



SEVENTH FRAMEWORK
PROGRAMME
THEME 7 TRANSPORT



Grant Agreement Number:

233597

Acronym:

PROLOGUE

Full Title:

PROmoting real **Life** **O**bservations for **G**aining
Understanding of road user behaviour in **E**urope

Funding Scheme:

Collaborative project small scale

Project Co-ordinator

SWOV Institute for Road Safety Research

Duration:

01/08/2009 to 31/07/2011

Project Website:

www.prologue-eu.eu



Methodological and organizational issues and requirements for ND studies

Deliverable D2.2

Authors: Groenewoud, C., Schoen, E., Malone, K., Jonkers, E., Hoedemaeker, M., Hogema, J. (TNO)

Status: Final draft submitted to EC

Date: 25 May 2010

Dissemination Level: Public

Project Consortium

1		SWOV Institute for Road Safety Research (project co-ordinator)	NL
2		CERTH/HIT Hellenic Institute of Transport	GR
3		KfV Kuratorium für Verkehrssicherheit	A
4		Loughborough University	UK
5		Or Yarok	ISR
6		Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek - TNO	NL
7		TØI Institute of Transport Economics	NO
8		Test & Training International Planning and Service GmbH	A
9		Universitat de València	ES

Document Reference

Groenewoud, C., Schoen, E., Malone, K., Jonkers, E., Hoedemaeker, M., Hogema, J. (2010). *Methodological and organizational issues and requirements for ND studies*. PROLOGUE Deliverable D2.2. TNO, Soesterberg, The Netherlands.

Abstract

Naturalistic driving observation is a relatively new method for studying road safety issues, a method by which one can objectively observe various driver- and accident related behaviour. Typically, participants get their own vehicles equipped with some sort of data logging device that can record various driving behaviours such as speed, braking, lane keeping/variations, acceleration, deceleration etc., as well as one or more video cameras. In this way normal drivers are observed in their normal driving context while driving their own vehicles. Optimally, this allows for observation of the driver, vehicle, road and traffic environments and interaction between these factors.

The main objective of PROLOGUE is to demonstrate the usefulness, value, and feasibility of conducting naturalistic driving observation studies in a European context in order to investigate traffic safety of road users, as well as other traffic related issues such as eco-driving and traffic flow/traffic management.

The current deliverable describes the methodological issues related to naturalistic driving studies. It describes the experimental procedures, variables to be measured, experimental design, statistical analysis methods, organizational issues and legal and ethical issues for naturalistic studies. Maximal use is made of the extensive knowledge and experience that comes from the EU projects FESTA and EuroFOT, the 100car study and the SHRP2 preparatory safety studies.

Table of Contents

Executive Summary	5
1 Introduction	7
2 Procedures for selection of vehicles and participants	10
2.1 Selection of vehicles	10
2.2 Selection of participants	11
2.3 Geographical location	14
2.4 Recommendations	15
3 Variables to be measured	17
3.1 Driver variables	17
3.2 Vehicle variables	20
3.3 Situational conditions	22
4 Design	25
4.1 Study design	25
4.2 Accuracy and validity	26
4.3 Recommendations	26
5 Statistical analysis methods	27
5.1 Statistical insights from related studies	27
5.2 Statistical issues in ND research	29
6 Organizational issues	41
6.1 Organizational issues – planning	41
6.2 Organizational issues – stakeholders	44
6.3 Recommendations	47
7 Legal and Ethical issues	48
7.1 Relevant issues	48
7.2 Experience	51
7.3 Recommendations	52
References	54
Appendix I: Informed consent for drivers of leased vehicles	57
Appendix II: Informed consent for drivers of private vehicles	66
List of Figures	75
List of Tables	76
List of Abbreviations	77

Executive Summary

Naturalistic driving observation is a relatively new method for studying road safety issues, a method by which one can objectively observe various driver- and accident related behaviour. Typically, participants get their own vehicles equipped with some sort of data logging device that can record various driving behaviours such as speed, braking, lane keeping/variations, acceleration, deceleration etc., as well as one or more video cameras. In this way normal drivers are observed in their normal driving context while driving their own vehicles. Optimally, this allows for observation of the driver, vehicle, road and traffic environments and interaction between these factors.

The main objective of PROLOGUE is to demonstrate the usefulness, value, and feasibility of conducting naturalistic driving observation studies in a European context in order to investigate traffic safety of road users, as well as other traffic related issues such as eco-driving and traffic flow/traffic management.

The current deliverable describes the methodological issues related to naturalistic driving studies. It describes the experimental procedures, variables to be measured, experimental design, statistical analysis methods, organizational issues and legal and ethical issues for naturalistic studies. Maximal use is made of the extensive knowledge and experience that comes from the EU projects FESTA and EuroFOT, the 100 car study and the SHRP2 preparatory safety studies.

Procedures for selection of vehicles and participants:

When conducting a Naturalistic Driving study (ND-study), one of the first issues to consider is the selection of vehicles and participants. Relevant research questions are how driver characteristics and vehicle type relate to driving behaviour and accidents, including interaction between vehicles and vehicle versus vulnerable road users. In general, with very large sample sizes the chance of finding an effect is increased. However, there are two major drawbacks on just using large sample sizes; first every driver/participant needs a car equipped with the system and with a data logging system, which is expensive and second, we might find very small effects to be statistically significant, but probably not relevant. The appropriate sample size for a ND-study depends on a number of choices that have to be made in the final set-up. These are, for instance, the hypotheses that are going to be tested and the design choice.

Participant age, gender, and annual mileage have important implications when considering specific types of crashes expected to occur during the data collection period.

Psychometric measurements are generally included in order to relate psychological factors to driving behaviour, rather than select participants per se. Pre-screening participants according to a personality trait/attitude using psychometric instruments will allow the researcher to ensure that a range of drivers with the desired characteristics are included within the study.

Variables to be measured:

The deliverable provides a comprehensive list of variables that are relevant to be measured during a ND study. These variables are grouped by driver characteristics (e.g. age, gender), driver behaviour (e.g. number of lane changes), driver state (e.g. mental workload), driver distraction (e.g. time not looking at the forward field of view), vehicle variables (e.g. speed, car make) and environmental variables (e.g. road type, traffic density).

Design:

Another important issue is the design setup of an ND-study. Although naturalistic driving studies are typically observational, it still makes good sense to consider them in terms of experimental designs where an 'experimental group' is compared with a 'control group' in order to uncover the reasons why events did or did not happen. A case-crossover design is therefore recommended when conducting an ND-study.

Statistical analysis methods:

Usually, there are large amounts of data to be processed in ND studies. A major issue here is to spot important events in the data so as to relate these events to interesting explanatory variables. It is desirable to develop an automated procedure to spot these events. There are statistical tools that help developing an efficient procedure that minimizes false positive events and also minimize missed events.

An ND study will be set up with specific research questions in mind. These questions can only be addressed after data collection and proper event labelling has taken place. In view of the variability of data, answering the research questions nearly always involves the statistical modelling of data. That is, some response variable of interest is linked to possible explanatory variables using a statistical model. The size of the model's coefficients compared to their standard error is used as an evidence, or absence of evidence, for a supposed effect.

In this deliverable, we discuss statistical design considerations, statistical tools for recognizing events, and statistical modelling issues for ND studies.

Organisational issues:

For a ND study to proceed smoothly, good planning – which contains all scientific, technical, administrative and procedural activities and tasks that are needed to successfully complete the study – is essential. What we learn from previous experience is that usually studies do have such a plan. The practical realization differs per study.

From the literature we can conclude that it is important to do a stakeholder analysis, and to involve the stakeholders during the whole project. Good communication is essential.

Legal and ethical issues:

Carrying out a ND study gives rise to a number of legal and ethical issues. These issues need to be tackled to ensure the privacy of the participants, to ensure that the vehicles are safe to operate on the public highway, to cover liabilities and responsibilities in case of an accident, and so on. The arrangement with the participants needs to be formalized by a legal contract or letter of agreement. It is recommended to only release information when a court demands it (in case of an accident).

It is recommended to make a risk assessment plan and obtain legal advice at an early stage of the project.

1 Introduction

The term “Naturalistic Driving” refers to a relatively new technique involving the unobtrusive collection of driver behaviour in relation to the driving task in naturalistic settings. The methodology generally involves drivers using their own vehicles on a day-to-day level and their usage and driver behaviour is recorded usually by video camera. The data that is captured using this method can then be analysed in a number of ways. Perhaps the most useful aspect of the data recording is that it allows researchers to analyse driver behaviour in relation to various “critical incidents” that may occur during particular journeys. For example, a driver may need to swerve and decelerate rapidly in order to avoid a collision. The data recordings allow insight into the driver behaviour this preceded the critical incident and issues such as distraction and inattention, which may have been a factor in creating the critical incident, can then be verified. Collection of these data on a wide-scale will ultimately facilitate the development of countermeasures that can prevent crashes where issues such as distraction and inattention but also other causal factors (such as fatigue) are prevalent. For example, driver warning systems could be implemented into vehicles that are capable of recognising when the driver is indulging in inherently unsafe behaviour. Therefore, the results will lead to a better understanding of road safety and help to achieve an intrinsically safe road transport system by improving safety through improved in-vehicle technologies, development of self-explaining roads and advances in driver training techniques.

Naturalistic Driving observations provide information that would be difficult to obtain otherwise. For well known risk factors such as distraction, inattention and fatigue, naturalistic observations are actually the only method that would provide reliable information about their prevalence and their true relationship with crashes, i.e. the actual risk level. Other issues for which naturalistic observations would be an ideal method include

- The effect of road design characteristics, or weather conditions on the interaction between driver and vehicle;
- Comparing the driving style of specific road user groups, e.g. novice drivers, elderly;
- The prevalence of mobile phone or other in-car information devices and the relationship with particular behaviour patterns or crashes;
- The effect of passengers on distraction, particular driving behaviour or incidents/crashes;
- The interaction between motorised vehicles and vulnerable road users

The main objective of PROLOGUE is to demonstrate the usefulness, value, and feasibility of conducting naturalistic driving observation studies in a European context in order to investigate traffic safety of road users, as well as other traffic related issues such as eco-driving and traffic flow/traffic management.

PROLOGUE Task 2.2

The current deliverable describes the methodological issues related to naturalistic driving studies. It describes the procedures for selection of vehicles and participants, variables to be measured, design, statistical analysis methods, organizational issues and legal and ethical issues for naturalistic studies. Maximal use is made of the extensive knowledge and experience that comes from the EU projects FESTA and EuroFOT, the 100car study and the SHRP2 preparatory safety studies. The reports can be found in the list of references.

In the 100-car study one hundred subjects who commuted into or out of the Northern Virginia/Washington, DC, metropolitan area were recruited as primary drivers. They could either have their private vehicles instrumented or receive an instrumentally leased vehicle to drive for the duration of the study. Drivers were recruited with flyers and classified ads. Drivers under the age of 30 who did not drive a vehicle of an appropriate make and model were given a leased vehicle (22 vehicles), while drivers who drove the appropriate makes and models had their private vehicles instrumented (78 vehicles).

One goal of the study was to record as many crash and near-crash events as possible; this was facilitated by selecting subjects with higher than average crash- or near-crash risk exposure.

Key references to the 100cs include Neale et al. (2002), Dingus et al. (2006), Hanowski et al. (2006), and Klauer et al. (2006a). The research objectives of the study were:

- 1 Characterization of crashes, near-crashes, and incidents for the 100-Car study.
- 2 Quantification of near-crash events.
- 3 Characterization of driver inattention.
- 4 Driver behaviour over time.
- 5 Rear-end conflict causal factors and dynamic conditions.
- 6 Lane change causal factors and dynamic conditions.
- 7 Inattention for rear-end lead-vehicle scenarios.
- 8 Characterize the rear-end scenarios in relation to Heinrich's Triangle.
- 9 Evaluate performance of hardware, sensors, and the data collection system.
- 10 Evaluate the data reduction plan, triggering methods, and data analysis.

In chapter 2 (the procedures for selection of vehicles and participants) the deliverable focuses on driver characteristics as well as self-selection, sample size and geographical issues. Chapter 3, about variables to be measured, gives an overview of the recommended driver, vehicle and situational variables to be measured. Chapter 4 addresses the setup of an ND-study and also validity issues to take into account. Chapter 5 then focuses on the statistical issues when conducting an ND-study. The organizational issues are discussed in chapter 6, followed by the legal and ethical issues in chapter 7.

Definitions

Throughout the document, several terms referring to traffic events will be used that may be difficult to distinguish from each other. Therefore, the most prevalent terms are defined below.

Conflict: A traffic conflict is a traffic situation involving an interaction between two road-users (or between a road-user and the road-environment) that is bound to lead to an imminent collision unless one road-user at least performs an evasive action (Amundsen and Hyden, 1977).

Collision: Any contact with an object, moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated. Includes other vehicles, roadside barriers, objects on or off the roadway, pedestrians, cyclists, animals, etc (100 car study).

Near-miss: A narrowly avoided collision that required a rapid, evasive manoeuvre by the subject vehicle, or any other vehicle, pedestrian, cyclist, animal etc. A rapid, evasive manoeuvre is defined as a steering, braking, accelerating, or any combination of control inputs that approaches the limits of the vehicle capabilities (100 car study).

Incident: An occurrence that could have resulted in a crash or near crash, if the circumstances would have been more adverse (EuroFOT).

2 Procedures for selection of vehicles and participants

When conducting a Naturalistic Driving study (ND-study), one of the first issues to consider is the selection of vehicles and participants. Several aspects must be taken into account when it comes to recruiting participants and selecting suitable vehicles. Most of these aspects are to a great extent depending on the research questions of the study. Relevant research questions are how driver characteristics and vehicle type relate to driving behaviour and accidents, including interaction between vehicles and vehicle versus vulnerable road users. This chapter addresses the issues to take into account when making a selection of vehicles and participants for a ND-study.

Although discussed in separate paragraphs, vehicles and subjects are not independent. Selecting vehicles reduces subject variance and vice versa.

2.1 Selection of vehicles

In the 100 Car study, the Virginia Tech Transportation Institute (VTTI) began determining the vehicle requirements by first establishing the primary criteria that should be considered in selecting vehicles. The ease and cost of data collection system installation is a factor for vehicle selection. Specific vehicle models were limited to six to accommodate data systems installation requirements. This reduced the number of cable and connector sets, custom mounting brackets, and software configurations required for installation. Instrumentation of the cars was quite time-consuming. It took four months before all 100 cars were instrumented.

Several factors were considered when determining what vehicle types were most optimal. The most critical factors included vehicle type, vehicle demographics, vehicle location, data collection system installation issues, information obtainable from the in-vehicle network, and make and model requirements.

Model years were restricted to sequential sets in which the selected models have the fewest design revisions. In addition, using only three vehicle makes increase the commonality of each vehicle's onboard data network, thus requiring fewer software configurations. Recent year models were selected to enhance the data retrieval available over the vehicle's onboard data network. Additional factors were the crash rate of various body types and the distribution of particular vehicles in the Washington, DC/Northern Virginia area.

Although VTTI purposefully chose vehicle makes and models that were popular in the northern Virginia area, this selection still narrowed the number of drivers who could have their private vehicles instrumented. Furthermore, the vehicle models selected for inclusion in the study were typically not driven by younger participants. Even when the younger drivers drove the particular type of vehicle, the vehicle was typically an older model year requiring the creation of different mounting brackets. Rather than continue to create brackets that would only be used for a couple of vehicles, leased vehicles were used by a large portion of this study group. This option represented the most efficient way of incorporating the younger driving population within the study. Any future studies should be aware of the importance of a large and diverse subject pool, and avoid geographical areas with relatively small populations, unless the experimenters are prepared to customize the data acquisition system (DAS) for a large number of vehicles makes and models.

2.2 Selection of participants

Although for a FOT there is often a need to select certain groups of participants, in a ND-study this is not necessarily the case. Depending upon the research questions, there might be a need to select certain groups of participants for a naturalistic driving study, or one needs to assure that the selected participants represent the total population. It is clear, for example, that if one is interested in the effects of a certain system specifically designed to aid elderly drivers, one needs to select a sub group of the total population consisting of elderly drivers. However if one needs to have a representative sample of the total population, one also needs to know whether the sample does not contain, for example, too many elderly, or too many males. In this section, different aspects of participant selection are described as well as the estimation of the number of participants that a sample must contain in order for the study to be robust.

In the 100 car study, participant age, gender, and annual mileage had important implications when considering the number of rear-end crashes expected to occur during the data collection period. The location of the participant's residence or work place was important when considering the difficulties of locating and downloading the data from the vehicles. Vehicle type was very important for private vehicle subjects as each vehicle had to be either be a Toyota (Camry or Corolla) or Ford (Taurus or Explorer).

Since driver attrition and/or removal of the driver from the study were also important aspects of subject management and required additional drivers to be recruited throughout the 100 car study, it is recommended that ND-studies should always have a small number of "standby" participants who can be called on relatively short notice to replace any drivers removed from the study. Thus, the selection of participants and initial screening should continue beyond the placement of the desired number of vehicles on the road.

Subject compliance issues were also present in the 100 car study. Despite numerous efforts to explain the study protocol to drivers and to relay the importance of their compliance, several drivers chose not to do so completely.

These examples point to the importance of the person or persons who are in direct contact with the participants and who serve as the interface between the participants and the organization performing the study. These employees should be well trained in working with participants and with the resolution of the unique issues that are likely to arise in a study of this length and magnitude.

Demographics

Age and gender are the most used driver traits. It is necessary to keep in mind that if a variable is investigated for an "average" male, the results may not be transferable to an elderly female. Therefore age and gender need to be defined in an early phase.

Among the socioeconomic factors, income, education, employment and marital status are often used. These factors can influence not only the effects of the system described in terms of driver/driving behaviour, but also influence the exposure to different situations and the acceptance and willingness to pay for a specific system. Here it is necessary to either make sure that there is homogeneity among the participants or to make sure that there is a random sample that represents the driver population in focus.

Permanent impairments could be related to the vision or hearing but also to impairments that reduce a participant's mobility or cognitive functions. In order to reduce the risk of confounding, drivers with this type of impairment should not be included. On the other hand, if the evaluation has an interest in "working for all" participants with permanent impairments that are common in the population (e.g. colour blindness), the ND-study should consider these.

In some cases, a system under examination could be one that supports, for example, driver fatigue. It is then critical to select participants that drive under relevant conditions, such as night shift workers, drivers with sleep disorders, professional drivers, commuters and young drivers. On the other hand, there is a risk for confounding effects if a participant is using legal drugs which can induce temporary impairments and affect the data collected.

Driving experience

The definition of the driving experience recommended for the participants depends on the question whether or not in-car systems are being used. One of the research questions deals with the question how in-vehicle support systems are related to driving behaviour and accidents. When conducting a study using a relatively new in-car device for usability for example, it is important that participants have no experience with the device. However, driving experience in general and with other devices can be quite important, as interference with the driving task may occur (which could result in unreliable test results and even higher accident risk with inexperienced drivers due to distraction). It has been shown (for example Lansdown, 2002) that novice drivers have higher cognitive workload levels than experienced drivers. This latter group is better able to automate the driving task, which causes them to 'save attention' more than the inexperienced drivers. Thus, it can be said that when a study concerns in-car devices that may cause distraction from the primary driving task, participants should be experienced drivers, i.e. moderate to high-mileage drivers. Furthermore, it would be interesting to know the type of use of the drivers. Driving behaviour of drivers who use the car mostly for commuting trips might be quite different from those who use the car mostly for rural trips.

Depending on the purpose of the study, it can be interesting to select drivers who have an accident history as they may display certain driving characteristics. Similarly, participants without an accident history can also provide interesting information, as a control group. Accident history of the participants needs to be specified in order to diminish side-effects and to have equal groups.

Personality and attitudes

Psychometric measures are generally included in order to relate psychological factors to driving behaviour, rather than select participants per se. However, especially when certain intelligent transport systems are included in the study, there may be some benefit in basing recruitment on personality and attitudes.

Pre-screening participants according to a personality trait/attitude using psychometric instruments will allow the researcher to ensure that a range of drivers with the desired characteristics are included within the study. Finding participants who score near the extremes of measures, however, is likely to be difficult. The "Big-5" test battery however has been used with great success in a study by Ulleberg (2002). He managed to discern 6 subgroups of drivers which he labelled "Considerates", "Anxious", "Socially deviant", "Sensation seekers", "Aggressive" and "Strategic/selfish" drivers. Of these, the "socially deviant" and the "aggressive" drivers were involved more than average in accidents, while "considerates" and "anxious" were lower than average. "Sensation seekers" and "strategic/selfish" were about average. This extended variability of driver traits is very important to have in mind because researchers in the field of traffic safety tend to "dichotomize" the driver population, i.e. to regard drivers either as "normals" or as "sensation seekers", with the latter being the one with increased number of accidents. This hypothesis is, however, also questionable, see for example Jonah (1997), who states that sensation seekers could be frequently speeding without having more

accidents than average. However, when sensation seeking is linked to social deviant personality traits, the number of accidents is elevated.

Having selected appropriate instruments for recruitment, however, it is possible to identify certain groups or individuals that are more likely to exhibit a particular trait or attitude. For example, if researchers are interested in recruiting drivers who express positive attitudes towards speeding, the literature would suggest targeting young males. Male drivers perceive the negative outcomes of speeding as less likely than female drivers (Parker, Manstead, Stradling, & Reason, 1992a), younger male drivers perceive greater social pressure to speed (Conner, Smith, & McMillan, 2003) and younger drivers evaluate the positive outcomes of speeding more positively than older drivers (Parker et al., 1992a). Thus targeting these demographics would produce a sample of drivers with the desired attitudes.

Having administered the chosen psychometric instrument, researchers must then decide how to classify participants. Generally there is little consensus about how best to categorize drivers according to these measures. Many studies recruit a sample of drivers and dichotomise participants into low and high scorers based on a median split. The limitations of imposing such an artificial design however are obvious. Take, for example, a sensation seeking scale, it is possible that the recruitment procedures would attract a sample of drivers who all scored low on the scale. Simply dividing participants on the median of these scores would wrongly classify half of the sample as 'high' sensation seekers. Elsewhere, others have defined high and low scorers as those participants scoring in the upper and lower quartiles. Unfortunately for many psychometric instruments there exists no established point between 'high' and 'low' in the literature.

When screening using a particular scale, it is important to apply a theoretically justified approach to the categorization of participants. In FESTA, some examples of good practice are given. These are the sensation seeking scale Form V (Rudin-Brown & Parker, 2004; Zuckerman, 1994), the locus of control (Rudin-Brown & Parker, 2004) and the theory of planned behaviour. See FESTA deliverable 2.3 for more details about the criteria of these scales.

Before deciding to recruit on personality/attitudinal measures, researchers should consider that when increasing the inclusion criteria for any study it is inevitable that there will be a progressive shrinking of the research participant population. This 'funnel effect' lowers the number of participants eligible to take part within the research. It may therefore be necessary to screen a large number of drivers in order to recruit a relatively small number of participants with the appropriate characteristics, particularly since certain individuals will be less inclined to volunteer to trial certain systems. Inevitably selecting participants on additional measures such as these will increase the burden associated with the recruitment phase of any study.

Self-selection

In general, the participation in an ND-study is voluntary which means that the strategy of recruiting participants can have a biasing effect. For example, offering financial compensation might be an incentive for participants with a low income.

Zhou and Lyles (1997) investigated self-selection bias in driver performance studies. They found that, compared with nonparticipants, participants in performance studies are more active, more likely to travel and drive, less likely to avoid driving in certain circumstances, and less likely to have vision problems. They conclude: "The implication is that project participants represent more highly mobile and confident drivers than would be found in a random sample of the general population. However, project participants also had higher percentages of total accidents and violation points and were involved in more severe accidents than nonparticipants. These problems may be somewhat mitigated, though, by higher driving exposure for participant drivers".

The application of conventional measures of experimental control becomes more and more impossible as the setting moves from the lab to the field. For this reason, ND-studies can (by definition) never be true experiments but could be considered as quasi-experiments. Ruling out the threats to internal validity makes it necessary in this case to have well founded assumptions or hypothesis about external events (e.g. legislative measures) or influences on participants driving behaviour (e.g. increase of the oil price) during the study period and a rather complete documentation. This should make it possible to test (and at best rule out) at least a posterior alternative hypotheses to the initially formulated predictions.

2.2.1 Sample size

In general, with very large sample sizes the chance of finding an effect is increased. However, there are two major drawbacks on just using large sample sizes; first every driver/participant needs a car equipped with the system and with a data logging system, which is expensive and second, we might find very small effects to be statistically significant, but probably not relevant.

The appropriate sample size for a ND-study depends on a number of choices that have to be made in the final set-up. These are, for instance, the hypotheses that are going to be tested and the design choice.

In Europe, USA and Japan, a number of field operational tests have been conducted. The sample size in these FOTs differed widely. In most of the cases practical issues, as the availability of equipped cars and data logging systems, influenced these choices. When trying to ensure that the chosen sample size is representative for the behaviour of a group of drivers and that it is possible to statistically prove effects that exist, one should undertake a power analysis.

2.3 Geographical location

A large number of potential issues to do with the geographical location of a ND-study can be identified, which will have varying degrees of importance within any ND-study. The most important point in relation to the geographical area is that it must be chosen based specifically on the objectives of the ND-study, and in particular, in relation to the validity of the data that is being collected. For example, drivers (and what they are used to) in Sweden and Italy differ in many ways (other cars, different traffic situations, other weather conditions etc). There are two overall considerations:

- Do you need to consider a particular geographical aspect because it is relevant to the types of vehicles and or systems being studied?
- Does a geographical aspect need to be considered to ensure that the results obtained can be generalized to the wider 'population' of interest (i.e. external validity)?

The starting point is to consider the overall objectives of the study, including the types of cars and systems that will be incorporated into the trial. The factors above should then be discussed, one-by-one, in a multidisciplinary team, to assess whether they are key variables that need to be included within a ND-study. When discussing whether geographical variables need to be included within a ND-study, you need to consider whether to include certain geographic characteristics in order to obtain the data needed. Will a specific geographical factor influence the driver behaviour with a system? It is useful to break this down into sub-questions, e.g.:

- Within the confines of interest within the ND-study, could the driver be exposed to variations in this geographical factor?
- And - could variations in this geographical factor impact on the driver?

The second major consideration is that of generalization of the results. In particular it is necessary to ensure that geographical aspects are included to ensure that the data collected during a specific study can be generalized to the wider population of interest. When designing the geographical environment, one needs to consider each of the above factors in relation to the following question:

If this factor is ignored, is it possible that the results we get would be different to those I would expect to get with my intended population?

The third factor to consider is whether the geographical factor is of particular interest in terms of data analysis. If it is desirable to analyse results according to the presence or absence of a particular factor, then the geographical environment(s) must include that factor (and possibly variation thereof). An example would be an investigation of the use of travel-related information to a driver. A research hypothesis may suggest that the use of this information would depend on the amount of other value-adding information available in the environment – for example that navigation-related functions are particularly valuable in rural areas with relatively few signposts. The study would therefore need to include those rural areas, and non-rural areas for comparison.

Best practice could involve several steps:

- identify the variables you want to capture in the study;
- derive hypotheses you want to test;
- choose a geographical location(s) that allows to test the hypothesis.

2.4 Recommendations

- Create a large and diverse subject pool
- Choose vehicle makes that are popular
- Avoid geographical areas with relatively small populations
- Have a small number of “standby” participants to replace drivers removed from the study
- According to socioeconomic factors, make sure that there is homogeneity among the participants or to make sure that there is a random sample that represents the driver population in focus
- Decide in an early stage whether or not to include drivers with vision or hearing impairment (“working for all” versus reducing confound risk)
- When a study concerns in-car devices that may cause distraction from the primary driving task, participants should be experienced drivers (moderate to high-mileage drivers)
- Be aware of a possible ‘funnel effect’ (increasing inclusion criteria results in shrinking of the participant population) when it comes to selecting on personality traits
- When studying driver characteristics, keep in mind that an ND-study is voluntary which means that this self-selection can have a biasing effect

- Find out if you need to consider a particular geographical aspect because it is relevant to the types of vehicles and or systems being studied
- Find out if a geographical aspect needs to be considered to ensure that the results obtained can be generalized to the wider 'population' of interest (i.e. external validity)

3 Variables to be measured

This section describes the variables to be measured in a functional way. How these variables are technically measured is part of deliverable 2.1.

3.1 Driver variables

The first important level of variables is driver variables. These can be divided into different groups. In the following section, the variables that measure driver characteristics, behaviour and distraction will be discussed.

3.1.1 Driver characteristics

In FESTA, four categories of driver characteristics were distinguished. Although FESTA focussed on field operational tests, the findings from that study are also relevant for naturalistic driving studies.

- Demographic characteristics: gender, age, country, educational level, income, socio-cultural background, life and living situation, etc.
- Driving experience, and driving situation and motivation: experience in years and in mileage, professional, tourist, with or without passengers and children etc.
- Personality traits and physical characteristics: sensation seeking, locus of control, cognitive skills, physical impairments or weaknesses etc.
- Attitudes and intentions: attitudes towards safety, environment, technology etc.

These characteristics are not independent; some are even highly related and influence each other. Combinations of these different characteristics may influence driving behaviour quite differently. For example, an elderly driver with a sensation seeking personality may take much less risk when driving with his grandchildren than when he was young and driving alone.

Characteristics may be stable and unchangeable, such as gender, or more volatile, such as attitude. Some of the driver characteristics can be measured very easily, such as age, but others are more complex, such as personality traits. Even simple demographic characteristics are not always easy to use for classifying drivers into groups, for example drivers who lived in different countries.

Studies often focus on characteristics of individual drivers. However, drivers are not alone on the road. There are other road users and there may be passengers in the car, who may influence the driver's behaviour. For example studies have shown that young male drivers behave differently depending on the presence of passengers and on whether those passengers are male or female (for example Padlo et al., 2005; Preusser et al., 1998). There are also other kinds of influence, such as the (perceived) opinion of important others, such as parents, and more general social influences. So there is an interaction between the characteristics of the individual driver and those of other people.

In general it is useful to gather as many characteristics of drivers as practically possible. Even if no specific impacts are expected of certain characteristics, some outcomes may be explained better with more knowledge about the participants. A minimum set of data such as age, gender, income group and educational level is easy to gather from participants.

Next information is needed about driving experience. The role of the factor “driver experience” has been discussed in the literature in particular within the context of the explanation for the strikingly high crash risk of young novice drivers. However, there is clear evidence that driving experience has a significant impact on individual crash risk even if effects of age are controlled. Generally speaking, there seems to be a dramatic decrease of crash risk during the first months after licensing independently of driver age even if the starting level decreases with increasing driver age (Maycock et al., 1991). The variable “driving experience” describes the amount of practice a driver has gathered while performing the task of driving a vehicle which can be considered as the acquisition of a complex skill. It is usually measured by means of self-reports. The amount of practice, i.e. the mileage of an individual driver can be collected by asking the subject for an estimation of his/her overall mileage since licensing or the current mileage per year. However, beware that these self-reports are not very reliable.

For further understanding of driver behaviour one may consider to use questionnaires on attitudes, driving behaviour and personality traits. A well-known questionnaire about (self-reported) driving behaviour is the Driver Behaviour Questionnaire (DBQ). Some widely used personality tests are the Five Factor Model (FFM) test and the Traffic Locus of Control (TLOC) test (Özkan & Lajunen, 2005). Special attention may be given to the personality trait of sensation seeking, which is correlated with risky driving. The Sensation Seeking Scale (SSS) measures this trait. These questionnaires are available in many different languages, but they are not always standardized and cultural differences may play a role. Personality traits are very easy to measure, just by administering a short questionnaire. However, the concepts and interrelations of factors are very complex, and results should be treated with caution.

Below is a list of driver variables the authors labelled necessary to be measured in a naturalistic driving study per category, based on the literature. These driver characteristics are mostly independent measures or covariates.

Table 3.1: Demographic and personality variables

Stable variables (driver traits)	Unstable variables (driver states)
Age	Physical condition
Gender	Locus of control
Country of living	Self-reported driving behaviour (DBQ)
Education	Attitudes/intentions towards speeding, safety, environment
Income	
Aggressiveness	
Cognitive skills	
Risk perception	
Professional driver	
Driving experience (years, total kilometres, kilometres per year)	
Masculinity/femininity	
Field dependence	
Sensation seeking scale	

Variables in bold print are considered highly recommended

3.1.2 Driving behaviour

The second important category of variables is that of driving behaviour. Variables of this type are especially important when trying to relate driver characteristics to driver actions. In the 100 car study, Dingus et al. (2006) investigated rear-end conflicts in relation to lane changes. They found that several near-misses and incidents occurred when there was a cut-in to the lane in front of the subject vehicle, and also when the subject vehicle changed lane behind a lead vehicle. This suggests that the (lane-changing) behaviour is an important variable in relation to events.

Whereas driver characteristics are mostly independent variables or covariates, driving behaviour variables are mostly dependant measures.

Table 3.2: Variables describing driving behaviour

Driving behaviour variables
Frequency of performed left and right lane changes (number per kilometre and hour)
Frequency of active overtaking (number per kilometre and hour)
Frequency of passive overtaking (number per kilometre and hour)
Deviation from desired lane
Frequency of route changes (number per kilometre and hour)
Travel time uncertainty
Delay
Following/free state profile
Speed profile
Actual route
Use of car horn
System interaction and driving behaviour related responses to alarm/warning
Reaction time to alarm/warning

Variables in bold print are considered highly recommended

3.1.3 Distraction and driver state

Driver distraction is an issue widely studied in many experimental settings. For example, Olson et al. (2009) found that in their study, some type of distraction was listed as a potential contributing factor in more than 80 percent of all safety-critical events. Therefore, in a ND-study, it is a main issue that should absolutely be measured. The variables that must and can be used to do so are listed in the table below.

Table 3.3: Variables describing driver distraction and state

Driver distraction and state variables
Mental workload
Fatigue/drowsiness
Distraction from primary driving task (eye-tracking, glance duration, fixation)
Head-tracking
Number and position of hands on steering wheel
Presence and use of in-car devices (e.g. mobile phone, navigation system etc)
Driver identified events
Presence, number and age of passengers
(moving) object inside vehicle
Performing tasks other than the primary driving task:
- eating/drinking
- adjusting radio or other in-car device (e.g. climate controls, CD etc)
- dialling or texting on mobile phone

Variables in bold print are considered highly recommended

3.2 Vehicle variables

The second important level of variables is vehicle variables. These can be divided into variables that describe vehicle condition and vehicle parameters.

3.2.1 Vehicle condition

The vehicle condition should be measured in order to differentiate between certain groups or to relate to other measures, such as the vehicle parameters as described in the next section.

Table 3.4: Variables describing vehicle condition

Vehicle condition variables
Vehicle type (manufacturer, model, vehicle age)
Vehicle Identification Number (VIN)
Vehicle mass: driver + passengers, load besides driver and passengers, trailer connected or not, amount of fuel in the tank
Presence of safety systems (ACC, LDWS etc.)
Air conditioning: use/not use
Wiper status: use/not use
Other auxiliaries: use/not use
Cooling fan: operating/ not operating
Type of transmission
Type and amount of in-vehicle systems

Variables in bold print are considered highly recommended

3.2.2 Vehicle parameters

The vehicle parameters are a main issue in conducting a ND-study. Many of the parameters stated in the table below are highly recommended to include in the study. These variables can provide information that can be used to identify events such as crashes or near-crashes.

Table 3.5: Variables describing vehicle parameters

Vehicle parameters
Speed
Acceleration (longitudinal, lateral & gyro)
Deceleration (incl. sudden braking)
Percentage throttle
Percentage clutch
Percentage brake
Brake force
Gear position
Steering wheel angle
Turn signal
Lateral position
Lane departure
Time to line crossing (TLC)
Distance to vehicle in front
Distance to vehicle behind
Distance to other surrounding vehicles
Side vehicle detection
Time headway (forward & rear headway detection)
Space headway (forward & rear headway detection)
Time to collision (forward & rear TTC)
Post encroachment time (PET)
Travel time (including stop time)
Travel distance (mileage)
Waiting time at intersections
Friction

Variables in bold print are considered highly recommended

3.3 Situational conditions

FESTA identified a series of additional ‘situational variables’ in addition to the driver behaviour and vehicle variables. These include environmental, road condition, traffic condition and video data variables that must also be measured or recorded in a ND-study as they provide key background information that complements the driver behaviour data and is sometimes needed to derive the driver behaviour data. These situational variables are discussed below.

3.3.1 Environmental variables

The environment is a critical element in a ND-study, since it will determine the data that are collected and the ability to fulfil the objectives.

This section provides recommendations for which environmental variables to include within a ND-study.

Table 3.6: Variables describing environmental variables

Environmental variables
Precipitation (i.e. snow, rain, fog)
Time: date; time of the day
Daylight/dark conditions
Air pressure (measured with vehicle sensor)
Air temperature (measured with vehicle sensor)
Humidity (measured with vehicle sensor)
Wind speed

Variables in bold print are considered highly recommended

3.3.2 Road condition

Three main road categories should be differentiated (urban, rural, and motorway) and within each category of road sub-categories can be taken into account depending on the level of detail wished in the analyses. At a more general level, we may classify the roads according to their structural (for example, number of lane, horizontal and longitudinal profile, etc.), legal (speed limits, right of way) and functional (convergent or divergent; stable or transitional) characteristics.

Ideally a map and a database of the region of deployment of the study should be established in order to reduce the time needed afterwards for collecting this type of data (on the basis of the video recording of the road scene). An electronic map containing at least the type of roads and the speed limits in force (and location of speed cameras) would greatly facilitate the task.

Table 3.7: Variables describing road condition

Road condition variables
Road surface conditions (in Sweden by use of the winter model)
Road distance (a GPS will probably not have accuracy enough)
Road and traffic conditions based on GPS and time:
Gradient; horizontal curve; junction; roughness;
Macro texture
Road type
Environment (Urban/interurban/rural)
Number of lanes
Width of lanes
Base capacity and saturation flows
Central barrier
Sight distance
Speed limit
Location of speed cameras
Current traffic management: road markings, signs, rumble stripes, etc
Bus stops or parked cars along the street
Hard shoulder
Intersections:
- frequencies
- intersections types (signals/roundabouts/yield/stop)
- exit roads
Number of stops on route

Variables in bold print are considered highly recommended

3.3.3 Traffic conditions

With regard to traffic conditions, a distinction needs to be made between:

- Traffic conditions in a general sense, which characterize a general level of constraints and which, in the same manner as the infrastructure zones, define the driving environment;
- Other road users and their behaviour, which characterize an individual level of interaction between the driver and one or more other road users in the driver's immediate proximity.

The traffic, as a general and contextual entity can be characterized using several dimensions, for example density (expressed in terms of the number of vehicles travelling in a given space), speed (the average speed of traffic) and composition. This latter one addresses the types of vehicle (light vehicle, heavy vehicle, van, motorcycle) and their relative proportions in a given traffic stream.

The interactions at individual level between the driver and one or more other road users in the immediate vicinity can also be characterized using several dimensions. These could be the category to which they belong (light vehicle, heavy vehicle, van, motorcycle, pedestrians), their speed and acceleration (direction and rate) and their manoeuvres and behaviour (merging into the driver's lane or pulling out into a lane, merging from an entry slip road, braking, etc.).

Some elements of general traffic conditions could be estimated indirectly by using day and time of the day and/or when available by collecting data from relevant road authorities. Most of the relevant traffic and interactions elements need to be collected from the video recording of the road scene (in front and behind the instrumented vehicle). The most important practical consideration is that this video analysis is very time consuming, and you need to be very selective over choosing which sections of data to analyse. This type of analysis should be driven by the hypotheses you are trying to test.

Table 3.8: Variables describing traffic conditions

Traffic condition variables
Traffic density
Speed distribution, average speed and standard deviation
Traffic composition
Traffic signal picture
Traffic flow
Other (unrelated) incidents that may affect traffic flow
Category of road users in vicinity (pedestrian, cyclist, light/heavy vehicle, etc)
Speed/acceleration of road users in vicinity
Behaviour of road users in vicinity

Variables in bold print are considered highly recommended

3.3.4 Video/RADAR data

In addition to all the variables mentioned above, video data needs to be collected. This data will not be used to analyse like the variables described before, but to clarify odd outcomes found during the analysis of the variables. By installing cameras with a view on the face of the driver, the interior of the car, the forward view, the rear view and a view from both sides of the car, more insight to the obtained data can be gained. Radar can be used to identify objects in the front of the cars, their range, and the rates at which the range changes.

Auditory data of conversations and from other activities in the car in the cabin are also strongly recommended to be included. When considering distraction factors, talking with passengers is found to be the most prevalent factor, and the most frequent contributing cause of accidents when comparing with other internal distraction factors such as reading a map, searching for road signs, changing a CD, operating a radio, smoking etcetera (Sagberg, 2001). The data sample could sample the last two minutes and overwrite and start recording again every second minute to protect privacy of conversation content.

4 Design

4.1 Study design

Although naturalistic driving studies are typically observational, it still makes good sense to consider them in terms of experimental designs where an ‘experimental group’ is compared with a ‘control group’ in order to uncover the reasons why events did or did not happen.

Since a naturalistic driving study typically is observational, it is quite natural to compare it with studies performed in epidemiology. The overall strategy used in epidemiology is: observations of events in groups of individuals who share a particular characteristic, comparisons of rates of the events among groups, and then inferences regarding the basis for any differences seen.

In the paradigmatic observational cohort study, the investigator defines two or more groups of people that are free of ‘disease’ and that differ according to the extent of their exposure to a potential cause of disease. These groups are referred to as the study cohorts. When two groups are studied, one is usually thought of as the exposed or index cohort – those individuals who have experienced the assumed causal event or condition – and the other is then thought of as the unexposed, or reference cohort. There may be more than just two cohorts, but each cohort would represent a group with a different level or type of exposure. For example, an occupational cohort study of chemical workers might comprise cohorts of workers in a plant who work in different departments of the plant, with each cohort being exposed to a different set of chemicals. The investigator measures the incidence times and rates of disease in each of the study cohorts, and compares these occurrence measures.

Many cohort studies begin with but a single cohort that is heterogeneous with respect to exposure history. Comparisons of disease experience are made within the cohort across subgroups defined by one or more exposures. Examples include studies of cohorts defined from membership lists or administrative or social units, such as cohorts of doctors or nurses, or cohorts defined from employment records, such as cohorts of factory workers.

Compared to epidemiologic studies, a naturalistic driving study therefore most closely resembles a cohort study. In this case the ‘disease’ could best be renamed an ‘event’ like a crash or a near-crash, and the ‘exposure’ consists of one or more causes (e.g., inattention, age) leading up the event.

Relative risk and latent trajectory models can be used to investigate whether subgroups of drivers in the sample exhibit different event developments over time. Case-crossover designs allow for the comparison of pre-event and non pre-event factors by choosing for each pre-event period a second time period when an event did not occur, and which is as similar as possible to the time period when an event did occur in all other respects. In this design, ‘cases’ are used as their own ‘controls’. See also section 5.2.5 (explaining events).

The case-crossover design is a method to assess the effect of transient exposures on the risk of onset of acute events. Control information for each case is based on his/her past exposure experience (at a time period when events did not occur and which is as similar as possible to the time period when events did occur in all other respects), and a self-matched analysis is conducted. In case-crossover designs the ‘cases’ are therefore used as their own controls, see Mittleman, MacLure, and Robins (1995) and Rothman, Greenland and, Lash (2008). The need for a careful definition of both event and non-event situations is stressed, in order to get proper cause-effect relations.

Observer effect

Observations as a research method involve the act of noting and recording something, such as a phenomenon. The particular strengths of observations are that they capture ongoing processes, e.g. actual behaviours of individuals, habits and routines that are not consciously reflected upon and therefore difficult to elicit by different question based data collection methods. A weakness is that the observation may have an effect on the individual's behaviour. Video-recording may have the same negative effect, even though it can be anticipated that users, over time, will forget the camera and act according to their original habits and routines.

Study duration

Several aspects are important when considering the length of a ND-study. For example, when the number of participants is disappointing, a longer period of data collection is required. It is also likely that the aforementioned observer effect will fade away in time. In the 100 car study, the data set included 12 to 13 months of data collection for each vehicle. Depending on the kind and number of questions to be answered in the study and the number of participants (and vehicles), a minimum duration of 12 to 18 months would be recommended.

4.2 Accuracy and validity

In an experimental study where it is possible for the researcher to manipulate the experimental conditions attempts are made to obtain a random sample from the population in which interest lies because it is then straightforward to generalize the results found in the sample to the population. Accuracy in estimation implies that the value of the parameter that is the object of measurement is estimated with little error. Errors in estimation are traditionally classified as either random or systematic. Although random errors in the sampling and measurement of subjects can lead to systematic errors in the final estimates, important principles of study design emerge from separate consideration of sources of random and systematic errors.

Systematic errors in estimates are commonly referred to as biases; the opposite of bias is validity, so that an estimate that has little systematic error may be described as valid. Analogously, the opposite of random error is precision (which is related to power), and an estimate with little random error may be described as precise. Validity and precision are both components of accuracy.

As a general rule, the results of a study should allow a clear decision if the hypothesized relationships between variables exist or not, i.e. if the hypotheses can be accepted or has to be rejected. In the best case, the researcher is able to attribute the changes he/she observed at the dependent variable without any doubts to the manipulation of the independent variable. The internal validity of an experimental or quasi-experimental study describes the extent to which this inference is unequivocally possible due to the study design. Another aspect is external validity which describes the extent to which results can be generalized to other persons, situations and points in time. The internal validity of a study increases to the extent to which such alternative explanations can be ruled out. See section 5.2.7 for more details about statistic validity issues.

4.3 Recommendations

- Use a case-crossover design for the ND-study.
- The length of the study should be at least 12 months.

5 Statistical analysis methods

This chapter discusses statistical methods that can be used to explain behaviour or events such as accidents or near-accidents, and methods that can help in the automatic detection of these events from signals measured during a trip. We first discuss statistical insights from the 100 car study and from FESTA (Section 5.1). We supplement these insights with our own subject-matter knowledge to give an overview of statistical analysis issues in Section 5.2. In particular, we discuss statistical design issues for selection of participants or road stretches to be included in an ND study (Subsection 5.2.1), statistical tools to optimize automated event recognition (Subsection 5.2.2), methods to control false positive detections in multiple statistical testing (Subsection 5.2.3), a general framework for statistical models that are needed to answer the research questions of WP 1 (Subsection 5.2.4), statistical issues in explaining events (Subsection 5.2.5), and statistical issues in explaining behaviour (subsection 5.2.6). We conclude with a discussion on reliability and validity of the study results (Subsection 5.2.7).

5.1 Statistical insights from related studies

5.1.1 The 100 car study

Neale et al. (2002) is the report on the study design. Statistical methodology is addressed only in very general terms. Dingus et al. (2006) report on the results of the 100cs. The report addresses in particular the results for the above goals 1 up to 9. (As far as we can see, goal 10 is not yet publicly available.) In reporting the results, there was a clear emphasis on readability for the general user. For this reason, the report presents summary frequency tables relevant to the respective goals.

The reports of Hanowski et al. (2006) and Klauer et al. (2006a) focus on specific issues in the 100cs. Hanowski et al. addresses interactions between light vehicles and heavy vehicles. Here too, we can learn little about the statistical methods underpinning the conclusions.

Finally, Klauer et al. (2006a) addresses the issue of driver inattention. Six specific objectives were identified:

- 1 What are the prevalence as well as the types of driver inattention in which drivers engage during their daily driving? What is the relative near-crash/crash risk of driving while engaging in an inattentive task? Is the relative near-crash/crash risk different for different types of secondary tasks?
- 2 What are the environmental conditions associated with driver choice of engagement in secondary tasks or driving while drowsy? What are the relative risks of a crash or near-crash when engaging in driving inattention while encountering these environmental conditions?
- 3 Determine the differences in demographic data, test battery results, and performance-based measures between inattentive and attentive drivers. How might this knowledge be used to mitigate the potential negative consequences of inattentive driving behaviours? Could this information be used to improve driver education courses or traffic schools?
- 4 What is the relationship between measures obtained from pre-test batteries (e.g., a life stress test) and the frequency of engagement in distracting behaviours while driving? Does there appear to be any correlation between willing-

- ness to engage in distracting behaviours and measures obtained from pre-test batteries?
- 5 What is the relative near-crash/crash risk of eyes off the forward roadway? Do eyes off the forward roadway significantly affect safety and/or driving performance?
 - 6 Are there differences in driving performance for drivers who are engaging in a distraction task versus those drivers who are attending to driving? Are some of the safety surrogate measures more sensitive to driving performance differences when driving distracted versus other safety surrogate measures?

There are separate chapters treating each of these issues. Several statistical techniques are mentioned explicitly. First, odds ratios (closely related to relative risks) are calculated. These are used to determine differences in attention between baseline time intervals and time intervals where some event took place. Second, t tests are used to determine differences in groups with a low and a high involvement in inattention-related crashes or near-crashes. Further, correlation between variables within these groups is studied. Finally, a discriminant analysis was used to determine which variables help to discern between the groups.

All of the above statistical techniques are related to models that explain events or group membership based on explanatory variables. In Section 5.2.3 of the present document, we elaborate on the specific models and we relate the modelling to the research questions of PROLOGUE as formulated in WP1.

5.1.2 FESTA

The aim of the support action FESTA as launched in the first call of FP7 was to provide guidelines for the conduct of a Field Operational test (FOT). A FOT is defined as a study undertaken to evaluate one or more systems, under normal operating conditions in environments typically encountered by the host vehicle(s) using quasi-experimental methods. A system is some device whose impact is to be assessed by the FOT.

A FOT differs from a naturalistic driving study because there are clear experimental groups with treatments that can be allocated to the participants. Nevertheless, there are some ideas that transfer easily to observational studies.

Lassarre et al. (2008) discuss issues relevant to data analysis and modelling. They propose a chain structure linking a hypothesis to be tested to the outcome of the test. The links of chain are data quality analysis, data processing, performance indicator calculation, hypothesis testing and global assessment. Statistical methods have been described for the calculation of performance indicators, and the testing of hypotheses. As to the performance indicators, there is a suggestion to treat the state of a system such as a vehicle as a Markovian chain. For this purpose, a trip is split up into time intervals. The 'chain' is just a sequence of zeros and ones, denoting whether or not the vehicle is in a particular state at some time interval k , say. A zero corresponds to no event in the interval, and a 1 corresponds to the occurrence of an event. The Markov assumption is:

$$P(S_k | S_{k-1}, S_{k-2}, \dots, S_1) = P(S_k | S_{k-1})$$

(Taylor and Karlin, 1994). In words, the state on interval k given the state on all previous intervals of the trip only depends on the state of the previous interval.

The Markov idea could be valuable for naturalistic studies as well. We return to this issue in Section 5.2.5 when discussing Research Questions 4A-7A.

Because of the focus on comparing several treatment groups, Lassarre et al. (2008) discuss the t test. This test compares variables of two groups defined by the presence or absence of the system. A statistically significant difference in variable points to a

systematic effect of using the system. The authors suggest that the group differences may depend on additional explanatory variables. So it might be beneficial to include all the variables in a statistical model linking a performance indicator to the treatment group and the additional variables. One could extend this idea to observational studies whenever we want to explain performance indicators by potentially causal variables. As the indicators are to be explained by multiple explanatory variables, the technique is now called multiple regression analysis; see Montgomery et al. (2006) for an introductory text.

5.2 Statistical issues in ND research

Any ND study will involve either a selection of participants or a selection of infrastructure elements for inclusion. The selection should include statistical design considerations to define the groups of participants or road stretches.

Usually, there are large amounts of data to be processed in ND studies. A major issue here is to spot important events in the data so as to relate these events to interesting explanatory variables. It is desirable to develop an automated procedure to spot these events. There are statistical tools that help developing an efficient procedure that minimizes false positive events and also minimize missed events.

An ND study will be set up with specific research questions in mind. These questions can only be addressed after data collection and proper event labelling has taken place. In view of the variability of data, answering the research questions nearly always involves the statistical modelling of data. That is, some response variable of interest is linked to possible explanatory variables using a statistical model. If a model coefficient deviates statistically significant from zero, this is taken as evidence for an effect of the corresponding explanatory variable.

There will usually be many statistical tests carried out in an ND study. Hence, one would expect that some tests come out as significant just by random variation. Therefore it is useful to include methods that control the false positive error rate.

Finally, one wants to be sure that the random error in the results is low and that there is no systematic error. So there should be a continuous awareness in the study as to the reliability and validity of the study results.

In this section, we discuss statistical design considerations, statistical tools for recognizing events, methods to control the false positive error rates when multiple statistical test are carried out, statistical modelling issues for ND studies, and the general concepts of reliability and validity of a study.

5.2.1 Statistical design for selection of participants and road stretches

An ND study may be based either on a number of participants or a number of road stretches followed over some period of time. Naturally, there will be questions on the relation between traits of the participants or characteristics of the road stretches and the occurrence of events. To increase the potential of the data for answering these questions, it is recommended to heed principles of experimental design as borrowed from industrial experimentation. We refer to Montgomery (2009) for a general introduction.

To illustrate why experimental design might be important for selection of participants, suppose that we are interested in the effect of the driver's gender on the occurrence of events over the observational period. If we include, say, two women and ten men, we would have a smaller chance in detecting gender differences than if we would include as many women as men. So one principle to heed is that of balance regarding the participants' traits of interest.

The above example will sound trivial to many. A less trivial example bears on the number of traits that can be included in a study without compromising the balance. Indeed, suppose that we select seven traits, with two possible options for each trait. The traits can be participant related as well as vehicle-related. To be more specific, suppose that we want to study gender (m/f), age (20-30, 50-60), education (high, low), sensation seeking scale (high, low), automatic gear (yes, no), vehicle age (high, low), and vehicle mass (high, low). Altogether, this yields $2^7 = 128$ possible combinations. However, using experimental design principles, it is possible to investigate the effects of all these traits in as few as 12 experimental groups with a so-called Plackett-Burman design (Mee, 2009). However, we recommend including several participants in each of the group. The number should be based on the size of the group differences that are to be detected. We recommend performing a statistical power analysis (Cohen, 1988; Bausell and Li, 2002) to decide on the number of participants. Such an analysis gives either the number of participants to detect a pre-specified difference between two groups with a specified probability of detection, or the probability of detection given the group size. An example follows.

Suppose that we think that about 50% of the younger drivers will be involved in a near crash event. Suppose further that we want to detect if older drivers have 20% less chance to be involved in this type of events. If we want to have a probability of detecting such a difference of at least 80%, then there are in total 130 participants needed. Note that these can be distributed events over the groups in the design. For the 12 experimental groups, this implies 11 participants per group.

The general principles outlined for participants also hold for the selection of cross roads or road stretches.

5.2.2 Statistical tools to optimize event recognition

A major issue in the processing of the data collected in an ND study is to spot important events so as to relate these to interesting explanatory variables. It is desirable to develop an automated procedure to spot these events. Trigger variables can be defined that might point to interesting events. For example, a very large longitudinal deceleration might point to a near collision. The problem here is to define exactly what is meant by 'very large' and also which potential trigger variables to use. If we set a low threshold, we end up with too many false positives. If we set up a high threshold, we end up with missed events. If we use too many trigger variables, we can also miss important events. This is because introduction of a new trigger further restricts the time intervals selected as having a potential event.

In the 100 car study, a selection procedure was developed based on a portion of 10% of the data; see Klauer et al. (2006a) for details. Briefly, six trigger variables were defined. Initially, they all were set at a very liberal level. These levels jointly resulted in a large number of putative events. Data reductionists classified all these events as crash, near crash, incident, or invalid. The thresholds were subsequently tightened by trial and error to minimize both the invalid events and the missed events. Note that there are various events considered in the study.

We propose to augment the procedure with statistical design methods as explained in the previous section. In particular, after having set all the trigger variables at their liberal level, we can perform automated selection based on a statistical design. The triggers are the experimental factors, and their liberal and restrictive levels are the settings of the factors. An 'experiment' corresponds to the classification of the valid and invalid events as labelled by the data reductionists into false positives and false negatives by the automated procedure defined by the trigger settings. For six triggers, we could employ a Plackett-Burman design (see previous section) to detect the effect of the trigger on the false alarms and the missed events. Triggers that are not influential are dis-

carded. Triggers that have a substantial effect on the detection are investigated more closely.

After selection of the appropriate triggers and their thresholds, a fully automated detection of the remaining 90% of the data can be performed. However, the events still should be screened manually by the data reductionists.

5.2.3 Multiple statistical testing

One way to control the false-positive rate over the entire study is by adjusting the error rate for the individual tests according to the total number of these tests. A particular simple way of doing this is the Bonferroni adjustment. This adjustment uses a critical value of the individual tests corresponding to the desired experiment-wise false-positive rate divided by the number of tests. So, in a study with 168 tests, say, one would set the individual false-positive rate at $0.01/168$ if the aim is to control the rate over the entire experiment to 0.01. However, this will result in too strict a demand on the individual tests because it uses the total number of tests. What it really would need is the number of tests for variables that do not affect event rates. It is obvious that this number is unknown in advance. It is equally obvious that the unknown number might differ from the total number of tests. For the example with the 168 tests, the Bonferroni adjustment to control the overall false positive rate to 0.01 implies an individual rate of $0.01/168 = 5.95 \times 10^{-5}$. If 50% of the variables would be inert, we would need a rate of twice this number, or 1.19×10^{-4} .

In an attempt to alleviate the problems with the Bonferroni adjustment, Benjamini and Hochberg (1995) proposed to use a so-called false discovery rate (FDR). This is the expected percentage of significant test results that correspond to falsely rejected hypotheses. This is how the procedure is carried out for an FDR of q :

- a. Let m be the number of hypotheses under consideration.
- b. Calculate the m P values and put these in ascending order.
- c. Find the largest value i for which $P(i) \leq i \cdot q / m$, call this value k .
- d. Reject all $H(i)$ $i = 1 \dots k$.

It is easy to make a graphical representation of this procedure. We plot the P values against $i \cdot q / m$. As a reference line, we add $y = x$. We then look for the right most intersection of this line with the line that connects the plotted P values. Significant P values are on the left of the intersection point.

The FDR procedure is evidently less restrictive than the Bonferroni procedure, because it is only the smallest P value (cf. $i=1$) that must obey the Bonferroni adjustment.

If the largest P value (cf. $i=m$) already leads to significance at the q % level, than all the hypotheses under consideration are rejected.

5.2.4 Statistical modelling to address specific research questions

Any ND study involves a good number of vehicles. In most cases each participants is the main driver of the vehicle. In the sequel, we will assume that this is always the case, and data from additional drivers is ignored. Each participant will make multiple trips in the vehicle. So we have trips nested within participant/vehicle. Each trip can be divided in time intervals of some length. So we have time intervals nested within trip nested within participant/vehicle.

The variation of the variables in an ND study can be classified according to the nesting level at which the variation is present; see Fig. 5.1 for a representation. We have participant and vehicle variables that are assumed constant over all observations from one and the same participant/vehicle combinations. These include driver traits such as age, gender and driving experience, and vehicle traits such as type and age (the figure is based on participant traits only).

A second group of variables consists of those that are constant over a particular trip, but variable within a participant/vehicle combination. Examples are the length of a trip, and the number of passengers during a trip.

Finally, there are variables that vary dynamically over the time intervals of a trip. Many of the behavioural parameters are of this type.

A few remarks about the above categories are in place. First, the division into the categories may not seem clear-cut. For example, a driver can use a cooling fan during the whole of a trip, or only during a part of the trip. We suggest to group the variables according to the hierarchical level at which they could be varied.

Next, a variable from a lower level in the hierarchy could be aggregated into a higher level by calculating derived variables summarizing the whole of the original level. For example, one could calculate the average and the standard deviation of speed over the time points in a particular trip to simplify the statistical analysis.

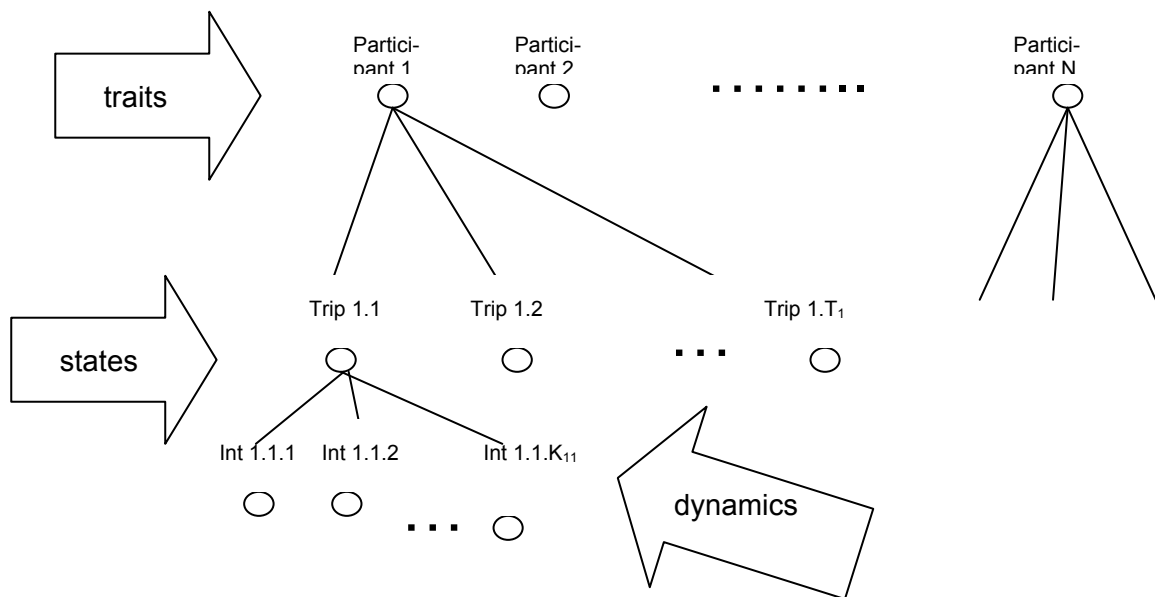


Figure 5.1: Strata of variation in ND studies

At each of the levels in the hierarchy, the variability of the variables is partly systematic and partly random. For any particular variable, the systematic part of the variability is the part that can be explained in a model relating the variable to explanatory variables. The remaining part of the variation is modelled as random variation. So models of variables measured at the participant/vehicle, trip, or time stratum should contain one, two, and three terms modelling the random variation. WP 1 of PROLOGUE poses research questions (RQs) on accidents, near accidents or critical incidents (type A), and research questions on behaviour (type B). There are seven groups of type A and eight groups of type B; see Table 5.1. In the next two subsections, we discuss the statistical issues arising from each type of research questions.

Table 5.1: Research Questions from PROLOGUE, WP 1

Research questions related to accidents, near-accidents and critical incidents / type A
RQ1A How are driver characteristics related to accidents?
RQ2A How is vehicle type related to accidents?
RQ3A How is roadway design related to accidents?
RQ4A How are in-vehicle support systems related to accidents?
RQ5A How is distraction related to accidents?
RQ6A How is fatigue/drowsiness related to accidents?
RQ7A How are weather conditions related to accidents?
Research questions related to behaviour/ type B
RQ1B How are driver characteristics related to driving behaviour?
RQ2B How is vehicle type related to behaviour?
RQ3B How is roadway design related to behaviour?
RQ4B How is in-vehicle safety systems related to driving behaviour?
RQ5B How is distraction related to driving behaviour?
RQ6B How is fatigue/drowsiness related to driving behaviour?
RQ7A How are weather conditions related to driving behaviour?
RQ8B Are drivers conducting environmentally friendly driving (eco-driving)?

5.2.5 Research questions of type A: explaining events

In this subsection, we discuss statistical models aiming at explaining the events as a function of variables characterizing traits, states, or dynamics (see Fig. 5.1). Before addressing the research questions, we would like to emphasize that ‘explaining events’ in an observational study means ‘establishing a strong correlation with events’. A strong correlation, such as the one between the number of storks in a country and the birth rate in the human population, does not in itself establish cause-effect conclusions. For these conclusions, we would need additional subject-matter reasoning, or statistically designed experiments focussed on administering or withholding some treatment of the subjects. This is not to be taken as disparaging observational studies. Indeed, many explanatory variables are observational. It even seems unethical to administer treatment to increase, say, drowsiness, in order to study its effect on events. We conclude that ‘establishing a strong correlation with events’ is the best we can do in many situations.

Research Questions 1A/2A Relation of driver characteristics and vehicle type to events

Examples of research questions in group 1A and 2A include ‘how are driver characteristics related to near-crashes?’ and ‘investigation of interaction between light vehicles and heavy vehicles resulting in near-accidents’.

Characteristics of driver and vehicle are measured at the upper hierarchical level. These are to be connected to events at the bottom level. We suggest addressing the dependency on driver characteristics jointly with vehicle properties, because there is a fixed linkage between driver and vehicle. Hence driver characteristics may be correlated with vehicle characteristics. For this reason, we would want to correct the effect of one variable for effects of others. For example, the age of the driver might be correlated with vehicle type. If we analyse the effects jointly, we can obtain the effect of age corrected for vehicle type and the effect of vehicle type corrected for age. If we would conduct two separate analyses, we would mix up both effects.

There are two ways to connect driver/vehicle characteristics to events.

First, we can model the total number of events for each of the participants as a function of the driver characteristics, the vehicle characteristics, and the total mileage for the driver during the whole of the study. We would have as many data points as there are participants. The dependent variable is the number of accidents. We suggest employing a generalized linear model (GLM; McCullagh and Nelder, 1989, Dobson and Barnett, 2008) to connect the number of events with the driver/vehicle characteristics. The components of such a model are

$$E(Y_i) = \mu_i,$$

$$\text{Var}(Y_i) = \Phi\mu_i$$

$$g(\mu_i) = x_i' \beta$$

with Y the number of events, $E(\cdot)$ the expected value operator, and e a normally distributed random variable. The index i bears on vehicle/driver combination. Further, x_i is a vector containing the explanatory variables, and β is the vector of true effects of these variables. The explanatory variables are linked to the expected number of events through the link function $g(\cdot)$. The usual link function to model counts is the (natural) logarithm.

The parameter Φ in the above GLM is called the dispersion parameter. For Poisson distributed variables, $\Phi = 1$. A value of $\Phi > 1$ models the clustering of events. In our case, this could be needed because some subjects are more accident-prone than others. An alternative probability distribution that accounts for clustering is the negative binomial distribution. This distribution is slightly more complicated to fit, however.

When working with raw counts, we correct for the total mileage of a participant by including this variable in the model. Instead, one could consider working with counts divided by total mileage. This is not recommended, because it assumes a simple linear dependence that cannot be tested afterwards. By explicitly including the mileage one can test whether some curvilinear trend remains after correcting for linear trend.

GLMs are implemented in most major statistical packages such as GenStat and SAS. The model results in a list of fitted coefficients of the explanatory variables, together with their standard errors. For example, the effect of age on the log-number of accidents might be -0.034 ± 0.012 . This would imply that an increase in age with 10 years would decrease the number of accidents with a factor of $\exp(10 \times 0.034) = 1.4$.

The second way of connecting driver/vehicle characteristics to accidents is through modelling of individual trips. For each of the trips, the dependent variable is an indicator variable for the occurrence of an accident. There are as many data points as there are trips. There may be explanatory variables at the trip level as well as at the participant level. We recommend modelling the probability π of an event using a so-called gener-

alized mixed model (GLMM; Breslow and Clayton, 1993). It is an extension of a GLM in a specific way. Both a GLM and a GLMM have fixed effects for the contributions of the explanatory variables, such as age, sex, education and the length of a trip. The random effects model variation that cannot be captured by the fixed effects. A GLM has a single random effect that operates on each observation independently. In our example, we could use a GLM to model the trips of a single participant/vehicle combination. A GLMM has an additional random effect that operates on all the data from one vehicle/participant. It is needed to model variation between the vehicle/participant combinations that cannot be explained by age, sex, etc. The effect is usually assumed to have a normal distribution. Thus, the components of the model are

$$Y_{ij} \sim \text{binomial}(\pi_{ij}, 1),$$

$$E(Y_{ij}) = \pi_{ij},$$

$$\text{Var}(Y_{ij}) = \pi_{ij}(1 - \pi_{ij})$$

$$g(\pi_{ij}) = x'_{ij} \beta + e_i,$$

with Y the event indicator, $E(\cdot)$ the expected value operator, and e a normally distributed random variable. The indices i and j bear on vehicle/driver combination, and trip. Further, x_{ij} is a vector containing the explanatory variables, and β is the vector of true effects of these variables.

The usual link function to model probabilities is the logit, or $\log[\pi/(1 - \pi)]$. This quantity is also known as the log-odds ratio; it is also used in logistic regression analysis. Klauer et al. (2006a) analyze the odds ratio to investigate the risk due to driver inattention. The back-transformation $\exp(\text{logit}(\pi)) / [1 + \exp(\text{logit}(\pi))]$ always results in an estimated value of π between 0 and 1, as appropriate for probabilities.

The length of the trip and the number of interactions should always be included in the model. At the participant / vehicle level one can introduce explanatory variables to answer the research questions proper. In addition, one could also introduce trip characteristics to determine the joint effect of a participant characteristic and a trip parameter. For example, there might be a synergistic effect of age and length of a trip to the probability of an accident.

GLMMs are implemented in most major statistical packages such as GenStat, SPSS, and SAS; see Molenberghs and Verbeke (2005) for an application with SAS. In addition to a list of fitted coefficients of the explanatory variables and their standard errors, the model gives the random variation due to differences in participants and vehicles not accounted for by the explanatory variables in the model. Typical outcomes of the analysis would state that the log-odds of having a near crash changes with 0.017 per additional life year of the participant. A 10 years' age difference could then result in a change of event rate from 1:7890 to 1:9322.

The two ways of connecting driver/vehicle characteristics with events differ in complexity and detail. Modelling the total number of accidents for the participants has the advantage of being conceptually simple. Modelling on the trip level could add more detail but it is also more complex.

Research Questions 3A Roadway design

This group of research questions include investigation of events in intersections and investigation of lane changing related to events. We can approach the group of research questions from two different perspectives. First, one can choose fixed stretches of roadway that differ in design parameters and observe the number of events over a given period of time. Second, one can include design parameters as covariates in the analysis of events of the participants included in the observational study. So there

could be a roadway perspective as well as a participant perspective when approaching the research question.

As regards the roadway perspective, there should be a variety of roadway stretches that differ in important design parameters (see Section 4.3.2). One could choose the stretches according to some factorial design to optimize the power of detecting effects of the design parameters. The data consist of as many records as there are road stretches. The number of events should somehow be corrected for the number of passages of the particular road stretch under observation. This is best done by including the number of passages, or a linearly related number, as a covariate. As in the previous section, we would require a GLM using a Poisson distribution for the number of events and a convenient link function. Note that the question of selecting the road stretches is closely related to selecting participants to address RQ1A and RQ2A.

Viewed from a participant's perspective, RQ3A might be addressed by studying the events and the related roadway design parameters, or by studying the probability of an event in time interval t given the conditions of one or more previous time intervals. We refer to the next section for the probability approach. We now discuss the pure event-based approach.

The event-based approach supposes one record per event. The roadway design has to be captured in clearly defined parameters, such as the number of lanes. The research question translates to the hypothesis of an equal distribution of events over the various categories defined by the roadway design. It is essential to correct for the total number of passes for the roadway section, because a very busy section will naturally have a higher probability of some event taking place.

The above data can conveniently be analysed with frequency table methodology based on log-linear models. The null hypothesis is that there is an equal distribution of events for all of the categories defined by the design parameters. This can be tested by including each of the design parameters in turn in the model. Influential parameters will result in a statistically significant test statistic. For example, suppose that we have to do with roundabouts with 4, 5, and 6 turns. Suppose further that there are 90 crashes on these roundabouts. Then we could test the hypothesis that these are distributed equally over the three types of roundabouts. A further refinement would be the inclusion of the number of passes through the roundabouts, and whether they are for motorized traffic only or not.

One can also address interactions between design parameters by including appropriate terms in the model. We refer to Agresti (2007) for a general account on frequency table methodology.

RQ4A-RQ7A In-vehicle support systems, distraction, fatigue/drowsiness, weather conditions.

We now turn to studying the probability of an event in time interval t given the conditions of one or more previous time intervals. This approach is recommended to study RQ4A-RQ7A, because the research questions bear on conditions that can change within a trip.

Suppose that each of the many trips in the ND study is divided in time intervals. One typical interval could be denoted with T_{ijk} . That is, the interval bears on the k th time interval of the j th trip of participant i . We denote with S_{ijk} the 'state' of this interval. A 0 corresponds with 'no event', and a 1 corresponds with an event. We want to ascertain whether the probability $P(S)$ of state S is related to explanatory variables in the previous time intervals. However, we do not want to include all the previous time intervals, for two reasons. First, it is not likely that an event at interval k depends on variables more than a few time intervals in the past. Second, if we would include explanatory

variables dating back many time intervals, we would end up with extremely sparse data. Estimating probabilities from these data is risky, to say the least.

To restrict attention to the explanatory variables on just, say, one previous interval, we adopt the so-called Markov assumption:

$$P(S_k | S_{k-1}, S_{k-2}, \dots, S_1) = P(S_k | S_{k-1}).$$

(Taylor and Karlin, 1994). Here, we suppress the indices i and j for the moment.

The formula implies that the state on interval k given the state on all previous intervals of the trip only depends on the state of the previous interval. We are particularly interested in the state on interval k given that there was no event in these previous intervals. For this purpose, we split a trip of K intervals, say, into $K-1$ records. The record for interval k contains the state of that interval and the explanatory variables of interval $k-1$. We do not use data of a trip after some event has taken place. In so doing, each trip either results in $K-1$ records with $S=0$, or with 1 record having $S=1$ and $K-2$ records having $S=0$.

A Markov chain proper looks back only to the previous time interval. There may well be cases for which this is insufficient. For example, continuing with one activity from many time intervals back could induce fatigue. We could handle this either by looking further back, or by extending the one time interval for looking back and calculating aggregate summaries for that interval.

The records of each trip all have the state of the interval as a 0/1 response variable. Explanatory variables will vary according to the research question to be addressed. For example, to assess the effect of weather condition, one could include time of the day, road surface parameters and parameters that model the weather condition. We suggest using logistic regression (Agresti, 2007) or a GLMM including a random effect for trip; see Research Questions 1A/2A above. Both types of analysis quantify the probability of an event as a function of explanatory variables in previous time intervals.

If we include trip or participant as random variables in a GLMM, it is possible to include driver characteristics and interactions of these characteristics with explanatory variables at the trip or interval level. For example, a reviewer suggested a possible interaction between the experience of the driver, which is constant over a trip, and specific properties of the road, which vary over the trip. Indeed, experienced drivers could handle complicated bends more easily than inexperienced drivers. So a difference in event rate between complicated bends and simple ones may depend on the experience of the driver. This is modelled with interaction terms.

5.2.6 Research questions of type B: Explaining behaviour

Driving behaviour is operationalized by recording several variables over time. Each trip in an ND study gives rise to its own time series of behavioural variables. In much the same way as with the research questions on events, we can use appropriate summaries of the behaviour parameters and relate these either to variables at the participant/vehicle level, at the trip level or at the time interval level.

Research Questions 1B/2B Driver characteristics / vehicle type

There are two ways to connect driver/vehicle characteristics to driving behaviour.

First, we can define an aggregate variable over each of the participants and model this number as a function of explanatory variables characterizing the participant and the vehicle. For example, we can define the total number of lane changes as an aggregate

that characterizes dynamic behaviour. We can model this variable with a GLM in much the same way as the total number of events for the participants (see RQ1A/RQ2A). It is crucial to normalize this variable to account for differences between the total mileages of the participants.

As a second example of an aggregate variable, we can derive various summaries from each participant's speed profiles such as the standard deviation of the speed. These summaries will be continuous, as opposed to counts. As before, we employ a GLM (McCullagh and Nelder, 1989; Dobson and Barnett, 2008) to connect the variable with the driver/vehicle characteristics. The random variation can be taken to be distributed as gamma. That is

$$\text{Var}(Y) = \Phi\mu^2,$$

where $\text{Var}(Y)$ is the variance of Y , Y is the continuous variable, μ is its mean, and Φ is the dispersion parameter. We can calculate the dispersion parameter from the data.

A gamma distribution is convenient when the absolute standard deviation increases with the mean value of the variable. It has the attractive property that the relative standard deviation is a constant. So one could state that the standard deviation of some quantity is, say, 7.4 %.

A useful link function for non-negative data such as standard deviations or variances is the (natural) logarithm, because it maps a value of $\mu > 0$ on the real line. Reversely, any model prediction on the log-scale is transformed back to a strictly positive value. So the expected standard deviation is always positive.

The second way of connecting driver/vehicle characteristics to driving behaviour is through modelling of individual trips. Here, the dependent variable is summarized over a trip rather than over all the trips of a participant. There are as many data points as there are trips. There may be explanatory variables at the trip level as well as at the participant level. Variables for the trip level could include the length of the trip and trip-related variables such as whether or not there were passengers during the trip. We recommend a GLMM to model the random variation between the vehicle/participant combinations. The components of the model are:

$$Y \sim \text{gamma}(\mu, \Phi),$$

$$E(Y) = \mu,$$

$$\log \mu = x' \beta + e,$$

with Y the behaviour parameter, and e a normally distributed random variable.

At the participant / vehicle level one can introduce explanatory variables to answer the research questions proper. In addition, one could also introduce trip characteristics to determine the joint effect of a participant characteristic and a trip parameter. For example, there might be a synergistic effect of driver traits and driver states on driving behaviour.

We refer to the discussion of Research Questions 1A/2A for further remarks on GLMM.

Research Questions 3B-7B Roadway design, in-vehicle support systems, distraction, fatigue/drowsiness, weather conditions

The research questions 3B up to 7B all relate to dynamic explanatory variables, because these variables can change within a trip. (We assume that Question 3B cannot be addressed from the roadway perspective). This poses the problem of how exactly to define the explanatory variables and the dependent variables.

In our discussion of the corresponding research questions for events, we suggested defining time intervals and relating the occurrence of an event at interval k to variables

at a few previous intervals. For behaviour response parameters we recommend using concurrent time intervals. That is, we relate behaviour occurring at time interval k to roadway design, support systems, distraction, fatigue or weather conditions prevalent at the same interval. So the intervals should be sufficiently large to define a sensible behavioural variable, and sufficiently small to capture the change in explanatory variables.

Each trip will contribute multiple records, depending on the length of the interval. For this reason, there will be correlation between the observations (behaviour variables) within a trip. Results of any two intervals can be correlated because they are from the same participant, or because they are from the same trip, or because they are taken at near time points. So a statistical model may have to contain random variables for participants, and for trips within a participant, and may also have to provide for a serial correlation.

If no serial correlation is needed to adequately model the data, we arrive at a GLMM with two random components. If there is indeed a serial correlation to be catered for, we may have a problem, because the fitting algorithm might not be able to handle so many complexities at once. In that case, we recommend considering a transformation of the behavioural variables such that it is approximately normally distributed. All the random components in the model are then normal, and the model reduces to a mixed model (West et al., 2006), for which standard software is available.

This section discusses analyzing the effects of very different explanatory variables on behaviour. It will depend on the particular behavioural variable what set of explanatory variables should be considered. In general, we recommend including them all in the analysis and discarding explanatory variables only on the basis of a statistical test.

Research Questions 8B Environmentally friendly driving (eco-driving)

The study of eco-driving raises the same kind of issues as those for questions 1B/2B; see the discussion of these questions.

5.2.7 Reliability and validity of study results

An ND study, like any large-scale quantitative study, is worth the effort only if the results are both reliable and valid. This statement will sound trivial in many ears, because of the every-day usage of these words. However, in social studies the adjectives 'reliable' and 'valid', together with their nouns of 'reliability' and 'validity' stand for concepts with a technical meaning beyond the every-day usage. In this subsection, we discuss both concepts.

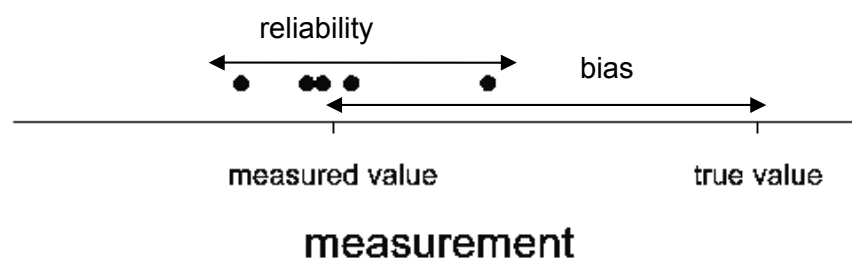
In the previous subsections, we discussed statistical models to explain events or behaviour using explanatory variables. The effects of these variables are expressed in a formula that quantifies the strength of a relationship by a list of fitted coefficients of the explanatory variables, together with their standard errors. We previously used the effect of age on the log-number of accidents as an example. A point estimate of this effect might be -0.034 . Recall that this implies an estimated decrease in the number of accidents with a factor of $\exp(10 \times 0.034) = 1.4$ with an increase in age with 10 years.

The *reliability* of the age effect is a measure of how consistent the effect measure is. Suppose that the mean value of the effect measures is given by the tick mark labelled 'measured value' in next page's figure. We show five independent measurements of the effect as fat black dots. A reliable study has a small spread between the dots. An appropriate measure of the reliability is the standard error of the effect. We postulated a standard error of 0.012. This implies that 95% of the realizations of an effect measure can vary between the 'measured value' $- 2 \times 0.012$ and the 'measured value' $+ 2 \times 0.012$.

The five realizations of the effect measure are based on five complete ND studies. In practice, of course, we carry out a single study, and the ‘measured value’ is unknown. Nevertheless, it is possible to calculate an estimate of the standard error. With this estimate we can calculate whether the effect measure is compatible with some hypothesized value of the ‘measured value’. The smaller the standard error, the nearer we are to the ‘measured value’.

The *validity* of the age effect is a bit more complicated. Suppose that the drivers selected for the study, for some reason were recruited from people that are very interested in science. Suppose further that these people are particularly healthy, and that they age slowly. Finally, suppose that we want to tell something about the age effect in the general population. In that case, the ‘measured value’ of the age effect deviates from the ‘true value’ we are really after. This cannot be repaired by doing several repeats of the ND study. The conclusions in our study are not valid for the general population, and the results are said to be biased.

We discussed validity and reliability in the context of a study with quantitative statistical models. They are more often discussed in the context of psychological testing (see, e.g., Kaplan and Saccuzzo, 2009). In psychological tests, reliability has much the same meaning as before. However, validity now relates to the extent that psychological concepts can be derived from a battery of test questions, and sampling from the wrong population seems to be less of an issue here.



6 Organizational issues

Naturalistic driving observation is a relatively new method for studying road safety issues, a method by which one can objectively observe various driver- and accident related behaviour. Typically, participants get their own vehicles equipped with some sort of data logging device that can record various driving behaviours such as speed, braking, lane keeping/variations, acceleration, deceleration etc., as well as one or more video cameras. In this way normal drivers are observed in their normal driving context while driving their own vehicles. Optimally, this allows for observation of the driver, vehicle, road and traffic environments and interaction between these factors.

This chapter provides an overview of the organizational issues that play a role and need to be addressed within a Naturalistic Driving (ND) study. Organizational issues are split into planning and stakeholders. Per topic there are three paragraphs. First there is a paragraph with relevant issues: findings from the literature. Then follows the subject experience, where relevant information from projects (past and current) on organizational issues is collected. After that, the information from literature and projects is compared: what are the differences and similarities. This is written down in the last paragraph which contains the conclusions.

6.1 Organizational issues – planning

6.1.1 Relevant issues

This overview of relevant issues was drawn primarily from the recently-completed FESTA project (FESTA Handbook v.2). This project gathered the state-of-the art knowledge and practice on setting up and running a Field Operational Test (FOT), and documented this information in a handbook and a set of appendices.

Chapter 2 of the FESTA handbook is on organizational issues (planning and running a FOT), and annex B of the handbook contains a complete FOT implementation plan including a checklist. It offers very structured help in conducting a pilot, and is therefore very relevant for a ND study.

For a Naturalistic Driving (ND) study to proceed smoothly, a plan of action must be developed which documents the scientific, technical, administrative and procedural activities and tasks that are needed to successfully complete it. It is important to know that costs overrun usually affect the number of participants, which in turn requires longer periods of data collection, which in turn affects the likelihood of completing the study in time.

Here the critical activities and tasks which are necessary to run a successful ND study are documented drawing on lessons learned from previous FOTs conducted in Europe, the United States, Japan, Australia and elsewhere and based on the FOT Implementation Plan from FESTA.

Table 6.1 lists the 22 activities identified in the FOT Implementation Plan. This table also applies to a ND study.

Generic guide to scheduling the 22 Activities described in the FOTIP, Annex B of FESTA Handbook

Table 6.1: Generic guide to scheduling the 22 Activities described in the FOTIP, Annex B of FESTA Handbook

		Set Up/Design	Preparation	Data Collection	Completion
1	Convene teams and people	█			
2	Define aims, objectives, research questions & hypotheses		█		
3	Develop project management Plan		█		
4	Implement procedures for stakeholders communication		█		
5	Design the study		█		
6	Identify and resolve legal and ethical issues		█	█	
7	Select and obtain Vehicles		█		
8	Select and obtain systems and functions to be evaluated	█	█		
9	Select and obtain data collection and transfer systems		█		
10	Select and obtain support systems		█		
11	Equip vehicles with technologies		█		
12	Implement driver feedback and reporting systems		█	█	
13	Select / implement relational database for storing data		█	█	
14	Test all systems to be used according to specifications			█	
15	Develop recruitment strategy and materials		█		
16	Develop driver training and briefing materials		█		
17	Pilot Test equipment, methods and procedures			█	
18	Run the FOT			█	█
19	Analyse the data				█
20	Write minutes and reports	█	█	█	█
21	Disseminate the findings				█
22	Decommission the study				█

6.1.2 Experience

100 Car study

In the 100 Car ND study Phase I and Phase II a planning was made for the different activities. The study was split into three phases: first the planning phase, then the conducting of the test and last the ‘deployment’ phase (recommendations for a large-scale field study). In Table 6.2 the organization of the study and its objectives can be found. The six objectives are accomplished through thirty tasks. The first sixteen tasks (phase I) are described in detail in the literature.

Table 6.2: Organization of the 100 Car Naturalistic Driving Study

PHASES	TASKS	OBJECTIVES
PHASE I: CONDUCT TEST PLANNING ACTIVITIES	1-16	<ol style="list-style-type: none"> 1. Establish Intelligent Vehicle Initiative (IVI) data needs to support rear-end and lane-change crash countermeasure development and benefits estimation. 2. Develop a methodology for data gathering, reduction, archiving, and analysis for field study data collection. 3. Develop and deploy a data collection system.
PHASE II: CONDUCT IVI FIELD TEST	17-27	<ol style="list-style-type: none"> 4. Conduct a naturalistic driving study. 5. Develop a philosophical and mathematical relationship between rear-end crashes and rear-end near-crashes as well as lane-change crashes and lane-change near-crashes.
PHASE III: PREPARATION FOR THE LARGE-SCALE FIELD DATA COLLECTION EFFORT	28-30	<ol style="list-style-type: none"> 6. Make recommendations to IVI for large-scale field study partnerships.

According to the 100 Car ND study the following needs to be done in order to organize a ND study (phase I):

- 1 Specify the details of the pre-crash and near-crash data to be gathered during the data collection phase;
- 2 Specify the number of cars to be instrumented, the number of camera views, the number of vehicle makes and models to be used, and the rate at which data was to be collected;
- 3 Determine the number of sites from which data can be collected, the rear-end crash frequency at various geographic locations, and the optimal location of the data collection site from the perspective of project resources;
- 4 Determine crash sampling requirements – how much data is needed;
- 5 Determine driver / vehicle demographic requirements – what are the ‘ideal’ driver characteristics that contribute to rear-end crashes;
- 6 Determine near-crash statistical power requirements (the 100 Car ND study did this by reviewing four previous research studies with instrumented vehicles in a natural driving environment);
- 7 Determine research design parameters, sampling rates and formats concept;
- 8 Determine vehicle types – critical factors include vehicle demographics, vehicle location and data collection system installation issues;
- 9 Develop participant recruiting specification – age, gender, vehicle types driven, number of miles driven per year, location of permanent residence or place of work;
- 10 Develop a test data collection plan – a synthesis of the previous points (Task 1 through 9) with revisions after comments of the contract sponsor;
- 11 Develop a test reduction plan (using an incident / near-crash data reduction method);
- 12 Develop a data analysis plan (including hardware, back-up and archiving aspects) based on the research questions;
- 13 Develop data collection system requirements. Data system requirements were categorized into four major areas: schedule requirements, general design requirements, performance requirements and test vehicle profile. The data collection system requirements are input for the hardware / software design specification;

- 14 Review and test technology and sensor alternatives;
- 15 Review and test trigger criteria methods – events in the dataset are located post hoc with editable triggers, which results in a comprehensive database;
- 16 Trade study analysis of hardware / software alternatives.

SHRP2

The Strategic Highway Research Program SHRP2 Safety Research Program aims to greatly increase the knowledge of driver behaviour. A large-scale ND study with over 3,000 participants has been initiated. Few documents about the planning of the FOT are available as the ND study is active.

6.2 Organizational issues – stakeholders

6.2.1 Relevant issues

FESTA

In the past, the impact assessment of FOTs focused on a narrow set of impacts of interest. Few looked at the stakeholder or supplier perspectives; some measured benefits but not (social) costs; very few started out with an impact table and formally identified what the expected “main effects” of the systems investigated would be; and some did not carry out a socio-economic impact assessment.

The recommendation is that FOTs and ND studies should be designed to be as complete as possible, both in terms of impacts and stakeholder views. This makes the study better and also helps the stakeholders. By addressing the research questions, FOTs and ND studies promise to furnish the major stakeholders (customers, public authorities, original equipment manufacturers (OEMs), suppliers, and the scientific community) with valuable information able to improve their policy making, services, user interface and market strategies.

For a ND study, a Stakeholder Analysis is recommended. During this analysis, the needs of the different stakeholders are identified and merged into a common requirements description. Stakeholders are those whose interests are affected by the issue or those whose activities strongly affect the issue, those who possess information, resources and expertise needed for strategy formulation and implementation, and those who control relevant implementations or instruments. Examples are the European Commission, governments and ministries of the participating countries, vehicle manufacturers, suppliers of systems used, research institutes, etc.

It is recommended to evaluate the stakeholders’ needs by means of questionnaires, workshops or well documented interviews of stakeholders’ representatives. It is also quite important to describe the selection process sufficiently to prevent from misjudgment.

Stakeholder Analysis

In the FOT-Net project, a practical approach for performing a stakeholder analysis was developed. This approach is applicable for various types of study, for example a ND study. The major elements of a stakeholder analysis are the following:

- Make a list of all stakeholders and classify them. Every stakeholder is different in background, goal, importance, and influence. Classifying stakeholders helps getting the ‘bigger picture’;

- List the interests of every stakeholder. With this information, the stakeholders that have conflicting interests or concerns, and the stakeholders that have interests that conflict with your (project) goal can be indicated;
- List the influence of every stakeholder;
- Describe the relations between the stakeholders, for example in a picture;
- Give stakeholders their role / place in or around the project or decide in what other way to involve them. Use information from the previous points for this.

6.2.2 Experience

SHRP 2

The SHRP 1 project, has formulated a lot of recommendations for the SHRP 2 project. The SHRP 2 project has done something with these recommendations and has also produced a couple of new recommendations. This paragraph contains two sections:

- Recommendations
- Stakeholders in SHRP 2

Recommendations

According to the SHRP 1 project, many other groups must be brought to the table:

- Local, regional, and state governments;
- Manufacturers and suppliers;
- The construction industry;
- Engineers and designers.

The roles of these groups will vary—from central to ancillary to no role at all—for different types of products.

Communicating early and often with the wide array of stakeholders mentioned above should help in identifying those who have concerns about the impact of a new way of doing things. These concerns need to be addressed early, openly and clearly. Sometimes there is no way to avoid a negative impact on some party, but often a mutual agreement can be reached that the innovation is beneficial for all parties or at least is not necessarily a threat to anyone.

Frequent updates should be provided to all interested parties. Even supportive stakeholders will need assistance in organizing themselves to facilitate implementation. The implementation of research results can take many years and a significant investment of resources, and it must continuously be accompanied by selling the benefits of the implementation. This is particularly true of research products or programs that are designed to last a long time.

Users and other stakeholders that are in a position to influence their acceptance should be involved in planning and carrying out implementation activities. Strong partnerships with and among stakeholders build trust and encourage implementation champions.

Active involvement of stakeholders also builds trust—a fundamental element of successful implementation in the highway community. The risk averse culture of the community, derived from its public responsibility and institutional incentives and disincentives accrued over the years, leads it to place much weight on trustworthy experts and on demonstrated experience before trying something new.

Involvement of stakeholders also includes coordination with other related programs and external stakeholder groups. Promotion of collaboration to expedite implementation, leverage resources, and increase the effectiveness of your products. Coordination includes working with standards-setting organizations, stewards of professional and technical manuals and guidebooks, and providers of technical training and certification, as appropriate.

Implementation cannot be reduced to communication—it is not enough to market innovations or to publish reports—but is a critical component. It is essential not only to provide information and answer questions but also to listen: to discern what users want, what they need but cannot articulate, where resistance may lie, and what positive lessons it can yield.

Communication must be a two-way process: the principal implementation agent must seek information from potential users about incentives and challenges they face and work cooperatively to leverage the incentives and overcome the challenges. The temptation to hear only from supportive stakeholders must be avoided; those who resist an innovation often have good reason for doing so, and much can be learned from them to make an innovation more attractive to more stakeholders. Communication must also be maintained throughout the implementation process. In addition to hearing about innovations when they first become available, stakeholders will want periodic updates on the progress of implementation; success stories; challenges overcome; and, most important, benefits achieved. Many communication mechanisms should be used to describe research results and products, to report on implementation activities, and to share information. These mechanisms can include extensive electronic information in e-newsletters and searchable websites, face-to-face interaction in workshops and focus groups, webinars, and wikis (websites that allow users to add and edit content).

Stakeholders in SHRP 2

Plans for and tentative outcomes of the program's early research efforts were consulted with a variety of stakeholders, to project the ultimate outcomes of the program, its potential users, and the incentives for and impediments to implementation that may be encountered.

The focus areas of SHRP2 were developed through almost 3 years of study and consultation with a broad array of stakeholders to ensure that the most critical needs would be addressed. A characteristic of SHRP 2 is that it is focused more on changing the way highway agencies do business than on producing a number of technology products.

Changing institutions and processes is risky, especially in the public sector. SHRP 2 produces methods and guidance, as well as technologies, designed to help agencies make the changes necessary to better serve their customers while managing the risk involved with institutional change.

Stakeholders will modify the ways they think and act, and possibly the ways their organizations are structured. The highway transportation community is large, complex, and generally risk averse. Small innovations are sometimes easier to implement than those that require a paradigm shift. Time and dedicated resources will be required to create an environment in which the highway community will embrace innovations that promise long-term benefits and may require a substantial change in organizational behaviour and business practices. SHRP 2 Renewal products must be applied systematically over an extended period of time for such change to take place.

Risk management: risk is inherent in trying anything new; some technological and methodological innovations can cause risk to be reallocated among stakeholders in a project. For instance, some innovative procurement procedures (such as warranties

and use of performance specifications) shift control over and therefore responsibility for, product quality from the transportation agency to the contractor. Unwillingness or inability to accept increased risk can be an impediment to implementing an innovative approach. Because of their responsibility to the public and the incentive structure they face, highway agencies tend to be risk averse. At both the individual and agency levels, there is little reward for success in innovation, and there are potentially huge penalties for failure.

6.3 Recommendations

- Before the ND study starts, make a plan of action which documents the scientific, technical, administrative and procedural activities and tasks that need to be done.
- Before the ND study starts, perform a stakeholder analysis to identify the needs, interests and influence of every stakeholder. With this information, the stakeholders that have conflicting interests or concerns (with each other or with the project goal) can be identified.
- Communicate early and often with a wide array of stakeholders and keep them updated.
- If necessary or useful, give stakeholders their role / place in or around the project.

7 Legal and Ethical issues

An essential component of Naturalistic Driving (ND) studies is the observation of various driver- and accident-related behaviours. Observation includes objectively and unobtrusively observing normal drivers in their normal driving context while driving their (own) vehicles. Typically, participants get their (own) vehicles equipped with some sort of data logging device that can record various driving behaviours such as speed, braking, lane keeping/variations, acceleration, deceleration etc., as well as one or more video cameras. These observation characteristics have implications for addressing legal and ethical issues within a ND Study.

This chapter begins with an overview of the legal and ethical issues that need to be addressed in a ND study. The second section shows how two ND studies and a Field Operational Test (FOT) addressed legal and ethical issues. This chapter concludes with a comparison between the relevant issues identified and what was done in practice.

7.1 Relevant issues

This overview of relevant issues was drawn primarily from the recently-completed FESTA project. This project gathered the state-of-the art knowledge and practice on the setting up of and running a Field Operational Test (FOT). This project documented this information in a handbook and a set of appendices. More detailed information on each of the legal aspects described below can be found in the appendix of the FESTA handbook and in FESTA Deliverable 6.4.

Carrying out a ND study gives rise to a considerable number of legal and ethical issues — obtaining the necessary permissions, ensuring that the vehicles are safe to operate on the public highway, going through any required ethical and human subject review procedures, obtaining participants' consent, complying with data protection laws, insuring the vehicles, insuring the project workers for indemnity and so on. It is not possible to provide a comprehensive guide to all the legal issues that can arise in a particular study, as these may be very dependent on the study design adopted. It is therefore imperative that the project obtain legal advice at an early stage. It should be noted that the regulations and laws vary from country to country and that even where there are European laws and regulations — for example on data protection and privacy — the interpretation of these may vary between countries. Thus projects carrying out a ND study in more than one country or carrying out studies that potentially involve cross-border traffic may need to consider the legal implications in all relevant countries. Another vital aspect is that projects fully consider health and safety aspects. It should be noted that not carrying out a prior risk assessment and therefore not giving proper consideration to the safety risks that may result from a ND study can expose an organisation to criminal prosecution, e.g. for corporate manslaughter in the event that an unforeseen disaster occurs.

Participant recruitment

In recruitment it is essential to ensure that participants have legal entitlement to drive the vehicles in question and are eligible for insurance. It may be wise to have insurance coverage for the fleet as a whole. If the participants are to drive their own vehicles or vehicles that belong to a fleet not under the control of the handling organisation, then insurance coverage needs to be confirmed. Coverage when travelling to other countries may be relevant. In some countries, it may be a requirement for the participants to undergo a medical examination to prove their capability to take part. In any case, it

would probably be sensible to ascertain if they have any medical conditions that might affect their ability to participate.

Participant agreement

There is a need to formalize the arrangement between the organisations responsible for the relationship with the participants and those participants themselves. The participants need to be informed in advance about the purpose of the study, the risks they may incur, the costs that are covered and not covered (and so have to be borne by them), whom to contact in case of breakdown, etc. It is not necessarily the case that the relationship with the participants will be set in the form of a legal contract; alternatively it may take the form of a letter of agreement. A lawyer can provide advice on this and should definitely be consulted. The agreement or contract may need to cover the potential liabilities and which party is responsible. One liability to consider is what happens in the event that a participant commits a traffic offence and/or incurs a traffic penalty (speeding ticket, parking ticket, etc.). Another liability is who is responsible for minor damage to the vehicle and payment of any insurance excess.

The issue of who is allowed to drive, e.g. other household members, and under what circumstances also needs to be considered. Only the participants will have been properly informed about their responsibilities. There is no way to ensure that any third parties are properly briefed.

Data protection

Data protection is stipulated by an EU directive of 1995 and is enshrined within the national laws of the various member states. These national laws may state specific requirements. There is no doubt that a ND study will give rise to data protection and privacy issues. No disclosure of the data, in such a way as to give rise to identification of the persons involved, can normally take place without prior consent. This can cause problems, even when the participants have been informed of in-vehicle video recording. If that video is subsequently passed on to a third party and the participant can be recognized from that video, there may be a problem.

Video recording (and also audio recording) can give rise to other problems. Passengers will not normally have given prior consent to being recorded, so it is questionable whether it is appropriate to have in-vehicle cameras with coverage of the passenger seats.

The data server must be protected from intrusion, and normally any personal ID information should be kept completely separate from the main database and stored with additional protection such as encryption. It has to be recognized that, even when data has been anonymized, it may be possible to deduce who has participated, e.g. from GIS data in the database.

Decide early in the project how to manage post-project data. Issues to consider are: What happens to data when the project ends? Who will have data usage rights? Who can access it? Who pays for possible storage? In projects with large amounts of stored data (several terabytes), the cost to store and manage data is not insignificant, and all project partners might not have the means to handle it afterwards. Where data is taken off-line, determine what metadata¹ should be kept, and how.

¹ Metadata describes other data. It provides information about a certain item's content. For example, an image may include metadata that describes how large the picture is, the color depth, the image resolution, when the image was created, and other data.

It is recommended to let participants know in an informed consent that their information will only be released when a court demands it (in case of an accident).

Risk assessment

The project needs a comprehensive risk assessment plan and will need to be able to demonstrate subsequently that the identified hazards have been properly managed. Organisations will normally have a safety management process for this.

System safety

It is obviously incumbent on those conducting a ND to ensure that the equipment that they have installed in a vehicle and the modifications that have been made to the vehicle systems do not give rise to any undue hazards. Hazards can arise from radio and electrical interference (where electro-magnetic compatibility tests should be conducted), from reducing vehicle crashworthiness (installations on the dashboard, interference with airbag deployment, and so on) and from HMI designs that cause distraction. The potential for failures to arise from modifications to and interaction with in-vehicle systems needs to be handled by means of a formal system safety assessment.

Approval for on-road use

Vehicles are subject to Whole Vehicle Type Approval processes and to Construction and Use regulations. Before it is certain that it is legal to operate a modified vehicle on public roads, a check must be made with the appropriate authorities, who may be the national government or a designated approval agency. Once a vehicle is certified to be legal to operate in one European country, it can normally be driven legally in other countries.

Insurance

Insurance requirements extend beyond the insurance of the vehicles and possibly of the participants. There is also a need for indemnity insurance to cover the ND as a whole. This may be provided by an employing organization's professional indemnity insurance, but it is vital to confirm that the large risks are covered.

Video data collection

Video data collection within the vehicle has been covered in the section about data protection. However, there are some additional points to consider. For example, there may be locations encountered where it is illegal or prohibited to video externally — border crossing, military locations, private premises. The possibility of this happening needs to be considered; external video may give rise to the same data protection issues as internal video. Many countries have regulations on the collection of outdoor video.

Ethical approval

Ethical approval to conduct a ND study may be even more difficult to obtain than legal approval. In many countries and in many organisations there are strict ethical approval and human subject review procedures. These procedures can be very time-consuming, so that time for the process needs to be considered in the project plan. Human rights legislation is also relevant, as is the Helsinki Declaration of 1964 and its subsequent revisions. This declaration enshrines the right of the individual to be informed and pro-

vide prior consent. The individual's protection and rights supersede any interests of scientific progress.

7.2 Experience

Two ND studies provide examples of how legal and ethical issues were addressed. The sections below provide information from each of the studies.

7.2.1 100-car study

The study had two types of consent forms, one for the drivers of private vehicles, the other for drivers of leased vehicles. Both consent forms required the drivers to have a valid drivers' license.

The paragraphs below describe the items covered by the consent forms.

Private vehicle

To participate, drivers must have a valid drivers' license and own a vehicle of which they are the primary driver for the experimental period of one year.

All participants were screened, took a health assessment, answered a sleep hygiene questionnaire, and were assessed using a Dula Dangerous Driving Index (Dula and Ballard, 2003). The screen also asked whether the driver would be the only driver of the vehicle, or whether other drivers may sue the vehicle. All drivers of the vehicle had to sign the consent form. In each instrumented vehicle, there was a camera that records the face, so primary and additional drivers could be recognized (of course this was anonymized to 'driver 1', driver 2', etc.). For additional drivers, no demographic data was available. Most analyses (at least the ones that do consider age, gender or vehicle miles travelled) were performed only for primary drivers.

The consent form for the drivers of private vehicles appears in Appendix II. The consent form has a section called "procedures and subject responsibilities", covering study preparation, in-processing, equipment installation, data collection and downloading, equipment maintenance, what to do in the event of a crash or airbag deployment, equipment de-installation and out-processing. It also addresses (safety) risks to the participant, benefits and the extent of anonymity and confidentiality. The consent form informed participants that no other passengers in the vehicle were recorded by the camera. It also informed the participant that there was audio recording capability in the vehicle, but that it only recorded for one minute after the driver activated the incident push button. The audio microphone was directional and only recorded the voice driver's voice.

The form also provided explicit instructions that should the driver be involved in an accident, the data collection equipment in the vehicle would likely capture the events leading up to the event. Participants were instructed NOT to give the data collection equipment to police officers or any other party.

The consent form also covers compensation, the freedom to withdraw from the study. The form also explicitly states that the research project was approved by the appropriate body that reviews research involving human subjects. Finally, all vehicle drivers were required to sign the consent form.

Leased vehicles

The consent form for the drivers of leased vehicles appears in Appendix I. This form is very similar to that for the drivers of private vehicles, except that the “vehicle return” section replaces the “equipment de-installation” section. It also addresses automobile insurance from a leasing point of view.

Lessons Learned which are relevant for legal and ethical issues

Important lessons were learned with regard to protecting the confidentiality of the drivers in the study. To protect the drivers in the event of a crash, it was deemed important to obtain a Certificate of Confidentiality from the National Institutes of Mental Health (NIMH).

The purpose of this certificate was to prevent the data collected in the study from being subpoenaed so that it could not be used against a subject in court. However, obtaining the certificate imposed a constraint on the study. Specifically, it was an original goal of the study to instrument the vehicles to collect video of the entire cab of the vehicle as well as to collect audio continuously to better understand the effect of passengers on driver distraction. Nonetheless, administrators at NIMH felt that it was important to protect the confidentiality of anyone in the vehicle who could be recorded via video or audio recordings. To have the driver administer and submit informed consent forms (or assent forms for minors) for every person who may get into the vehicle during the course of the year was considered infeasible and inappropriate. Posting a message inside the vehicle telling every person that they were being recorded was thought to have a negative effect on the naturalistic data collection approach with regard to the driver.

Therefore, the choice was made to use camera placement and angles that would only collect data on the driver and to only have audio recording active when the driver activated the incident push button. Obviously, from the perspective of understanding the degree to which passengers are creating a distraction in the vehicle, the data collected are not as complete as initially desired.

7.2.2 SHRP2

The information provided here is based on a telephone interview with Dr. Kenneth Campbell, SHRP2 Safety Chief Program Officer, and documents he provided. In the US, strict rules determine the use of confidential data, such as that collected during a ND study. Each organization involved must submit its proposed work with human subjects to its Institutional Review Board (IRB) in order to obtain permission to carry out the work in the study. As the research is being funded by a US government body, the researchers must be sure to defend the database against requests for release.

In addition, SHRP2 aims to make the data collected during the study available to qualified researchers beyond the SHRP2 study period. Usually, a consent form signed by a participant must list the names and affiliations of the researchers that will carry out the analysis. Because the SHRP2 aims to make this data usable 30 years into the future, and the names of the future researchers are not yet known, this is a new aspect of the consent form. This is especially challenging because the usual way that IRBs operate means that that organization is not well suited to safeguarding data after it has been collected.

7.3 Recommendations

- Obtain legal advice at an early stage of the project. In case of a large project, consider involving a legal expert in the project. It is not possible to foresee all legal issues that can arise in the study. Note that regulations and laws vary from country to country

- Ensure that participants have legal entitlement to drive the vehicles in question and are eligible for insurance.
- Formalize the arrangement between the organizations responsible for the relationship with the participants and those participants themselves. It is not necessary to do this in the form of a legal contract; alternatively a letter of agreement can be used. The participants need to be informed about the purpose of the study, the risks they may incur, the costs that are covered and not covered, whom to contact in case of a breakdown, etc. The agreement or contract needs to cover potential liabilities and which party is responsible.
- Consider the issue of who is allowed to drive the vehicle.
- Make sure the data is protected and ensure privacy. The data server must be protected from intrusion and personal ID information should be kept separate from the database and stored with additional protection such as encryption.
- It is recommended to let participants know in an informed consent that their information will only be released when a court demands it.
- Decide early in the project how to manage post-protect data.
- Make a comprehensive risk assessment plan to be able to demonstrate subsequently that identified hazards have been properly managed.
- Ensure that the equipment that is installed in the vehicle and the modifications that have been made to the vehicle systems do not give rise to any undue hazards.
- Check with the appropriate authorities that it is legal to operate the modified vehicles on public roads.
- Gain ethical approval to conduct the ND study.

References

- Agresti, A. (2007). *An introduction to categorical data analysis* (2nd edition; New York: Wiley).
- Amundsen, F.H. and Hyden, C., ed. (1977). *Proceeding of First Workshop on Traffic Conflicts*, Institute of Transport Economics, Oslo/Lund Institute of Technology, Oslo, Norway
- Bausell, B.R., and Li, Yu-Fang (2002) *Power analysis for experimental research* (Cambridge: Cambridge University Press)
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* 57, 289-300.
- Breslow, N.E., and Clayton, D.G. (1993). Approximate Inference in Generalized Linear Mixed Models. *Journal of the American Statistical Association* 88: 9–25.
- Cohen J. (1988). *Statistical power analysis for the behavioural sciences* (2nd edition; Hillsdale NJ: Lawrence Erlbaum Associates).
- Conner, M., Smith, N., & McMillan, B. (2003). Examining normative pressure in the theory of planned behaviour: Impact of gender and passengers on intentions to break the speed limit. *Current Psychology*, 22(3), 252-263.
- 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., Knippling, R.R. (2006). *The 100-Car Naturalistic Driving Study. Phase II – Results of the 100-Car Field Experiment*. Report No. DOT HS 810 593. National Highway Traffic Safety Administration. Washington D.C, USA.
- Dobson, A.J., and Barnett, A.G. (2008). *An introduction to generalized linear models*. (3d edition; New York: Chapman and Hall/CRC)
- Dula, C.S. and Ballard, M.E. (2003). Development and Evaluation of a Measure of Dangerous, Aggressive, Negative Emotional, and Risky Driving. *Journal of Applied Social Psychology* 33 (2), pp. 263–282.
- Faraway, J.J. (2006) *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models* (Boca Raton: CRC).
- FESTA (2008). Handbook version 2, Deliverable 6.4 of the EC project Field operational support Action (FESTA), grant agreement number 214853.
- FESTA (2008) Deliverable 2.3 (D2.3), Primer on experimental procedures, grant agreement number 214853.
- FESTA (2008) Deliverable 6.4 (D6.3), FOT requirements, legal aspects planning and development, grant agreement number 214853.
- FOT-Net project, <http://www.fot-net.eu/>.
- Hanowski, R.J., Olson, R.L., Hickman, J.S., and Dingus, T.A. (2006). The 100-car naturalistic driving study; a descriptive analysis of light vehicle-heavy vehicle interactions from the light vehicle driver's perspective. Contract No. DTNH22-00-C-07007. Blacksburg, VA: Virginia Tech Transportation Institute.
- Jonah, B.A. (1997). Sensation seeking and risky driving: A review and synthesis of the literature. *Accident Analysis and Prevention*, 29 (5), pp. 651-665
- Implementing the Results of the Second Strategic Highway Research Program, special report 296, Transportation Research Board, 2009.

- Kaplan, R.M., and Saccuzzo, D.P. (2009). *Psychological testing: principles, applications and issues* (7th edition, Belmont CA: Wadsworth)
- Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., and Ramsey, D.J. (2006a). *The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car naturalistic driving study data*. Contract No. DTNH22-00-C-07007. Blacksburg, VA: Virginia Tech Transportation Institute.
- Lansdown, T. C. (2002). Individual differences during driver secondary task performance: verbal protocol and visual allocation findings. *Accident Analysis and Prevention*, 34, 655 - 662.
- Lassarre, S., Dozza, M., Jamson, S., Lai, F., Saad, F., Vadeby, A., Victor, T., Brouwer, R., Carsten, O., Disilvestro, A., Feenstra, P., Hogema, J., Lenard, J., Minett, C., Morris, A., Van Noort, M., Schrijver, J., Taddei, A. (2008). FESTA - D2.4 Data analysis and modelling.
- Maycock, G., Lockwood, C. R. and Lester, F. (1981). The accident liability of car drivers. Research Report 315. Transport and Road Research Laboratory, Crowthorne, U. K.
- McCullagh, P., and Nelder, J.A. (1989). *Generalized linear models* (2nd edition; New York: Chapman and Hall/CRC).
- Mee, R.W. (2009). *A comprehensive guide to factorial two-level experimentation* (New York: Springer).
- Mittlerman, M.A., Maclure, M., and Robins, J.M. (1995). Control sampling strategies for case-crossover studies: An assessment of relative efficiency. *American Journal of Epidemiology*, Volume 142, 1, p.91-98.
- Molenberghs, G., and Verbeke, G. (2005). *Models for Discrete Longitudinal Data* (New York: Springer).
- Montgomery, D.C. (2009). *Design and analysis of experiments* (7th edition; New York: Wiley).
- Montgomery, D.C., Peck, E.A., and Vining, G.G. (2006). *Introduction to Linear Regression Analysis* (4th edition, New York: Wiley).
- Neale, V.L., Klauer, S.G., Knipling, R.R., Dingus, T.A, Holbrook, G.T., and Petersen, A. (2002) The 100 Car Naturalistic Driving Study; Phase I – Experimental Design. Contract No. DTNH22-00-C-07007. Blacksburg, VA: Virginia Tech Transportation Institute.
- Olson, R. L., Hanowski, R. J., Hickman, J. S., Bocanegra, J. (2009). *Driver distraction in commercial vehicle operations*. Report No: FMCSA-RRR-09-042. U.S. Department of Transportation, Federal Motor Carrier Safety Administration, Washington D. C.
- Özkan T, & Lajunen T. (2005). Multidimensional Traffic Locus of Control Scale (TLOC): factor structure and relationship to risky driving. *Personality and Individual Differences*, 38, 533-545.
- Padlo, P., Aultman-Hall, L., and Stamatiadis, N. (2005). Passengers and other factors affecting the safety of young and older drivers, *Transportation research record Vol. 1937*, Transportation Research Board, Washington, DC pp. 7–13
- Parker, D., Manstead, A. S., Stradling, S. G., & Reason, J. T. (1992a). Determinants of intention to commit driving violations. *Accident Analysis & Prevention*, 24(2), 117-131.
- Preusser, D.F., Ferguson, S.A. and Williams, A.F. (1998). The effect of teenage passengers on the fatal crash risk of teenage drivers, *Accident Analysis and Prevention*, 30, pp. 217–222

- Rothman, K.J., Greenland, S., and Lash, T.L. (2008). *Modern epidemiology*. Third edition. Philadelphia : Wolters kluwer | Lippincox Williams & Wilkins.
- Rudin-Brown, C. M., & Parker, H. A. (2004). Behavioural adaptation to adaptive cruise control (ACC): Implications for preventive strategies. *Transportation Research Part F*, 7, 59-76.
- Sagberg, F. (2001). Accident risk of car drivers during mobile telephone use. *International Journal of Vehicle Design*, 26 (1), pp. 57–69.
- SHRP 1 (1996) *RoadSavers*, Publication No. FHWA-SA-96-045, Federal Highway Administration, Washington, D.C.
- SHRP 1 (1997) *From Research to Reality: Assessing the Results of the Strategic Highway Research Program*, Publication No. FHWA-SA-98-008, Federal Highway Administration, Washington, D.C.
- SHRP 1 (in press) *Economic Benefits of SHRP Research Program Background Report*, Texas Transportation Institute.
- SHRP2 (2009). S02 Integration of Analysis and Methods and Development of Analysis Plane, Phase 1 Report. University of Iowa, Iowa State University, Montana State University.
- Taylor, H., and Karlin, S. (1994). *An introduction to stochastic modelling* (New York: Academic Press).
- Telephone Interview with Dr. Kenneth Campbell, SHRP2 Safety Chief Program Officer, by Dr. Kerry Malone, on January 7, 2010
- The 100-Car Naturalistic Driving Study (2002). Phase I – Experimental Design. U.S. Department of Transportation, National Highway Traffic Safety Administration.
- The 100-car naturalistic driving study (2006). Phase II – Results of the 100-car field experiment. Springfield: U.S. Department of Transportation, the National Highway Traffic Safety Administration.
- Ulleberg, P. (2002). Personality subtypes of young drivers. Relationship to risk-taking preferences, accident involvement, and response to traffic safety campaign, *Transportation Research Part F*, 4, p. 279–297.
- West, B., Welch, K.B., and Galecki, A.T. (2006). *Linear Mixed Models: a practical guide using statistical software* (New York: Chapman and Hall/CRC)
- Zhou, M. and Lyles, R.W. (1997) Self-selection bias in driver performance studies. *Transportation Research Board, Issue number 1573*
- Zuckerman, M. (1994). *Behavioural expressions and biosocial bases of sensation seeking*. Cambridge: University of Cambridge Press.

Appendix I: Informed consent for drivers of leased vehicles

APPENDIX A: INFORMED CONSENT FOR DRIVERS OF LEASED VEHICLES

INFORMED CONSENT FOR PARTICIPANTS IN RESEARCH PROJECTS INVOLVING HUMAN SUBJECTS

Title of Project: Naturalistic Driving Study

Research Conducted by: Virginia Tech Transportation Institute (VTTI)

Research Sponsored by: National Highway Traffic Safety Administration (NHTSA)

Investigators: Dr. Tom Dingus, Dr. Vicki Neale, Sheila Klauer, Dr. Ron Knipling, Heather Foster

I. PURPOSE OF THIS RESEARCH PROJECT

The objective of this study is to collect data on driving behavior. There are no special tasks for the driver to perform; instead, the driver is requested to merely drive as they regularly would to their normal destinations. This instrumentation is designed such that it will in no way interfere with the driving performance of the vehicle and will not obstruct the driver in any way. Due to the number of vehicles that are being instrumented and the time period involved, it is likely that crashes and the events leading up to them will be recorded.

One hundred high-mileage drivers are being recruited to participate in this research. All age groups and both men and women are being asked to participate. To participate, drivers must have a valid drivers' license and own a vehicle of which they are the primary driver for the experimental period of one year.

II. PROCEDURES AND SUBJECT RESPONSIBILITIES

The following describes procedures for the study and participant responsibilities:

Preparation for study:

1. Review entire study information package.
2. Read this informed consent form carefully; make a note of any questions. You may call Heather Foster of VTTI (703-538-8447) to discuss any questions.
3. Sign and date this form.
4. Ensure that any person likely to drive the instrumented vehicle has signed this consent form. (If you wish to add another driver at a later time, an informed-consent form can be obtained from VTTI.)
5. Provide close-up pictures (head-shots) of all consenting drivers.

In-processing (requires two hours):

6. Call Heather Foster of VTTI at 703-538-8447 to schedule an appointment for in-processing.
In-processing will ordinarily be scheduled for 8-10 a.m. or 4-6 p.m. on selected weekdays, and 9-11 a.m. on Saturdays, at the VT Northern Virginia Center, 7054

Haycock Road, Falls Church, VA 22043. (Parking is available in the Visitors Parking Lot.)

7. Bring the following to the subject in-processing:
 - Signed informed consent form (this document)
 - Valid driver's license
 - Social Security Number
 - Two forms of identification
8. Listen to a short overview orientation to the study, and "Q&A" discussion. Sign remaining administrative forms; a copy of all signed forms will be provided to you for your records.
9. Review insurance protocol for the leased vehicle.
10. Take a vision exam.
11. Take a hearing exam. (Note: a free hearing exam is available for all prospective drivers, family members, and other frequent passengers, provided they agree to the re-testing in the event of an air bag deployment.)
12. Complete surveys regarding your health, sleep hygiene, stress levels, overall personality, and driving behaviors and practices.
13. Take one or more brief performance tests.
14. Schedule VTTI delivery of the leased vehicle to your home or workplace.

Data collection during driving:

15. Wear your safety belt at all times.
16. Drive your vehicle as you normally would.
17. Do not wear sunglasses unless absolutely necessary.
18. In the event of a safety-related incident, [i.e. a crash, near-crash, driving error, or unsafe condition involving you vehicle or adjacent vehicles], press the red incident button located above the rear-view mirror after the incident as soon as it is safe to do so. For one minute, a microphone (directed toward the driver) will be activated; during this time, please briefly describe what happened, and why. In particular, what was the driving error that caused the incident?

Data downloading:

Note: the location of your vehicle will be known to VTTI researchers via a radio transmitter providing Global Positioning System (GPS) coordinates. This information will be used to locate vehicles for data downloading.

19. Permit VTTI researchers to access the vehicle (at your home or work location) every 1-4 weeks to download data. Most data downloads will require a data line to be plugged into a data port near the vehicle's rear license plate on the outside of the vehicle. (No access to the inside of the vehicle is required.) Subject to your approval, data downloads will be completed between 7 a.m. and 11 p.m.

Equipment and vehicle maintenance:

20. In the event of equipment malfunction or damage, notify VTTI as soon as possible.
21. Permit a service call at your home or office for repairs (if preferred, vehicle may be brought to Hurley's). If repairs cannot be made in a service call, bring the vehicle to Hurley's for repairs. VTTI will provide \$10 to cover Metro fare or other transportation needs.

22. Buy regular, unleaded gasoline for the vehicle. Perform regular safety checks; e.g., once monthly, check tire pressure, oil level, and other fluids. Have oil changes and other preventive maintenance performed per a schedule and instructions provided to you by VTTI.

In the event of a crash: Study Procedures (applies to all collisions, regardless of severity):

23. Contact VTTI as soon as possible after the crash. (Accident reporting instructions and phone numbers will be left in the glove box of the leased vehicle.)
24. Participate in a short phone interview with VTTI about the crash. In addition, since you are driving a vehicle owned by the State of Virginia, there are two reporting requirements following accidents, one for this study and one for the state (Virginia Tech Motor Pool), which will be explained to you during in-processing.
25. Schedule an appointment for hearing re-testing, to be conducted **as soon as possible** after the crash. Re-testing is conducted at Professional Hearing Services (6231 Leesburg Pike Suite 512 Falls Church, VA 22044 Phone 703-536-1666). Re-testing results will be provided to you and to VTTI.
26. Encourage all passengers whose hearing has been tested to schedule this re-testing.
27. If the crash is police reported, request a copy of the Police Accident Report from the police, and provide a copy to VTTI. VTTI will remove all personal identifiers to ensure confidentiality. "Personal identifiers" include names, addresses, phone numbers, and license plate numbers.
28. Request and provide copies of medical report(s) associated with your crash injuries and treatment. For some crashes, crash and injury information may already be available to NHTSA, and thus to this study, in conjunction with other NHTSA-sponsored studies in the Northern Virginia area.
29. Permit VTTI and/or Hurley's to check and test the vehicle instrumentation.

In the event of a crash: Virginia Tech Motor Pool Procedures

30. Follow the instructions in the glove compartment.
31. Contact VTTI as soon as possible, we will assist you in filing the Virginia Tech Motor Pool accident report.

In the event of an air bag deployment:

32. Permit a Special Crash Investigation team from NHTSA to inspect the vehicle.
33. Participate in an in-person interview with the Crash Investigation team.

Vehicle Return:

VTTI will contact you at the end of the 12-month study, to schedule out-processing and return of the leased vehicle.

34. Bring your leased vehicle to the VT North Virginia Center to return. VTTI will provide \$10 to cover Metro fare or other transportation.

Out-processing/study completion (requires one hour):

35. Complete out-processing administrative paperwork.
36. Complete short questionnaires regarding stress levels, driving behavior and

performance over the past year, and study evaluation.

Equipment Installation and Data Collection

You are being asked to drive with the instrumentation for approximately one year. The data on the vehicle will be downloaded via a data port located behind the rear license plate or via short range wireless communication (if there is no access to the vehicle). Once the data is downloaded, it will be stored on a project specific data server that will be accessed only by research staff affiliated with the project.

The data collection system is designed to require no maintenance and will not require you to perform any maintenance. However, if a diagnostic check of the data confirms a disruption of the data collection, a hardware engineer will be assigned to correct the problem. To perform the maintenance, VTTI or Hurley's will contact you to receive permission to work on the vehicle and schedule the repair. We will try to avoid interfering with your commuting schedule.

Automobile Insurance

In the Commonwealth of Virginia, responsibility for automobile insurance resides with the owner of the vehicle.

In the event of an accident or injury in a Virginia Tech automobile, the University will provide automobile liability coverage for property damage and personal injury. The total policy amount per occurrence is \$2,000,000. This coverage (unless the other party was at fault, which would mean all expense would go to the insurer of the other party's vehicle) would apply in case of an accident for all volunteers and would cover medical expenses up to the policy limit. In the event of an accident, you must notify the police and the VT Motor Pool (contact information will be left in the glove compartment of the leased vehicle).

VT also carries as a part of its automobile liability insurance a "Med Pay" endorsement that will pay up to \$5,000 in medical expenses, until fault in an accident is determined, at which time all medical expenses would go to the insurer of the vehicle at fault.

If you are working as an employee for another company, you may be deemed to be driving in the course of your employment, and your employer's worker's compensation provisions may apply in lieu of the Virginia Tech and Commonwealth of Virginia insurance provisions, in case of an accident. The particular circumstances under which worker's compensation would apply are specified in Virginia law. If worker's compensation provisions do not apply in a particular situation, then Virginia Tech and Commonwealth of Virginia insurance will provide coverage.

Medical Insurance

Participants in a study are considered volunteers, regardless of whether they receive payment for their participation; under Commonwealth of Virginia law, workers compensation does not apply to volunteers; therefore, if not in an automobile, the participants are responsible for their own medical insurance for bodily injury. Appropriate health insurance is strongly recommended to cover these types of expenses.

If you should become injured in an accident, whether in or out of an automobile, the medical treatment available to you would be that provided to any person by emergency medical services in the vicinity where the accident occurs.

A Virginia Tech automobile accident report form is located in the glove compartment of the vehicle you will be driving and outlines what you should do if you become involved in an accident and are not incapacitated.

Automatic Collision Notification

The vehicle will also be equipped with an automatic collision notification system, triggered by collision impacts. The system is intended to notify VTTI in the event of a collision impact. When serious impacts are detected by VTTI staff, they will notify local emergency services. *However, VTTI cannot guarantee continuous 24-hour coverage or coverage of all vehicle locations.* Therefore, in the event of a crash, you should not expect an emergency response based on this system. *Notify police and emergency services as you otherwise would following a crash.* However, this automatic collision notification system *may* enable emergency service to be dispatched to you faster after a crash.

III. RISKS

The risk to you is no more than you would normally incur while driving. All data collection equipment is mounted such that, to the greatest extent possible, it does not pose a hazard in any foreseeable way. None of the data collection equipment will interfere with any part of your normal field of view. The addition of the data collection systems to the vehicle will in no way affect the operating or handling characteristics of the vehicle.

Please note that you are being asked not to wear sunglasses unless absolutely necessary; however, if at any time you are suffering from glare problems (e.g., from the sun shining directly into your face) and cannot see the roadway and your surrounding environment, sunglasses are recommended.

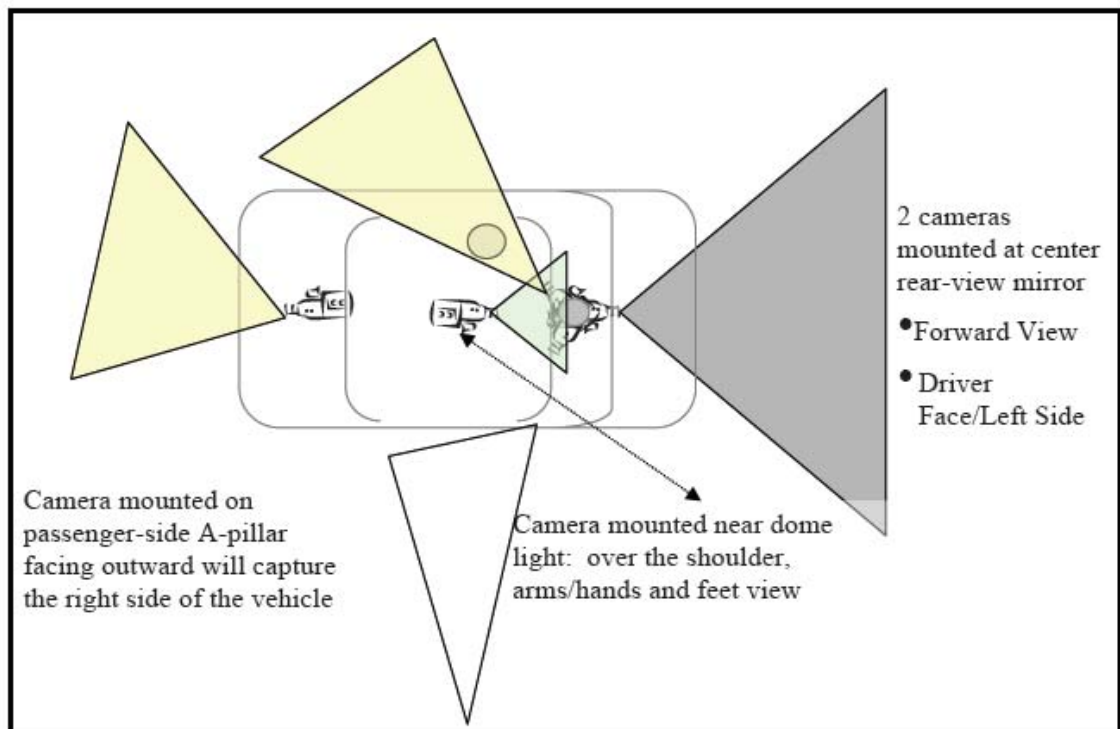
IV. BENEFITS

While there are no direct benefits to you from this research, you may find the experiment interesting. No promise or guarantee of benefits is being made to encourage participation. Your participation will help to improve the body of knowledge regarding driving behavior and performance.

V. EXTENT OF ANONYMITY AND CONFIDENTIALITY

Video information will be taken during the course of data collection. The data gathered in this experiment will be treated with confidentiality. Driver names will be separated from the collected data. A coding scheme will be employed to identify the data by subject number only (e.g., Driver No. 3).

While you are driving the vehicle, a camera will record your face and the left exterior side of vehicle, the right exterior side of the vehicle, the forward view, the rear-view, and the instrument panel view. This is shown below. Note that no other passengers in the vehicle will be within the camera view. Also, there is audio recording capability in the vehicle, but it will only record for one minute when you activate the incident push button. Please note that the audio microphone is directional and will only record your voice from the driver's seat.



The data from this study will be stored in a secured area at the Virginia Tech Transportation Institute. Access to the data will be under the supervision of Dr. Tom Dingus, Dr. Vicki Neale, Sheila Klauer, Dr. Ron Knipling, and Heather Foster. Data reductionists assigned to work on this project will also have access to your data. Data reduction will consist of examining driving performance under various conditions. During the course of this study, the video will not be released to anyone other than individuals working on the project without your written consent. Following the study, some data may be made available to the contact sponsor, the National Highway Traffic Safety Administration, for research purposes only. Please note that NHTSA is under the same obligation to keep your data confidential.

If you are involved in a crash while participating in this study, the data collection equipment in your vehicle will likely capture the events leading up to the event. The data collection equipment SHOULD NOT be given to police officers or any other party. You are under NO LEGAL OBLIGATION to mention that you are participating in this study.

We will do everything we can to keep others from learning about your participation in the research. To further help us protect your privacy, the investigators have obtained a Confidentiality Certificate from the Department of Health and Human Services. With this Certificate, the investigators cannot be forced (for example by court subpoena) to disclose information that may identify you in any Federal, State, or local civil, criminal, administrative, legislative, or other proceedings. Disclosure will be necessary, however, upon request of DHHS for audit or program evaluation purposes.

You should understand that a Confidentiality Certificate does not prevent you or a member of your family from voluntarily releasing information about yourself or your involvement in this research. Note however, that if an insurer, employer, or someone else learns about your participation, and *obtains your consent* to receive research information, then the investigator may not use the Certificate of Confidentiality to withhold this information. This means that you and your family must also actively protect your own privacy. In addition to the Confidentiality Certificate, we have also obtained approval through the NHTSA Human Use Review Panel for your protection.

Finally, you should understand that the investigator is not prevented from taking steps, including disclosing information to authorities, to prevent serious harm to yourself or others. For example, if we learned about offenses such as child abuse or habitual driving under the influence, we would take appropriate action to protect you and someone else, even though we will still maintain privacy of the data.

VI. COMPENSATION

You will be compensated \$125 per month for approximately 12 months of participation in this study. If you choose to withdraw from participation prior to the 12-month period, you will be compensated for the proportion of time that you have participated. You will also receive a \$300 study completion bonus at the end of the 12-month period and equipment de-installation. This bonus will be provided at the out-processing.

In addition to this compensation, you will be given \$10 for travel on the days that instrumentation is installed and removed.

VII. FREEDOM TO WITHDRAW

You are free to withdraw from the study at any time without penalty. If you choose to withdraw, you will be compensated for the portion of the time of the study.

VTTI has the right to terminate your participation in the study at any time. For example, VTTI may withdraw you from the study if the quantity or quality of data is insufficient for study purposes or if you pose a threat to yourself or to others. Subjects withdrawn from the study will receive pro-rated payment (at \$125 per month) and will be required to schedule equipment de-installation as soon as possible.

VIII. APPROVAL OF RESEARCH

This research project has been approved, as required, by the Institutional Review Board for Research Involving Human Subjects at Virginia Polytechnic Institute and State University, by the Virginia Tech Transportation Institute.

IRB Approval Date

Approval Expiration Date

IX. DRIVER'S RESPONSIBILITIES

I voluntarily agree to participate in this study. I understand the procedures and responsibilities described above, in particular in **Section II, Procedures and Subject Responsibilities**.

X. DRIVER'S PERMISSION

I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent:

Signature of Driver: _____
Date: _____

Signature of Additional Driver: _____ Date: _____

Signature of Legal Guardian if any additional driver is minors: _____ Date: _____

Signature of Additional Driver: _____ Date: _____

Signature of Legal Guardian if any additional driver is minors: _____ Date: _____

Signature of Additional Driver:

_____ Date:

Signature of Legal Guardian if any additional driver is minors:

_____ Date:

Should I have any questions about this research or its conduct, I may contact:

Heather Foster 703-538-8447 hfooster@vtti.vt.edu
Research Specialist/Northern Virginia Center, Virginia Tech Transportation Institute

Dr. Ronald R. Knipling 703-538-8439 rknipling@vtti.vt.edu
Northern Virginia Site Manager/Falls Church, Virginia Tech Transportation Institute

Dr. Vicki L. Neale 540-231-1514 vneale@vtti.vt.edu
Co- Principal Investigator, Virginia Tech Transportation Institute

Dr. David M. Moore 540-231-4991 moored@vt.edu
Chair, IRB
Office of Research Compliance
Research & Graduate Studies

All drivers must be given a complete copy (or duplicate original) of the signed Informed Consent.

Appendix II: Informed consent for drivers of private vehicles

[REVISED 10-22-02]

APPENDIX A: INFORMED CONSENT FOR DRIVERS OF PRIVATE VEHICLES

INFORMED CONSENT FOR PARTICIPANTS IN RESEARCH PROJECTS INVOLVING HUMAN SUBJECTS

Title of Project: Naturalistic Driving Study

Research Conducted by: Virginia Tech Transportation Institute (VTTI)

Research Sponsored by: National Highway Traffic Safety Administration (NHTSA)

Investigators: Dr. Tom Dingus, Dr. Vicki Neale, Sheila Klauer, Dr. Ron Knipling, Heather Foster

I. PURPOSE OF THIS RESEARCH PROJECT

The objective of this study is to instrument drivers' personal vehicles to collect data on driving behavior. There are no special tasks for the driver to perform; instead, the driver is requested to merely drive as they regularly would to their normal destinations. This instrumentation is designed such that it will in no way interfere with the driving performance of the vehicle and will not obstruct the driver in any way. Due to the number of vehicles that are being instrumented and the time period involved, it is likely that crashes and the events leading up to them will be recorded.

One hundred high-mileage drivers are being recruited to participate in this research. All age groups and both men and women are being asked to participate. To participate, drivers must have a valid drivers' license and own a vehicle of which they are the primary driver for the experimental period of one year.

II. PROCEDURES AND SUBJECT RESPONSIBILITIES

The following describes procedures for the study and participant responsibilities:

Preparation for study:

1. Review entire study information package
2. Read this informed consent form carefully; make a note of any questions. You may call Heather Foster of VTTI (703-538-8447) to discuss any questions.
3. Sign and date this form.
4. Ensure that any person likely to drive the instrumented vehicle has signed this informed consent form. (If you wish to add another driver at a later time, an informed consent form can be obtained from VTTI.)
5. Provide close-up pictures (head-shots) of all consenting drivers.

In-processing (requires two hours):

6. Call Heather Foster of VTTI at 703-538-8447 to schedule an appointment for in-processing.

In-processing will ordinarily be scheduled for 8-10 a.m. or 4-6 p.m. on weekdays, and 9-11 a.m. on Saturdays, at the VT Northern Virginia Center, 7054 Haycock Road, Falls Church, VA 22043. (Parking is available in the Visitors Parking Lot)

7. Bring the following to the subject in-processing:
 - Signed informed consent form (this document)
 - Valid driver's license
 - Proof of insurance for your vehicle
 - Vehicle registration

 - Social Security Number
 - Two forms of identification
8. Listen to a short overview orientation to the study, and Q&A discussion. Sign remaining administrative forms; a copy of all signed forms will be provided to you for your records.
9. Take a vision exam.
10. Take a hearing exam. (Note: A free hearing exam is available for all prospective drivers, family members, and other frequent passengers, provided they agree to the re-testing in the event of a crash.)
11. Complete surveys regarding your health, sleep hygiene, stress levels, overall personality, and driving behaviors and practices.
12. Take one or more brief performance tests.
13. Schedule your vehicle for equipment installation. (see below)

Equipment installation:

14. Bring your vehicle to Hurley's Auto Audio (1524 Springhill Road, McLean, VA 22102, Phone 703-790-8744) for equipment installation this will require a full day. We will provide \$10 to cover Metro fare or other transportation needs.

Data collection during driving:

15. Wear your safety belt at all times.
16. Drive your vehicle as you normally would.
17. Do not wear sunglasses unless absolutely necessary
18. In the event of a safety-related incident, [i.e. a crash, near-crash, driving error, or unsafe condition involving your vehicle or adjacent vehicles], press the red incident button located above the rear-view mirror after the incident as soon as it is safe to do so. For one minute, a microphone (directed toward the driver) will be activated; during this time, please briefly describe what happened, and why. In particular, what was the driving error that caused the incident?

Data downloading:

Note: the location of your vehicle will be known to VTTI researchers via a radio transmitter providing Global Positioning System coordinates. This information will be used to locate vehicles for data downloading.

19. Permit VTTI researchers to access your vehicle (at your home or work location) every 1-4 weeks to download data. Most data downloads will require a data line to be plugged into a data port near the vehicle license plate on the outside of the

vehicle. (No access to the inside of the vehicle is required.) Subject to your approval, data downloads will be completed between 7 a.m. and 11 p.m.

Equipment maintenance:

20. In the event of equipment malfunctioning or damage, notify VTTI as soon as possible.
21. Permit a service call at your home or office for repairs (if preferred, vehicle may be brought to Hurley's). If repairs cannot be made in a service call, bring the vehicle in to Hurley's for repairs. We will provide \$10 to cover Metro fare or other transportation needs.

In the event of a crash (applies to all collisions, regardless of severity):

22. Contact VTTI as soon as possible after the crash. (Accident reporting instructions and phone numbers will be placed in glove box during equipment installation.)
23. Participate in a short phone interview with VTTI about the crash.
24. Schedule an appointment for hearing re-testing, to be conducted as soon as possible after the crash. Re-testing is conducted at Professional Hearing Services (6231 Leesburg Pike Suite 512 Falls Church, VA 22044 Phone 703-536-1666). Re-testing results will be provided to you and to VTTI.
25. Encourage all passengers whose hearing has been tested to schedule this re-testing.
26. If the crash is police-reported, request a copy of the Police Accident Report from the police, and provide a copy to VTTI. VTTI will remove all personal identifiers to ensure confidentiality. "Personal identifiers" include names, addresses, phone numbers, and license plate numbers.
27. Request and provide copies of medical report(s) associated with your crash injuries and treatment. For some crashes, crash and injury information may already be available to NHTSA, and thus to this study, in conjunction with other NHTSA-sponsored studies in the Northern Virginia area.
28. Permit VTTI and/or Hurley's to check and test the vehicle instrumentation.

In the event of an air bag deployment:

29. Permit a Special Crash Investigation team from NHTSA to inspect the vehicle.
30. Participate in an in-person interview with the Crash Investigation team.

Equipment de-installation:

VTTI will contact you at the end of the 12-month study, to schedule equipment de-installation and out-processing.

31. Bring your vehicle to Hurley's Auto Audio for equipment de-installation, which will require a full day. We will provide \$10 to cover Metro fare or other transportation needs.
32. Inspect your vehicle at Hurley's and sign form to verify that all recording equipment has been removed, and that the vehicle has been restored to its original state. Keep copy for your records.

Out-processing/study completion (requires one hour):

33. Complete out-processing administrative paperwork.
34. Complete short questionnaires regarding stress levels and driving behavior and

- performance over the past year, and study evaluation.
35. Receive final payment for your participation.

Equipment Installation and Data Collection

You are being asked to drive with the instrumentation for approximately one year. No holes will be drilled into your vehicle to mount equipment. Instead, holes holding existing apparatus will be used. The data collection system is approximately 8" x 18" x 24." The computer/data storage system is housed in the back of the trunk and mounted to the trunk "roof" (not to the trunk lid). A camera module will be mounted above the rear-view mirror and an incident push-button will be located on the camera module. This will be done without drilling holes or making any permanent modifications to the vehicle. Wires will not be visible.

As part of the data collection system, forward- and rearward-looking radar will be installed behind the front and rear license plates. For the radar to function, we will need to replace you state license plate with plastic plates for the duration of the study. You will be provided with a temporary registration and an authorization letter from the state DMV for your records. At the end of the study your original license plates will be reinstalled on the vehicle.

The data on the vehicle will be downloaded via a data port located behind the rear license plate or via short range wireless communication (if there is no access to the vehicle). Once the data is downloaded, it will be stored on a project specific data server that will be accessed only by research staff affiliated with the project.

The data collection system is designed to require no maintenance and will not require you to perform any maintenance. However, if a diagnostic check of the data confirms a disruption of the data collection, a technician will be assigned to correct the problem. To perform the maintenance, VTTI or Hurley's will contact you to receive permission to work on the vehicle and schedule the repair. We will try to avoid interfering with your commuting schedule.

Insurance

Please note that since you are driving your own vehicle, Virginia Tech is not liable for the expenses incurred in any accident you may have. In the event of an accident, you are not responsible for coverage of the instrumentation in the vehicle.

Participants in a study are considered volunteers, regardless of whether they receive payment for their participation. Under Commonwealth of Virginia law, workers compensation does not apply to volunteers; therefore, the participants are responsible for their own medical insurance for bodily injury. Appropriate health insurance is strongly recommended to cover these types of expenses.

If you should become injured in an accident, whether in or out of an automobile, the medical treatment available to you would be that provided to any person by emergency medical services in the vicinity where the accident occurs.

Automatic Collision Notification

The vehicle will also be equipped with an automatic collision notification system, triggered by collision impacts. The system is intended to notify VTTI in the event of a collision impact. When serious impacts are detected by VTTI staff, they will notify local emergency services. *However, VTTI cannot guarantee*

continuous 24-hour coverage or coverage of all vehicle locations. Therefore, in the event of a crash, you should not expect an emergency response based on this system. *Notify police and emergency services as you otherwise would following a crash.* However, this automatic collision notification system *may* enable emergency service to be dispatched to you faster after a crash.

III. RISKS

The risk to you is no more than you would normally incur while driving. All data collection equipment is mounted such that, to the greatest extent possible, it does not pose a hazard in any foreseeable way. None of the data collection equipment will interfere with any part of your normal field of view. The addition of the data collection systems to the vehicle will in no way affect the operating or handling characteristics of the vehicle.

Please note that you are being asked not to wear sunglasses unless absolutely necessary; however, if at any time you are suffering from glare problems (e.g., from the sun shining directly into your face) and cannot see the roadway and your surrounding environment, sunglasses are recommended.

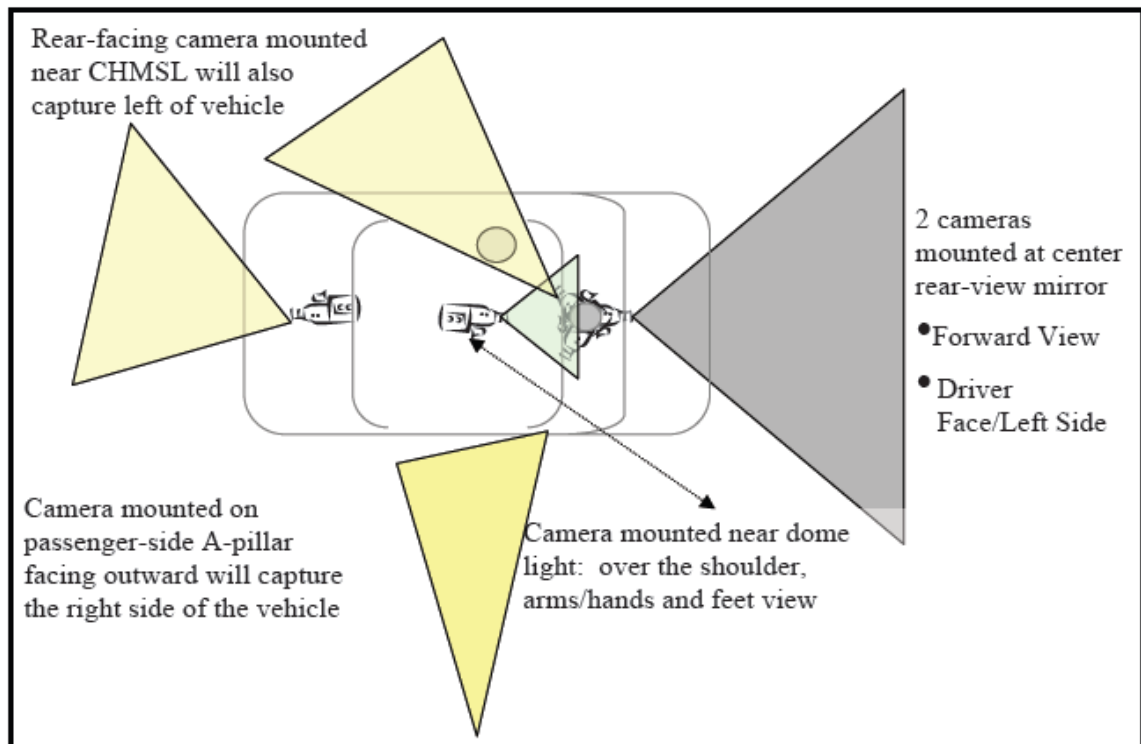
IV. BENEFITS

While there are no direct benefits to you from this research, you may find the experiment interesting. No promise or guarantee of benefits is being made to encourage participation. Your participation will help to improve the body of knowledge regarding driving behavior and performance.

V. EXTENT OF ANONYMITY AND CONFIDENTIALITY

Video information will be taken during the course of data collection. The data gathered in this experiment will be treated with confidentiality. Drivers' names will be separated from the collected data. A coding scheme will be employed to identify the data by subject number only (e.g., Driver No. 3).

While you are driving the vehicle, a camera will record your face and the left exterior side of vehicle, the right exterior side of the vehicle, the forward view, the rear-view, and the instrument panel view. This is shown below. Note that no other passengers in the vehicle will be within the camera view. Also, there is audio recording capability in the vehicle, but it will only record for one minute when you activate the incident push button. Please note that the audio microphone is directional and will only record your voice from the driver's seat.



The data from this study will be stored in a secured area at the Virginia Tech Transportation Institute. Access to the data will be under the supervision of Dr. Tom Dingus, Dr. Vicki Neale, Sheila Klauer, Dr. Ron Knipling, and Heather Foster. Data reductionists assigned to work on this project will also have access to your data. Data reduction will consist of examining driving performance under various conditions. During the course of this study, the video will not be released to anyone other than individuals working on the project without your written consent. Following the study, some data may be made available to the contact sponsor, the National Highway Traffic Safety Administration (NHTSA), for research purposes only. Please note that NHTSA is under the same obligation to keep your data confidential.

If you are involved in a crash while participating in this study, the data collection equipment in your vehicle will likely capture the events leading up to the event. The data collection equipment **SHOULD NOT** be given to police officers or any other party. You are under **NO LEGAL OBLIGATION** to mention that you are participating in this study.

We will do everything we can to keep others from learning about your participation in the research. To further help us protect your privacy, the investigators have obtained a Confidentiality Certificate from the Department of Health and Human Services. With this Certificate, the investigators cannot be forced (for example by court subpoena) to disclose information that may identify you in any federal, state, or local civil, criminal, administrative, legislative, or other proceedings. Disclosure will be necessary, however, upon request of DHHS for audit or program evaluation purposes.

You should understand that a Confidentiality Certificate does not prevent you or a member of your family from voluntarily releasing information about yourself or your involvement in this research. Note however, that if an insurer, employer, or someone else learns about your participation, and *obtains your consent* to receive research information, then the investigator may not use the Certificate of Confidentiality to withhold this information. This means that you and your family must also actively protect your own privacy. In addition to the Confidentiality Certificate, we have also obtained approval through the NHTSA Human Use Review Panel for your protection.

Finally, you should understand that the investigator is not prevented from taking steps, including disclosing information to authorities, to prevent serious harm to yourself or others. For example, if we learned about offenses such as child abuse or habitual driving under the influence, we would take appropriate action to protect you and someone else, even though we will still maintain privacy of the data.

VI. COMPENSATION

You will be compensated \$125.00 per month for approximately 12 months of participation in this study. If you choose to withdraw from participation prior to the 12-month period, you will be compensated for the proportion of time that you have participated. You will also receive a \$300 study completion bonus at the end of the 12-month period and equipment de-installation. This bonus will be provided at the out-processing.

In addition to this compensation, you will be given \$10 for travel on the days that instrumentation is installed and removed.

VII. FREEDOM TO WITHDRAW

You are free to withdraw from the study at any time without penalty. If you choose to withdraw, you will be compensated for the portion of the time of the study.

VTTI has the right to terminate your participation in the study at any time. For example, VTTI may withdraw you from the study if the quantity or quality of data is insufficient for study purposes or if you pose a threat to yourself or to others. Subjects withdrawn from the study will receive pro-rated payment (at \$125 per month) and will be required to schedule equipment de-installation as soon as possible.

VIII. APPROVAL OF RESEARCH

This research project has been approved, as required, by the Institutional Review Board for Research Involving Human Subjects at Virginia Polytechnic Institute and State University, by the Virginia Tech Transportation Institute.

IRB Approval Date

Approval Expiration Date

IX. DRIVER'S RESPONSIBILITIES

I voluntarily agree to participate in this study. I understand the procedures and responsibilities described above, in particular in **Section II, Procedures and Subject Responsibilities**.

X. DRIVER'S PERMISSION

I have read and understand the Informed Consent and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent:

Signature of Driver: _____ Date: _____

Signature of Additional Driver:

Date: _____
Signature of Legal Guardian if any additional driver is minors:

Date: _____

Signature of Additional Driver:

Date: _____
Signature of Legal Guardian if any additional driver is minors:

Date: _____

Signature of Additional Driver:

Date: _____

Signature of Legal Guardian if any additional driver is minors:

Date: _____

Should I have any questions about this research or its conduct, I may contact:

Heather Foster 703-538-8447 hfooster@vtti.vt.edu
Research Specialist/Northern Virginia Center, Virginia Tech Transportation Institute

Dr. Ronald R. Knipling 703-538-8439 rknipling@vtti.vt.edu
Northern Virginia Site Manager/Falls Church, Virginia Tech Transportation Institute

Dr. Vicki L. Neale 540-231-1514 vneale@vtti.vt.edu
Co-Principal Investigator, Virginia Tech Transportation Institute

Dr. David M. Moore 540-231-4991 moored@vt.edu
Chair, IRB
Office of Research Compliance
Research & Graduate Studies

All drivers must be given a complete copy (or duplicate original) of the signed Informed Consent.

List of Figures

Figure 5.1: Strata of variation in ND studies.....	32
--	----

List of Tables

Table 3.1: Demographic and personality variables.....	18
Table 3.2: Variables describing driving behaviour.....	19
Table 3.3: Variables describing driver distraction and state.....	20
Table 3.4: Variables describing vehicle condition.....	20
Table 3.5: Variables describing vehicle parameters.....	21
Table 3.6: Variables describing environmental variables.....	22
Table 3.7: Variables describing road condition	23
Table 3.8: Variables describing traffic conditions	24
Table 6.1: Generic guide to scheduling the 22 Activities described in the FOTIP, Annex B of FESTA Handbook	42
Table 6.2: Organization of the 100 Car Naturalistic Driving Study.....	43

List of Abbreviations

CO:	Confidential Deliverable
DAS:	Data Acquisition System
FOT:	Field Operational Test
GLM:	General Linear Model
GLMM:	General Linear Mixed Model
ND:	Naturalistic Driving
PU:	Public Deliverable
QA:	Quality Assurance
RE:	Restricted Deliverable
VTTI:	Virginia Tech Transportation Institute
WP:	Work Package