vehicle manoeuvres (such as speed, acceleration/deceleration, direction), driver behaviour (such as eye, head and hand manoeuvres), and external conditions (such as road, traffic and weather characteristics). Thus, the ND approach allows us to observe and analyse the interrelationship between driver, vehicle, road and other traffic participants in ordinary situations, in conflict situations and, more rarely, in some actual crashes. This type of information is necessary for identifying new and promising measures, not only for reducing road transport casualties, but also for reducing the environmental impact of road transport.

Traditionally, road behaviour research has been using driving simulators, instrumented vehicles, self-reports, analyses of crash statistics, and increasingly also in-depth crash investigation. These methods have greatly contributed to the understanding of road user behaviour. However, they also have several limitations. For example, results from driving simulator studies cannot always be easily transferred to real traffic situations, since both the traffic environment and the vehicle characteristics are only approximations of reality. This is especially true in the simpler and static-based simulators. In instrumented vehicle studies subjects drive in real traffic but in a special, highly equipped vehicle with, usually, an experimenter on-board. This makes subjects aware of the fact that they participate in an experiment which may affect their driving behaviour. The results of self-reports may be biased by socially desirable responses as well as by perceptual and memory limitations.

Compared to these traditional road safety research methods ND has many important advantages. ND also offers much wider perspectives in understanding normal traffic behaviour in normal everyday traffic situations. Participants are not involved in an experiment; there is no observer present, there are no experimental interventions or aims that participants can guess and act for. Furthermore, there is the possibility to observe...
conflicts, near-crashes or possibly even actual crashes without potential biases of post-hoc reports. As such an ND study can contribute to clarifying the prevalence of, for example, fatigue and distraction among drivers/riders and the related crash risk; the interaction between road and traffic conditions and road user behaviour; to understanding the interaction between drivers/riders and vulnerable road users; to specifying the relationship between driving style and vehicle emissions and fuel consumption; and many other aspects of traffic participation that are difficult to study by means of traditional research.

1.3. Experiences with ND

ND is a fairly new research method that was developed in the late nineties of the previous century and that has been developed and refined continuously since then. The method has become technically possible because of the tremendous developments in information and communication technologies, improvements in storage capacities, data-mining, image processing, low-cost camera technology, etc. in the last couple of decades.

The first major ND study was conducted in the USA by Dingus et al. (2006) who instrumented the cars of one hundred drivers who commuted on a regular basis in the Northern Virginia/Washington D.C metropolitan area. They gathered data over a 12-month period and during this time the vehicles were driven over 3,000,000 km with a total of 43,000 hours of exposure involving 67 (mostly minor) crashes and 761 near-crashes. All vehicles were instrumented with a Data Acquisition System (DAS) engineered by Virginia Tech Transportation Institute (VTTI).

Continuing developments now enable the collection, transfer and analysis of up to even 1 petabyte of data, i.e. one million gigabytes, as is expected to be the case in a large-scale ND study that is currently being performed in the USA. This large-scale study is part of the second Strategic Highway Research Program (SHRP2: www.shrp2nds.us; Antin, 2011; Boyle, 2009). The aim of this follow-up project in the USA is to collect almost 4,000 vehicle years of data. The data acquisition is close to finishing, and it is expected that at the beginning of 2014, 3,800 vehicle years of data (around 4 petabytes), over 50 billion km’s and 400 crashes will be available for analyses.

In Europe also there have been various, smaller-scale and more focused ND studies that included ND methodologies. Several projects funded by the European Commission are:

- PROLOGUE: aimed to assess the feasibility and usefulness of a large-scale ND study in Europe and to formulate recommendations for such a large-scale study (www.prologue.eu; Sagberg et al., 2011);
- INTERACTION: aimed at a better understanding of driver interactions with in-vehicle technologies (interaction-fp7.eu);
- 2BeSafe: focusing on the behaviour and safety of powered two-wheelers (www.2besafe.eu; Laporte & Espié, 2012);
- DaCoTA assessing the usefulness of the ND method for gathering large-scale, representative information about safety performance indicators and exposure in the different EU Member States (www.dacota-project.eu; Thomas et al, 2013);
- SeMiFOT: aimed at implementing and developing the Naturalistic Field Operational Test (N-FOT) method as a method to understand crash causation and the effect of new safety systems (Victor et al, 2010).
- Large Field Operations Tests such as euroFOT (studying the use of Advanced Driver Support Systems, www.eurofot-ip.eu) and TeleFOT (studying the use of nomadic devices, www.telefot.eu) used NDS methods to investigate the behaviour of drivers while driving with new intelligent transport systems. While Naturalistic Driving Studies tend to focus on crash-explanatory factors, Field Operational Tests generally focus on evaluation of systems or functions, however, the data collected in both types of studies can be used for many alternative purposes, such as analysis of Environment, Efficiency and Mobility impacts.

In 2012 the European Commission decided to fund the first large scale European Naturalistic Driving Study: the UDRIVE project (www.udrive.eu). The acronym stands for “eUropean naturalistic Driving and Riding for Infrastructure and Vehicle safety and Environment”. UDRIVE’s objectives are two-fold: to identify well-founded and tailored measures to improve road safety, and to identify approaches for reducing harmful emissions and fuel consumption in order to make road traffic more sustainable.
In this paper we will first provide some more details on the project. Next we will describe the methodology the project followed to design this complex study. The process of defining research questions, and the results will be discussed. The research questions form the basis of the study design, and of the specification of the recording equipment. Both will be described. An outline will be given of how the study will be conducted at seven different European Operation sites. Finally, it will be explained how the data gathered in the project will be used, and what impact is to be expected. Although the project has yet to start collecting data from drivers, we consider the design of the study as a major result which may help other initiatives to set up similar studies.

2. The UDRIVE project

The scientific and technical aims of UDRIVE are describing and quantifying road user behaviour in different European regions, in regular conditions and (near-)crashes, and providing a quantified estimate of the risk of particular safety-critical behaviours, focussing especially on the prevalence and effects of driver states, such as distraction and inattention, and the interactions between drivers/riders and high risk groups like pedestrians and cyclists. In addition, it aims at describing and quantifying road user behaviour in relation to emission levels and fuel consumption, focussing in particular on the effects of driving style, road and road network characteristics, and traffic conditions such as congestion, impaired visibility or adverse weather. The UDRIVE project does not only investigate the behaviour of drivers of passenger cars, but also that of truck drivers and motorcycle riders.

UDRIVE aims to provide recommendations for safety and sustainability measures related to regulation, enforcement, driver awareness, driver training, and road design. The UDRIVE results may lead to improved driver behaviour models and risk functions which can be used in traffic simulations. The project aims to making the traffic system safer and more sustainable, and identifying new approaches, measures and tools. Main results from the project will be the definition of measurable safety and environmental performance indicators for monitoring developments over time, the improvement of existing models of driver behaviour to be used for e.g. predicting effects of safety and environmental measures, and traffic flow simulations, and applications in commercial transport, including driver support systems and targeted training for safer and more fuel efficient driving. Finally UDRIVE will demonstrate how ND data can be used for commercial purposes, including safety and sustainability applications, behaviour based programmes and business models for future data collection.

The project is a 4 year large-scale integrating project, funded under the 7th EU Framework Programme. It started in 2012 and will end in 2016. Projects partners are the SWOV (coordinator), BASt, CDV, CEESAR, CIDAUT, DLR, ERTICO, FIA, IBDIM, IFSTTAR, KFV, LAB, RENAULT, Loughborough University, Or Yarok, SAFER, TU Chemnitz, TNO, University of Leeds, VOLVO. More information may be found at the project website: http://www.udrive.eu/.

3. Methodology

The methodology to plan, design and conduct the ND study, and to analyse the results, is derived from the FESTA methodology. This methodology was originally developed to conduct Field Operational Tests (Barnard & Carsten, 2010). In the European FESTA project (Field opErational teSt supporT Action) by a consortium of a large number of partners, both industrial and academic. The methodology is currently maintained by the European Support Action FOT-Net (www.fot-net.eu), and has been revised several times. Naturalistic Driving Studies are now included in this methodology. NDS and FOTs may be seen as different methods because the study design is different (participant selection, experimental conditions, vehicle sample etc.), and the research questions and hypotheses are different. In particular, the main difference relates to the degree of experimental control in the study. However, there is also substantial overlap, especially in the structured approach to develop the study, and the tools and analysis method used. Between the FOT and the NDS lies the Naturalistic FOT. This last one uses observation in a natural setting, typically to evaluate the relationship between driver-, vehicle-, or environment factors and crash risk, driving behaviour, and the effectiveness of countermeasures (Victor et al. 2010).
In Figure 1 the FESTA methodology is summarised. The methodology is described in detail in the FESTA handbook (FESTA, 2011). The methodology consists of a process which systematically details the steps to be taken to set-up the test (the left-hand side of the V), the actual data acquisition (the bottom of the V), and the analysis of the data and evaluation and interpretation of the results (the right-hand side of the V). The first part of the methodology to define the study follows a systematic research oriented approach. For Field Operational Tests the process starts with defining functions and systems to be investigated (for example, a forward collision warning), the use cases have to be defined, specific events in which a system is expected to behave according to the specified function, for example, car following. For Naturalistic Driving studies, this is not relevant, so the methodology starts with the next step: the definition of research questions and hypotheses. Defining research questions is a major step; below we will describe how this step was taken in UDRIVE. The study is then designed in detail and performance indicators are selected. Performance indicators are quantitative or qualitative indicators, monitored at regular or irregular intervals, and can be compared to one or more criteria (for example acceleration). The next step determines which specific measures and sensors to use. When the whole test is defined, the actual data acquisition can take place. Data will be stored in a database and analysed. Analysis leads to evaluation of whether the research questions have been answered. In a FOT the next steps concern evaluating the functioning of the systems and answering the question what the impact would be if the system were fully deployed in a large proportion of vehicles. Naturalistic driving studies have a different aim, and therefore these steps are not relevant. However, NDS also have scaling up activities, such as the aggregation of results, and the analysis of the implications of these results. For example, one can use findings to identify new and more efficient measures and tools to improve safety and sustainability of road transport. Another impact area is to demonstrate how naturalistic driving can be used in the industrial development of safety and sustainability functions and services.

At different points ethical and legal questions should be addressed, such as the privacy of the participants, responsibilities and data ownership. As in NDS participants are observed for a long period by data acquisition systems including video cameras, and data may also be gathered from the environment and other road users. It should be clear that these are not trivial issues. Moreover, in a European context a NDS project has to deal with different regulations and legislation in the member states.
In the FESTA methodology, two processes are depicted outside the V. The first is the horizontal Context bar. This includes (see also Sagberg et al, 2010) the identification of stakeholders, the selection of topics, the dissemination activities, and last but not least the identification of constraints, such as available technologies and budgets. As a Naturalistic Driving study is a long, complex and costly project, this part of the methodology is crucial for the success of the project. The FOTIP is depicted in the vertical bar: this is the FOT Implementation Plan, the FESTA handbook provides a wealth of practical details and a checklist to ensure that all necessary steps and actions are taken care of.

The UDRIVE project follows the FESTA methodology and has divided its subprojects (SP) according to (clusters of) FESTA V steps, see figure 2.

![Figure 2 The UDRIVE work programme](image)

Additional to the classical FESTA V, and indicated by the arrows between SP1 and SP4 and SP5 in figure 2, a close link between these SPs was established explicitly from the beginning in order to ensure that the analysis goals lying behind the research questions are clear to the analysts when analysing the data later on, and that known analysis constraints of NDS data are considered already during the development of the research questions.

### 4. Research questions and required data

Research questions were developed in UDRIVE on the basis of a review of existing research activities relevant for traffic safety and eco-driving. These research questions address a variety of factors. The elaboration of questions was done considering the following thematic areas:

1. Crash causation and risk
2. Normal driving/everyday driving
3. Distraction and inattention
4. Vulnerable road users
5. Driving style and eco-driving

For the analysis performed in UDRIVE a selection of these questions had to be made according to the resources, technical possibilities and data collected. This first comprehensive list enables us to derive technical requirements for the data acquisition system, and thus lay the foundation to enable researchers to address further
questions even after the end of the project. This list contains 39 questions distributed over the five thematic areas. Some examples of these questions are:

1. Research questions on crash causation and risk
   RQ2.1: To what extent are driver factors associated with risky behaviour?

2. Research questions on normal/everyday driving
   RQ2.4: To what extent are driver assistance systems used?

3. Research questions on distraction and inattention
   RQ3.3: What factors determine how drivers proactively allocate their attention in anticipation of how a driving situation will unfold?

4. Research questions on vulnerable road users
   RQ4.2: What are the conditions – as observed from the driver's vehicle – that determine whether the driver will yield to the pedestrian or the pedestrian to the driver?

5. Research questions on driving style and eco-driving
   RQ5.1: When do drivers brake and is it necessary to brake in each instance?

Certain questions will help addressing several goals (like for example RQ 2.1 will permit to analyse the risk of behaviour, manoeuvres, and contextual factors, and also study the influence of personality factors on the occurrence of critical events).

In order to answer the research questions the type of data needed to conduct the naturalistic driving study was also defined. For each thematic area the scenarios, the candidate Safety Critical Events (SCE), the baseline, and further characteristics of the dataset were identified. In terms of scenarios, the context in which data needs to be acquired was described. It includes the tasks performed by the road user (e.g. car following, negotiating bends; overtaking; free flow), and the characteristics of the environment (road type, traffic conditions, time of day). With regard to the candidate SCEs it was described which critical situations are relevant to be analysed, and how these can be selected (within boundaries). Furthermore, the baseline scenarios to be used as a comparable reference, and the characteristics of the dataset were clarified, in order to better understand how the research questions can be addressed.

5. Study plan

The research questions and the detailed data needs, as described in the section above, formed the basis for designing a study plan. The field trials will involve three types of vehicles, partners across seven countries, and aim at the recruitment of 290 participants in total. For the cars we aim to recruit drivers from households were more than one person drives the same vehicle.

<table>
<thead>
<tr>
<th>Type of vehicle</th>
<th>Country</th>
<th>Fleet size (number of DAS)</th>
<th>Number of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>France</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Germany</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Poland</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>UK</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td>Powered Two-wheelers</td>
<td>Austria</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Spain</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Truck</td>
<td>Netherlands</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

290

Requirements for the participants were defined with regards to experience, age, and environmental exposure. These factors are seen as the most relevant for the research questions, as these driver characteristics are known from the literature to be of significant influence on driving behaviour.
With respect to vehicles, engine size and intended travel patterns are the primary focuses. Two types of passenger cars, one type of powered two-wheeler, and two types of trucks will be included. Previous research shows that a range of driver personality characteristics strongly influence driver behaviour. Therefore, it is of great importance to measure these characteristics also by applying a set of standard questionnaires. These driver characteristics describe the respondents’ tendency to adopt selected target behaviours, such as violation, aggressiveness, compliance with traffic rules, anxiety, etc., or to show a generic tendency towards risky driving.

6. Data acquisition

Research questions and study plan fed into the specification (and subsequent procurement process). One common data acquisition system (DAS) will be used for data collection. This common DAS will be tailored to the needs of the project, and allow for the necessary adaptations to fit to the different vehicle types (e.g. rugged sensors or cameras for motorcycles or similar).

Basically, the DAS is composed of:
- 5 to 8 cameras depending on the vehicle type
- 1 smart camera for environment recording (except for motorcycles)
- CAN interface depending on the vehicle type
- 1 GPS/3G antenna
- 1 Accelerometer/Gyroscope
- 1 speed sensor (for motorcycles)

The amount of data collected implies a smart and efficient data management. The collected data will initially be collated at seven Operation Sites (OS), pre-processed at three Local Data Centres (LDCs) and then delivered to the Central Data Centre (CDC). Pre-processing corresponds to a harmonisation of all data. In addition, the LDCs are responsible for performing map matching (i.e., the addition of road and infrastructure attributes from digital maps). Thus the data is only temporarily stored at the LDCs in order to perform the pre-processing and map matching.

All collected and pre-processed data will be stored at a central data centre (CDC), so that all partners will have access to all collected data. The CDC will support data analysis, and, consequently, also store the results of processing, data annotation and analysis.

To facilitate processing and analysis of the data, a common software toolset will be developed. This toolset will separate data management tasks from analysis tasks and hence will not impose requirements of data management on the analysts themselves. Instead the analysts can concentrate on developing and implementing algorithms, without having to deal with the low level architecture of the database and data storage.

The UDRIVE data flow is given in Figure 3.
7. Expected results and conclusions

On the Operation Sites in the seven countries participants will be recruited and their vehicles will be instrumented with the DAS. They will drive and ride with this equipment for 21 months, and their data will be gathered. The operations at the sites are coordinated and the way in which participants are briefed and data are collected will be the same at all sites (except for differences related to the type of vehicle). This will ensure that UDRIVE functions as one study at multiple locations instead of seven different studies, and will make data comparable and complementary.

The data will be analysed and the results will provide answers to the research questions. We expect to find confirmation about driver behaviour already known from observations by human observers and from experimental and simulator studies. We also expect to find new insight and more precise insight on what is actually happening on European roads and how drivers and riders behave in normal situations, in critical and dangerous situations, and in interaction with other road-users. This knowledge is needed by stakeholders who make decisions about transport policies, development of new intelligent transport systems, and for training of drivers and riders. UDRIVE aims to provide recommendations for safety and sustainability measures related to regulation, enforcement, driver awareness, driver training, and road design. The UDRIVE results may lead to improved driver behaviour models and risk functions which can be used for traffic simulations. Finally UDRIVE will demonstrate how ND data can be used for commercial purposes, including safety and sustainability applications, behaviour based programmes and business models for future data collection.

After the end of the project the UDRIVE project will offer access (within the bounds of legal and ethical restrictions) to the collected data so that it can be consulted and used for subsequent analyses by road safety and environmental experts from all over the world. This will enable the exploitation of the data beyond the scope of the UDRIVE project.
Acknowledgements

UDRIVE is co-funded by the European Commission, DG Research and Innovation, in the 7th Framework Programme, GA No. 314050. The authors would like to thank all UDRIVE project partners for their contributions to the study.

References


