FESTA Handbook

Version 2

19 August 2008

Project:	Field opErational teSt supporT Action (FESTA)
Grant agreement no.:	214853
Workpackage:	WP6 Implementation Aspects
Task:	T6.4 FESTA Handbook
Deliverable no.:	D6.4
Document title:	FESTA Handbook
Deliverable nature:	Public
Deliverable status:	Final
Authors:	FESTA consortium

FESTA Consortium:

Centro Ricerche Fiat, University of Leeds, BMW Forschung und Technik GmbH, Daimler AG, Gie Recherches et etudes PSA Renault, Volvo Car Corporation, Volvo Technology Corporation, Robert Bosch GmbH, A.D.C. Automotive Distance Control Systems GmbH, Delphi France SAS, Loughborough University, Chalmers University of Technology, Institut National de Recherche sur les Transports et leur Sécurité INRETS, Statens Väg-och Transportforskningsinstitut VTI, Nederlandse Organisatie voor toegepast-natuurwetenschappelijk onderzoek TNO, Bundesanstalt fuer Strassenwesen BASt, Valtion Teknillinen Tutkimuskeskus VTT, INFOBLU SPA, Orange France, European Road Transport Telematics Implementation Coordination Organisation ERTICO, Universitaet zu Koeln

Acknowledgements

We would like to thank all the individuals who have contributed to this handbook. Special thanks are due to the various external experts for the valuable insights they contributed. They are Jim Sayer of UMTRI, Charlie Klauer of VTTI, Tom Triggs, Kristie Young, Eve Mitsopoulos-Rubens, Nebojsa Tomasevic and Karen Stephan of MUARC, Harri Peltola of VTT and Riku Kotiranta of Chalmers University of Technology.

List of Abbreviations

AC	Alternating Current
ACC	Adaptive Cruise Control
ACEA	European Automobile Manufacturers' Association
AD	Aftermarket Device
ADAS	Advanced Driver Assistance Systems
BCR	Benefit Cost Ratio
BLOB	Binary Large Object
C2C-CC	Car-to-Car Communication Consortium
CALM	Continuous Air interface for Long and Medium range communication
CAN	Controller Area Network
CBA	Cost Benefit Analysis
CPI	Consumer Price Index
DAQ	Data Acquisition
DAS	Data Acquisition System
DBA	Database Administrator
DBQ	Driver Behaviour Questionnaire
DC	Direct Current
DGPS	Differential Global Positioning System
ECU	Electronic Control Unit
EDR	Event Data Recorder
EMC	Electromagnetic Compatibility
FCW	Forward Collision Warning
FM	Frequency Modulation
FOT	Field Operational Test
GIS	Geographical Information System
GPS	Global Positioning System
GSM	Global System for Mobile communications
GUI	Graphical User Interface
I2V	Infrastructure to Vehicle
ICT	Information Communication Technology
IEEE	Institute of Electrical & Electronics Engineers
I/O	Input/Output
IRR	Internal Rate of Return
ISA	Intelligent Speed Adaptation
ITS	Intelligent Transportation System
IVIS	In-Vehicle Information Systems
IVSS	Intelligent Vehicle Safety System
JPEG	Joint Photographic Experts Group
LDW	Lane Departure Warning
LED	Light Emitting Diode
LIDAR	Light Detection and Rating
LIN	Local Interconnect Network
MCA	Multi Criteria Analysis

MJPEG	Motion JPEG
MOST	Media Oriented Systems Transport
MM	Minimum Mean
MP3	Moving Picture Experts Group Layer-3 Audio
MPEG	Moving Picture Experts Group
MPEG-4	MPEG standard, version 4
NAS	Network Attached Storage
ND	Nomadic Device
NPV	Net Present Value
NTP	Network Time Protocol
OEM	Original Equipment Manufacturer
PI	Performance Indicator
PND	Personal Navigation Devices
RAID	Redundant Arrays of Independent Disks
SAN	Storage Area Network
SMS	Short Message Service
SNTP	Simple Network Time Protocol
SP	Smart Phone
SQL	Structured Query Language
SSD	Solid State Drive
SSS	Sensation Seeking Scale
T-LOC	Traffic Locus of Control
TTC	Time To Collision
UTC	Universal Time Coordinated
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
VMS	Variable Message Sign
WBS	Work Breakdown Structure
WLAN	Wireless Local Area Network
WTP	Willingness To Pay

Table of Contents

1 Introduction

In Japan and in the United States Field Operational Tests (FOTs) have been introduced as an evaluation method for driver support systems and other functions several years ago with the aim of proving that such systems can deliver real-world benefits. In Europe too, FOTs have been conducted at a national or regional level, particularly on speed support systems and lane departure warning systems. These FOTs have proven to be highly valuable. Recently FOTs have been identified as an important means of verifying the real-world impacts of new systems at a European level and in particular to verify that European R&D has the potential to deliver identifiable benefits. This Handbook is the result of a joint effort of several research institutes, OEMs and other stakeholders from across Europe to prepare a common methodology for European FOTs. It is also highly relevant, and it is hoped useful, for FOTs conducted at a regional or national level within Europe as well as outside Europe.

For the purposes of this Handbook, a "Field Operational Test" (FOT) is defined as:

> A study undertaken to evaluate a function, or functions, under normal operating conditions in environments typically encountered by the host vehicle(s) using quasi-experimental methods.

This means that it must be possible to compare the effects that the function has on traffic with a baseline condition during which the function is not operating. In order to achieve this, the drivers' control over or interaction with the function(s) has to be manipulated by the research team. "Normal operating conditions" implies that the drivers use the vehicles during their daily routines, that data logging works autonomously and that the drivers do not receive special instructions about how and where to drive. Except for some specific occasions, there is no experimenter in the vehicle, and typically the study period extends over at least a number of weeks.

The main purpose of this Handbook is to provide guidelines for the conduction of FOTs. It walks the reader through the whole process of planning, preparing, executing, analysing and reporting an FOT, and it gives information about aspects that are especially relevant for a study of this magnitude, such as administrative, logistic, legal and ethical issues. Another aspect of the Handbook is to pave the road for standardisation of some aspects of FOTs, which would be helpful for cross-FOT comparisons. It has to be kept in mind, though, that many traffic parameters in different European countries differ substantially.

In Figure 1.1 the steps that need to be carried out during an FOT are presented. They will be explained in detail in the different chapters of the Handbook. For orientation purposes, a copy of the figure is provided in the beginning of each chapter highlighting which step of the FOT Chain is described in the current chapter. The FOT Implementation Plan takes up all the steps and integrates them into one big table which can be used as a reference when actually carrying out an FOT.

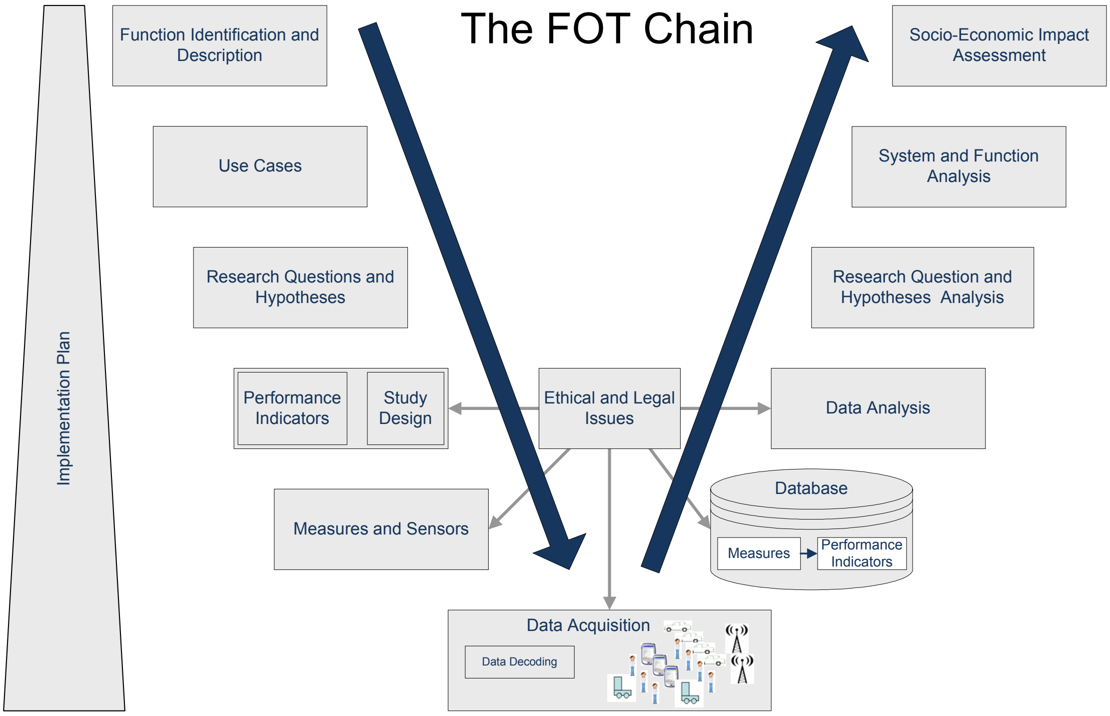


Figure 1.1: The steps that typically have to be considered when conducting an FOT. The large arrows indicate the time line.

The first steps, which include setting up a goal for the study and selecting a suitable research team, and also the last steps that include an overall analysis of the systems and functions tested and the socio-economic impact assessment, deal with the more general aspects of an FOT and with aggregation of the results. The further down on FOT Chain V-Shape the steps are located, the more they focus on aspects with a high level of detail, like which Performance Indicators to choose, or how to store the data in a database. The ethical and legal issues have the strongest impact on those high-level aspects, where the actual contact with the participants and the data handling takes place.

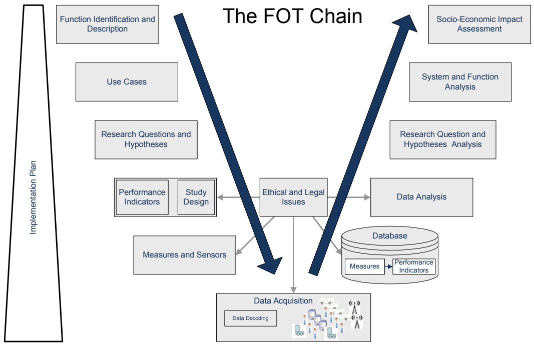
The FESTA Handbook is not meant to be a substitute for consultations with experts, organising a good and capable research team, and carrying out specific investigations into the legal and ethical issues that apply to the current question and situation. It is not an exhaustive action list, and each FOT has its own special issues and concerns that have to be dealt with on an individual basis. Nor is the advice in it necessarily perfect and representative of the state of the art. On many issues, there will be scope for disagreement with the recommendations or use of alternative sources of advice. But it would certainly be preferable for major departures from the advice to be justified to funding agencies and major stakeholders.

The FESTA project consortium decided early in the project that the primary focus of the FESTA Handbook would be on the evaluation of Advanced Driver Assistance Systems (ADAS) and In-Vehicle Information Systems (IVIS) for vehicles — both in the form of autonomous systems and of cooperative systems. It was also agreed that the FESTA Handbook should be relevant to the evaluation of Original Equipment Manufacturer (OEM), aftermarket and nomadic systems. The Handbook is therefore designed specifically to guide the evaluation of such systems, and is less relevant to the evaluation of electronic road infrastructure such as Variable Message Signs (VMS). However, it will be seen that many of the activities identified in the Handbook are common to the evaluation of most vehicle- and infrastructure-based ICT technologies).

In conclusion, the FESTA Handbook gives an overview and general guidelines concerning the conduction of a FOT. FOTs are designed to participate to the solution of a problem, and this handbook is intended to provide a formalised and practical framework, and not a cook book: the methodology described will necessarily have to be adapted to the specific case, in order to increase the efficiency of the approach or tackle data incompleteness or inconsistency. Even more, the results of a FOT may have to be integrated with external sources of information, to achieve a wider perspective, and an increased relevance for tackling the problem at stake .

In addition to the Handbook itself, more detailed work, which was produced during the FESTA project and which is referenced in the Handbook, is included in the annexes to the Handbook and the FESTA deliverables.

2 Planning and Running a Field Operational Test



2.1 Introduction

For a Field Operation Test (FOT) 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. Given that the lifecycle of a FOT typically evolves through many phases, there are many issues to consider. In this chapter, the critical activities and tasks which are necessary to run a successful FOT are documented — in the form of a "FOT Implementation Plan" (FOTIP) — drawing on lessons learned from previous FOTs conducted in Europe, the United States, Japan, Australia and elsewhere.

The FOTIP is contained in Annex B of the FESTA Handbook. In this chapter, the FOTIP is introduced, described, explained and discussed.

2.2 The FOT Implementation Plan

2.2.1 Purpose

The FOTIP is intended to serve primarily as a checklist for planning and running FOTs:

- to highlight the main Activities and Tasks that would normally be undertaken in successfully completing a FOT;
- to ensure that, in running a FOT, researchers and support teams are aware of critical issues that influence the success of the FOT;
- by drawing on the experiences of previous FOTs, to highlight the "dos" and "don'ts" of running a FOT; and
- to provide a consistent framework for planning, running and decommissioning FOTs.

The FOTIP presented in this Handbook is not intended to be prescriptive, but rather to serve as a generic guide in conducting FOTs. By their very nature FOTs are major projects – extensive and expensive. Significant previous FOTs that have not delivered their anticipated outcomes have not done so primarily because of failures to anticipate problems that compromised their successful execution. The FOTIP attempts to map out all known critical issues that need to be taken into account in planning and undertaking a FOT.

The history of FOTs suggests that no two will be the same, and that there often are many unforeseen Tasks and Sub-Tasks that arise during its lifecycle. The list of Tasks and Sub-Tasks contained in the FOTIP in Annex B of this Handbook is not, therefore, exhaustive. It is based on the collective wisdom of those that have been involved in planning and running previous FOTs. There may be specific requirements for future FOTs conducted in Europe that will need to be decided on a case-by-case basis.

The FOTIP at Annex B describes *what* needs to be done, and approximately *when*, in running a successful FOT. Other relevant chapters in the FESTA handbook describe in detail *why* these activities are necessary and *how* they are to be accomplished.

2.2.2 Description of the FOT Implementation Plan

The FOTIP at Annex B of this Handbook resembles a traditional Work Breakdown Structure (WBS), but without timelines. It is specifically designed in this way so that timelines can be inserted at a later date by those responsible for the overall planning and running of the FOT.

The FOTIP is divided into 5 columns:

- **Column 1 — Activities**. An Activity is a high level task e.g. "Convene FOT research and support teams" that is usually needed to run a FOT.
- **Column 2 — Tasks and Sub-Tasks**. A Task directly supports an Activity e.g. "Appoint FOT project manager". A Sub-Task directly supports a Task. Essentially, this column contains a series of action statements – "do this"; "do that"; etc. There are very few sub-tasks listed in this column, to contain the size of the document. The document is cross-referenced to other chapters of the FESTA Handbook, which identify the relevant Sub-Tasks that support these Tasks.
- **Column 3 — Person/Organisation Responsible for Activity**. This column identifies the person, team, organisation or combination thereof that would usually be responsible for completion of a Task. The FOT project manager is ultimately accountable for successful completion of all Tasks, and is therefore included for every Task. Column 3 contains a list of numbers, each of which denotes a particular person, team or organisation. The table in Annex B (immediately preceding the FOTIP) provides a legend for these numbers.
- **Column 4 – Critical Considerations (the "dos" and "don'ts").** This column contains critical advice for ensuring that an Activity or Task is successfully completed. e.g. "Be sure that the vehicle systems are designed so they do not drain the battery when the vehicle engine is not running." e.g. "Do not underestimate the amount of time required to recruit company drivers for the FOT."
- **Column 5 – General Advice.** This column provides general advice on how to maximise the likelihood of running a successful FOT e.g. "The FOT lifecycle is long. Hence, it is advisable to write separate reports on each critical stage of the FOT…" This column also contains explanatory notes, reference to other relevant documents (e.g. FOT reports) and cross-referencing to other chapters in the FESTA Handbook.

The Activities and Tasks identified in the FOTIP are consistent with those identified in the higher level "FOT Chain" that is described in Chapter 1 of this Handbook (see Figure 1.1), although the chronological order in which the Activities and Tasks are shown varies slightly between the two. For example, in the FOT Chain, in Figure 1.1, it is assumed that the first step when planning an FOT is the identification of systems and functions to be analysed. In the FOTIP, on the other hand, this task is identified later in the sequence of planning activities (within Activity 2), as there are other planning activities and tasks that necessarily precede the identification of systems and functions to be analyzed. The FOTIP identifies the scientific, technical, administrative and procedural activities for planning and running an FOT; the FOT chain summarizes the key, high level, scientific and technical steps undertaken when performing an FOT, and the sequential links between them.

2.2.3 Development of FOT Implementation Plan

The content of the FOT Implementation Plan derives from several research activities undertaken in Work Package 2.5 of the FESTA project:

- a comprehensive review of the literature on previous FOTs undertaken in different parts of the world: the United States and Canada; the Asia-Pacific region (including Australia and Japan); Europe; and Scandinavia. This included reference to FOT project plans, internal reports, meeting minutes and related documents, where possible. A special literature review of FOTs of nomadic devices was also undertaken, which encompassed all of these regions. References for the publicly available literature reviewed are listed at the end of the Handbook.

- a one-day workshop with FOT experts who had previously conducted FOTs, in Europe, the United States and Australia. This activity, along with the outputs of the literature reviews, identified critical Activities, Tasks and Sub-Tasks for successfully conducting FOTs, as well as the practical “dos” and “don’ts” of carrying out FOTs;

- an international teleconference with experts with experience in conducting FOTs and naturalistic driving studies. This augmented the information derived from the workshop;

- written feedback from FOT experts, who commented on an earlier draft of the FOT Implementation Plan; and

- internal consultation with other FESTA Work package leaders, to identify critical scientific, technical and administrative activities arising from other FESTA research activities undertaken in developing other chapters of the FESTA Handbook.

2.2.4 Assumptions underlying the FOT Implementation Plan

There is no one way of conducting a successful FOT. The review of the literature on FOTs revealed that many different approaches have been taken in planning, running, analysing and decommissioning FOTs. The FOTIP in Annex B of this Handbook draws together procedural activities that are most common to the known FOTs that have been conducted, and the collective wisdom of those who conducted them.

The FOT Implementation Plan is relevant to FOTs in which the ADAS and IVIS systems to be evaluated already exist as production systems in vehicles, or to studies in which the systems to be evaluated must be chosen by the FOT project team, purchased or developed, and installed (e.g. as in Regan et al., 2006).

The FOT Implementation Plan provides only a general guide to the sequence in which Activities, Tasks and Sub-Tasks should be performed. Some need to happen early in the project and others at the end. Some need to immediately precede others. Other tasks need to proceed concurrently with others. Decisions about the scheduling of Activities, Tasks and Sub-Tasks are the responsibility of the FOT Project Manager. Table 2.1 lists the 22 Activities identified in the FOTIP, and highlights the main dependencies that exist between them. Within Activities, it is up to the FOT Project manager to further decide which Tasks and Sub-Tasks should proceed sequentially and in parallel.

Some of the major Tasks listed in the FOTIP (e.g. “recruit participants”, within the Activity “Run FOT”) are given only a one-line description and, as such, may appear to be down played in the plan. A judgement had to be made about how much detail to include in the FOTIP. Where such one-liners exist, this is because either the Task in question is one that most researchers would normally be familiar with (e.g. recruiting study participants) or because the Sub-Tasks involved are described in detail in other Chapters of the FESTA Handbook. Where appropriate, any known difficulties and concerns associated with major Tasks for which only a one-line description is given are emphasised.

Table 2.1: A generic guide to scheduling the 22 Activities described in the FOTIP in Annex B of the 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				

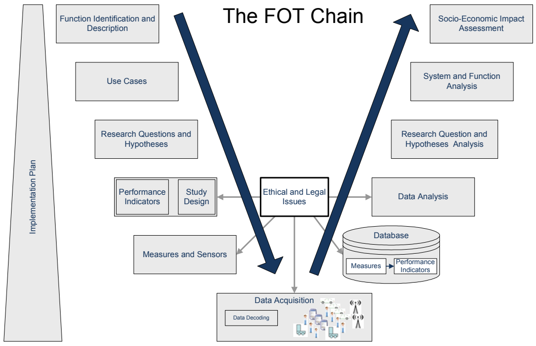
2.2.5 Using the FOT Implementation Plan

It is suggested that the FOTIP be used as follows:

- read through the FOTIP before starting to plan a FOT;
- use the FOTIP as a checklist for guiding the planning, design and running of the FOT — and as a quality control mechanism for ensuring during the study that nothing critical has been forgotten;
- read the FOTIP in conjunction with other chapters in the FESTA Handbook, and refer to other chapters and other FOT reports for detail; and

- if desired, use the FOTIP as the basis for the development of GANTT charts and other project management tools.

3 Legal and Ethical Issues



3.1 Introduction

Carrying out an FOT give 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 FOT, as these may be very dependent on the system(s) to be tested and 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 FOTs in more than one country or carrying out FOTs 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 an FOT can expose an organisation to criminal prosecution, e.g. for corporate manslaughter in the event that an unforeseen disaster occurs.

The differences in laws and regulations between the countries are not addressed here. As an example of what can arise on a national level, the view of one German lawyer with a high degree of knowledge in the area is included in Annex A for consideration.

In terms of the project timeline, legal and ethical issues need to be considered from the beginning to the end (and indeed afterwards in terms of data protection). So the discussion here does not neatly follow the FOT chain.

The FOT Implementation Plan, discussed in Chapter 2 and presented in table form in the Annex B, provides information about when in the FOT process the various legal and ethical issues need to be considered. The project plan needs to clearly identify who are the persons responsible for ensuring compliance.

3.2 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.

3.3 Participant agreement

There is a need to formalise 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 FOT, 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.

3.4 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 an FOT 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 recognised 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 man database and stored with additional protection such as encryption. It has to be recognised that, even when data has been anonymised, it may be possible to deduce who has participated, e.g. from GIS data in the database.

3.5 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.

3.6 System safety

It is obviously incumbent on those conducting an FOT 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.

3.7 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.

3.8 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 FOT as a whole. This may be provided by an employing organisation's professional indemnity insurance, but it is vital to confirm that the large risks are covered.

3.9 Video data collection

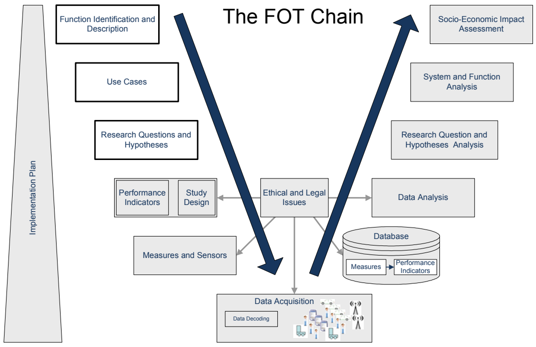
Video data collection within the vehicle has been covered in section 3.4. However, there are some additional points to consider. For example, there may be locations encountered where it is illegal or prohibited to video externally — border crossing, military locations, private premises. The possibility of this happening needs to be considered; it is likely to be more of a problem in truck FOTs.

External video may give rise to the same data protection issues as internal video. Many countries have regulations on the collection of outdoor video.

3.10 Ethical approval

Ethical approval to conduct an FOT 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 provide prior consent. The individual's protection and rights supersede any interests of scientific progress.

4 From Functions to Hypotheses



The final objective of an FOT is to evaluate in-vehicle functions based on Information Communication Technology (ICT) in order to address specific research questions. These research questions can be related to safety, environment, mobility, traffic efficiency, usage, and acceptance. By addressing the research questions, FOTs promise to furnish the major stakeholders (customers, public authorities, OEMs, suppliers, and the scientific community) with valuable information able to improve their policy-making and market strategies. Individuating the most relevant functions and connected hypothesis to successfully address the above-mentioned research questions is one of the major challenges in an FOT. In this Chapter, the process of individuating the functions to be tested in an FOT and the relevant connected hypotheses will be elucidated. Specifically, the reader will be guided in the process of 1) selecting the functions to be tested, 2) defining the connected use cases to test these functions, 3) identifying the research questions related to these use cases, 4) formulating the hypothesis associated to these research questions, and 5) linking these hypothesis to the correspondent performance indicators. The FOT chain shows specifically the steps reported above.

4.1 Systems and functions

In the last few years, the number of ICT functions available on standard vehicles has been rapidly increasing. ICT functions are intrinsically designed to provide the driver with new, additional information. However, the extent to which this increased amount of information from these ICT functions results in clear and positive effects on safety, environment, mobility, usage, and acceptance in real traffic situation is unknown. FOTs warrant to evaluate, for the first time, these ICT functions in a real traffic situation during naturalistic driving. In this handbook we refer to 1) in-vehicle, 2) cooperative, and 3) nomadic systems intended as a combination of hardware and software enabling one or more ICT functions. Depending on the different systems implementing a specific function, different challenges may have to be faced during the FOT design.

4.1.1 Vehicle systems

Vehicle Systems are a combination of hardware and software enabling one or more functions aimed at increasing driver's safety and comfort. Vehicle Systems promise 1) to increase driver safety by increasing driver's attention in potentially hazardous scenarios (such as the Forward Collision Warning), 2) to improve the driver's comfort by automating some of the operational driving tasks (such as the Adaptive Cruise Control function), 3) to increase driver mobility by furnishing timely traffic information (such as the Dynamic Navigation function), and 4) to increase safety in critical situation by automating the vehicle response (such as the Collision Mitigation function).

Vehicle Systems are becoming more and more standard equipment, even in middle class vehicles and commercial vehicles. However, their impact on the driver, the traffic system, the society, and the environment in the short-, but especially in the long term is not fully understood. FOTs can help quantifying the impact of Vehicle Systems on driver's workload and to understand how different functions interact with each other in a real complex traffic situation. Further, FOTs will expose these functions to many improbable scenarios which are not possible to be tested during the functions evaluation phase.

4.1.2 Cooperative systems

Cooperative Systems are vehicle systems based on vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I) and infrastructure-to-vehicle (I2V) communication technology. Communication technology has enabled a new class of in-vehicle information and warnings which are more precise in terms of time and location. Infrastructure based information tells the driver for example what is the appropriate speed to keep on a specific part of road or warns the driver in case of ice on the road or fog. Cooperative functions and systems have been developed in several projects under the 6th Framework Programme, namely CVIS, PReVENT, SAFESPOT and WATCHOVER.

One of the objectives of V2V, V2I, and I2V is to increase road safety. The development of safety-critical V2V systems in Europe has been mainly promoted by the Car-to-Car Communication Consortium (C2C-CC) which supports currently three 10MHz channels at 5.9 GHz according to the IEEE 802.11p standard. Further, a wider variety of information will be available with the CALM (Continuous Air interface for Long and Medium range communication) architecture. Due to the technical characteristics of Cooperative Systems, certain challenges have to be met while planning and performing related FOTs.

- Cooperative systems have some geographical limitations. High frequency communication is limited by the line of sight and this is limited by topography;
- How much radio propagation is dampened by atmosphere depends on the output power of the radios but bad weather might reduce the communication range a little;
- Communication will use its own communication congestion control, which is power control for reduction of range and avoiding redundant messages by intelligent protocols;
- Intelligent routing algorithms need to be developed to distribute messages in a certain geographical area or to dedicated vehicles;
- In the case of V2V applications, the end user can only benefit if there is a large number of similar-equipped vehicles. This is a burden on FOTs for cooperative systems. It will be necessary to establish well defined scenarios with application relevant situations and requires a minimum penetration rate.

It is most likely that the initial market drive will be for V2I-based applications until there is a critical mass of equipped cars on the road. FOTs can assess the technological and business feasibility of Cooperative Systems and may be necessary to complete the validation of such systems. In fact testing to validate Cooperative Systems may require very complicated set-up which may not be possible unless in a real-traffic set-up.

4.1.3 Nomadic devices

Nomadic Devices have become very popular in the last two years by the introduction of so called Personal Navigation Devices (PND) and Smart Phones (SP). They both run more or less autonomously on their own hardware platform. They are actually present in nearly every second vehicle and have become a quasi-standard device bought by the end-user.

One of the PND and SP critical characteristics is unequivocally the mounting in the car. Mostly they are mounted with a suction disc directly to the windscreen and therefore limit the drivers' field of view increasing the risk for accidents. Nevertheless, the popularity of these devices has not been abated. To limit the risk increase connected with the increasing popularity of PND and SP, the European Automobile Manufacturers' Association (ACEA) is trying to push within the ESOP (ESOP 06) a safer integration of PNDs and SPs.

The functionality of PNDs and SPs has been evolving over time. Initially, the PNDs objective has been limited to guide the driver from a geographic point A to a point B while, in the present time, the devices enable media player, media viewer, mobile phones, etc. and are in addition voice activated. SP were coming right from the other side with its main emphasis on telephony, supported by additional functions like address book and message sender and receiver. Additional functions (photograph, navigation, radio, media device centre etc.) were then continuously added.

Nomadic Devices are now available with telecommunication, Bluetooth, WiFi networks, assisted GPS, speech recognition and high-end operating systems with sufficient mass storage to host maps and related dynamic georeferenced applications. Results regarding the Nomadic Device integration in the vehicle environment through projects such as AIDE, GST and CVIS are already available. Further the upcoming FOT projects, such as TELEFOT or EUROFOT, will give perspectives and requests towards the methodological approach in FOTs to fulfil the user needs and safety considerations.

Usage aspects of Nomadic Devices from the driver behaviour and acceptance, environmental and traffic conditions need to be evaluated. The generation of indicators in the FESTA project aimed to compare on/off board solutions will help stakeholders to provide the drivers with the right nomadic services and the appropriate user interface.

In the light of the fast Nomadic Devices evolution, another main question that needs to be answered is how future integration of Nomadic Devices will look like. The end user is demanding more and more a seamless integration of external devices into the vehicular environment. For example, planning a route at home at the PC, downloading the data into the mobile device, approaching the vehicle and automatically loading the navigation system.

Due to the above-mentioned fast innovation cycle, the Nomadic Device FOTs will require a state of the art planning due to always new upcoming features and functions. This includes as well the consideration of the surrounding infrastructure since many functions rely on them. Weather forecast, traffic information, road conditions, speed advisory, etc… are all dependent on service providers.

4.2 General methodology

The main advantage of an FOT is that it has the potential to give insight in system performance in naturalistic driving situations, as free as possible from any artefact resulting from noticeable measurement equipment or observers in the car. Therefore the first step when planning an FOT is to identify systems and functions where considerable knowledge about their impacts and effects in realistic (driving) situations is of major interest, but is still lacking (see Section 4.2.1). Another domain for FOTs is the area of systems and functions which need a certain penetration rate to work at all, like especially cooperative system.

After the identification of the functions and system, which should be tested in an FOT, the goal is to define statistically testable hypotheses and find measurable indicators to test the hypotheses. To reach this goal, several steps need to be taken, starting from a description of the functions down to an adequate level of detail (see Section 4.2.1). This means that the main aspects of the functions, its intended benefits and the intrinsic limitations have to be described to fully understand objectives and limitations and to derive reasonable use cases.

Secondly, these use cases need to be defined (see Section 4.2.2). Use cases are a means to describe the boundary conditions under which a function is intended to be analysed. A general starting point is given by the functional specifications from the function description part. But it might also be of interest how a function performs when certain preconditions are not met and to identify unintended and unforeseen effects.

Starting from the use case definitions specific research questions need to be identified (see Section 4.2.3). Research questions are general question to be answered by compiling and testing related specific hypotheses. While research questions are phrased as real questions ending with a question mark, hypotheses are statements which can either be true or false. This will be tested by statistical means (see Chapter 9). One might already have a very clear idea from the beginning which hypotheses are to be tested in a very specific situation during the FOT. However, this very focused view might result in an extreme limited experimental design, where important unintended effects will not be considered. The process to define hypotheses developed in FESTA aims to prevent these potential issues (see Section 4.2.4).

Finally, hypotheses can only be tested by means of reasonable indicators (see Section 4.2.5).

These steps are shown as parts of the complete FOT and are elaborated further in the following sections. FESTA deliverables D3.1, D4.1 and D5.1 provide additional detail on the application of the FESTA methodology to identify functions and systems and to develop hypotheses for the experimental design. All steps, from the description of the systems and functions, the development of use cases and scenarios, as well as the research questions and hypotheses and the proposal of related performance indicators, have been accomplished.

4.2.1 Step 1: Selection and description of functions

Usually it is quite clear from the beginning what functions or at least what type of functions will be the object of an FOT. However, to select the specific functions but also in case the type of functions has not yet been decided, a Stakeholders Analysis is recommended. During this analysis, the needs of the different stakeholders need to be 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, like customers, public authorities, OEMs, suppliers, and the scientific community. It is of vital importance that all relevant stakeholders are included in the analysis to guarantee that the selection process will not itself bias from the beginning the appraisal of the gained results.

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 misjudgement.

The basis for all following steps is a sufficient description of the selected functions. For these purposes a spreadsheet template has been prepared and is presented in the Annex[2] to collect the necessary information. It provides two main parts: First, the functional classification, where a short high level description of the main aspects of the function should be given. This information is usually provided through the system specifications given by the system vendor or OEM. The second part of the description comprises of limitations, boundary conditions and additional information which is necessary to understand how the function works.

The boundary conditions part describes where and under which circumstances the system/ function will operate according to its specifications, where the FOT should take place and which type of data needs to be recorded during the FOT to enable a good interpretation of the results. It consists of:

- Infrastructure requirements, cooperative systems and nomadic device requirements. Here all required actors besides the actual system need to mentioned, which might have an impact on system performance, service availability or similar. It is intended to trigger the consideration of factors which are external to the system/ function under evaluation;
- Demographical Requirements/ Driver Requirements: Especially the user or driver recruitment needs to take into account, whether a function is particularly designed for a specific group of users or drivers. Drivers differ on a large variety of characteristics, which may all have an influence on how they drive and use different systems and services. These differences may be important to take into account when planning a FOT. Four categories of driver characteristics may be distinguished:
    - 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.

- Geographical Requirements/ Road Context: This description is necessary for systems which, concerning their functionality, depend strongly on the horizontal or vertical curves of the road layout or on the road type. For example, certain speed limit information systems depend largely on the availability of speed limit information in a digital map, which is up to now only commercially available on high class roads.
- Geographical Requirements/ environmental restrictions: Certain systems are especially designed for specific environmental conditions or, on the other hand, specifications might indicate that the system under evaluation will not work under certain environmental conditions. In this case the location of the FOT needs to be selected carefully and the relevant data must be recorded during the FOT. e.g. most of the functions using perception system will be affected by adverse weather conditions. If this is the case it is necessary to log respective data and take it into account for later data analysis.
- Geographical Requirements/ Traffic Context: The performance of certain systems might depend on the traffic context, that is, the traffic density (e.g. given by the Level of Service) or might even be designed to work in specific traffic densities only. Like the other geographical requirements, this needs to be taken into account when an FOT is planned, performed and the data is analysed.
- Other Limitations: All other limitations need to be mentioned, which might have considerable impact on the performance of functions or systems, since these limitations have major impact on the experimental design and data analysis.

4.2.2 Step 2: Definition of use cases and situations

FOTs will test technically mature ICT systems. Therefore, systems and functions to be tested are on the market or close to market and can be easily implemented. But the list grows too long if all possible implementation variations and technologies are considered separately. The use cases are putting the systems and functions at a suitable level of abstraction in order to group technology-independent functionalities and answer more holistic research questions described later.

Table 4.1: Use Cases, Situations, Scenarios, and their mutual dependence.

Subject	Definition	Comment	Example
Use Case	Target condition in which a system is expected to behave according to a specified function	A use case is a system and driver state, where “system” includes the road and traffic environment.	Car following
Situation	A combination of certain characteristics of a use case. Situations can be derived from use cases compiling a reasonable permutation of the use cases characteristics.	Thus a situation is a state of the environment or system.	Speed above 70Km/h + sunny day + FCW on
Scenario	A use case in a specific situation	Use case + situation = scenario	Car following with speed above 70Km/h + sunny day + FCW on

A use case is a textual presentation or a story about the usage of the system told from an end user’s perspective. Jacobson et al (1995) defined the use cases: “When a user uses the system, she or he will perform a behaviourally related sequence of transactions in a dialogue with the system. We call such a special sequence a use case.” Use cases are technology-independent and the implementation of the system is not described. Use cases provide a tool for people with different background (e.g. software developers and non-technology oriented people) to communicate with each other. Use cases form the basic test case set for the system testing. There are number of different ways to define a use case. Use cases in FESTA are very general descriptions, like e.g. “car following”. This general description needs to be refined to a reasonable level of detail. This refinement is done by describing so called situations (see Table 4.1). It is the detailed scenario description which triggers the development of specific hypotheses for later analysis.

The situational descriptors are selected in a way that relevant information can be gathered to distinguish between main differences while evaluating systems. The situational descriptors can be distinguished in static and dynamic, while the static describe attributes which will not change significantly during one ride of the vehicle, like age or gender of the driver. Nevertheless this information needs to be stated, since it is one of the main inputs to filter the huge amounts of data in the later stage of data analysis. The second part of attributes is dynamic, since it can change during a ride of the vehicle, like the system action status (system on or off), the traffic conditions, road characteristics or the environmental situation.

The situations are defined as a combination of certain characteristics of a use case. Situations can be derived from use cases compiling a reasonable permutation of the use cases characteristics. The identification of possible situations was covered from three viewpoints:

1. systems and vehicle specification,
2. environmental conditions specification and
3. driver characteristics and status specification.

The situational descriptors in FESTA conforms the following structure:

IDENTIFICATION AND DESCRIPTION

Use case name	A name for identification purposes.
Description	General description of the use cases with necessary depth of information to get a quick overview.
Occurrence	Information about the anticipated quantity of occurrences has implications for the amount of data to analyse.

SYSTEMS AND VEHICLES

System status	Depending on the hypotheses the analysis might concentrate on situations where the system is activated or present. *Example: ON/OFF (baseline) or IDLE/ON/OFF*
System action status	Depending on the hypotheses the analysis might want to compare the driving performance between different system statuses, e.g. whether the system is actively controlling the vehicle or not. *Example: acting/ not acting (meaning e.g. ACC controlling car speed or not)*
System/ function characteristics	Depending on the hypotheses, an analysis of system or driver performance with respect to special system/ function characteristics might be conducted, e.g. examining differences in system performance between nomadic devices (phone, Smartphone, PND...) or effects that depend on vehicle type. *Example: passenger vehicle/ truck/ bus*
Interaction between systems	System and especially driver behaviour might change depending on whether the system under evaluation is the only active support system or whether interactions between two or more systems are foreseen. *Example: interaction between Blind Spot Warning and Lane Departure Warning.*

ENVIRONMENTAL CONDITIONS

Traffic conditions	Performance of some systems might differ depending on traffic density. Others might only be reasonable with a minimal traffic density. *Example: Level of Service A and B*
Environmental situation	System performance differs depending on lighting and weather conditions like rain/ snowfall/ icy roads, etc. *Example: normal/ adverse weather conditions*

Road characteristics	e.g. type of road, gradient, super-elevation, curvature, curviness, since some systems are dedicated to improve driving performance in curves etc. *Example: urban roads/ rural roads/ highways*
Geographical characteristics	Information about geographical characteristics relevant for testing the systems. *Example: mountainous/ flat areas, metropolises with high street canyons.*

DRIVER CHARACTERISTICS AND STATUS

Driver specification	Characteristics of the users have an impact on the driving performance. 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. Information about driving experience is also important. 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 (T-LOC) test. 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. When evaluating the acceptance and use of new systems in the car, drivers' acceptability of technology is important. Both social and practical aspects play a role. Technology acceptance has different dimensions, such as diffusion of technology in the drivers' reference group, the intention of using the technology, and the context of use (both personal and interpersonal). Measuring acceptability can be realized via (existing) standardized questionnaires, in-depth interviews before and after "use" (driving), and focus groups.
Driver status	Mindset of the driver *Example: attentive/ distracted/ impaired*

Purpose, distance, duration	Describes the different attributes of a trip (time between ignition on and ignition off). All three aspects have an impact on driver behaviour and hence on patterns in the data.

A set of basic rules has been set for the design of the situations for an FOT:

1. Complementary: situations are not allowed to overlap.
2. Entirety: the sum of all situations should describe the complete use case.
3. Baseline: The same situation without the use of the systems (system off or non-present) is defined as the baseline. The baseline is the basis for the benefit assessment of the system and the comparison between systems. Therefore, for the same use case, there can be many baselines depending on the number of situations.
4. Comparability: functions compared in an FOT need to have the same use case and therefore same baseline and situations.
5. Variability of situation parameters: depending on the point of view (user, trip, vehicle, single FOT, multiple FOTs, etc…), attributes describing a situation can vary considerably or not.

This list is non-exhaustive and might be extended if necessary.

Finally, out of all the possible situations, one will need to select the relevant ones for scenarios of interest in an FOT. The scenarios are defined as a use case in a specific situation and therefore one or more scenarios should be considered from each use case. All other situations should be considered out of the scope of the FOT study. However, if possible data should still be collected in all situations in case an alternative study would like to reuse the same data.

During FESTA a list of functions and use cases was produced based on technically mature ICT systems and functions on the market. The list was consolidated based on the feedback from a stakeholders workshop and a dedicated questionnaire.

The process of defining the use cases will help the FOT for the next steps: the definition of the research questions and hypotheses and finally the identification of the needed indicators. The scenarios as they are defined at this stage of the FOT are not detailed enough for data analysis purposes. For this reason, after the definition of the indicators, the scenarios (and their situations) will need to be further described in terms of *events* for data analysis purposes. Only then, the scenarios can be classified with a quantitative measurement tools in function of the defined indicators.

4.2.3 Step 3: Identification of the research questions

The research questions specific to an FOT can only be identified once the overall goal of an FOT has been established.

In general terms the goal of any FOT is to investigate the impacts of mature ICT technologies in real use. The core Research Questions should therefore focus on <u>impacts</u> but there are other questions that 'surround' this core. The range of possible questions is listed below. This list below should be considered a first step in any FOT and not a comprehensive set of questions.

LEVEL OF SYSTEM USAGE

Which factors affect usage of the functions? Examples are

- Purpose of journeys where system is used
- Familiarity with routes where system is used
- Portion of journey for which system is used
- Types of road on which system is used
- Traffic density
- Headway
- Weather condition
- Ambient lighting

How do driver characteristics affect usage of the functions? Examples are

- Personal characteristics (e.g. age, vision)
- Socio-economic characteristics (e.g. family, friends, employment status)
- Journey-related characteristics (e.g. other car occupants, shared driving)

IMPACTS OF SYSTEM USAGE

What are the impacts on safety?

- exposure
- risk of accident or injury
- incidents and near accidents
- accidents?

What are the impacts on personal mobility?

- individual driving behaviour
- travel behaviour
- Comfort

What are the impacts on traffic efficiency?

- traffic flow (speed, travel time, punctuality)
- traffic volume
- Accessibility

What are the impacts on the environment?

- CO_2 emissions
- Particles
- Noise

IMPLICATIONS OF MEASURED IMPACTS

What are the implications for policy?

- Policy decisions
- Laws, directives and enforcement
- Future funding

- Public authority implications
- Emergency service implications

What are the implications for business models?

- Predictions for system uptake
- User expectations
- Pricing models

What are the implications for system design and development?

- HMI design and usability
- Perceived value of service
- Device design
- Communications networks
- Interoperability issues

What are the implications for the public

- Public information/education
- Changes in legislation
- Inclusive access to systems
- Data protection

4.2.4 Step 4: Creation of hypotheses

Once the key research questions for the FOT have been identified, hypotheses can be derived. The process of formulating hypotheses translates the general research questions into more specific and statistically testable hypotheses.

There is no process that can assure that all the “correct” hypotheses are formulated. To a large extent, creating hypotheses is an intuitive process, in which a combination of knowledge and judgement is applied. Nevertheless, a number of recommendations can be made about how this process should be conducted. These recommendations have been tested in a FESTA workshop and modified based on the experience of and feedback from that workshop.

Two complementary ways to develop hypotheses have been used. Both ways need to be followed, while it is not of importance which step is taken first. One of the steps follows the sequential check of specific areas in which functions can have an impact; the other step is fully based on the description of specific scenarios. While the one step results mainly in general hypotheses, the other step triggers the development of very specific hypotheses in specific driving situations or scenarios.

Deriving hypotheses from the scenarios

The main reasoning to describe functions, their use cases, situations and scenarios in detail according to Steps 1 and 2 is to trigger the generation of hypotheses for very specific scenarios. The hypotheses generation should be conducted by a team of experts, consisting of human factors experts, development engineers and traffic engineers and all of them need to fully understand the functions/ systems with all aspects and limitations.

Scenarios should be covered systematically. It is recommended that a structured approach be used and that the situations are checked sequentially for related hypotheses.

The six areas of impact

The six areas of impact defined by FESTA are based on Draskóczy et al. (1998). Although this approach was originally designed for formulating hypotheses on traffic safety impacts, it is in fact equally applicable for efficiency and environmental impacts.

The six areas are:

- Direct effects of a system on the user and driving.
- Indirect (behavioural adaptation) effects of the system on the user.
- Indirect (behavioural adaptation) effects of the system on the non-user (imitating effect).
- Modification of interaction between users and non-users (including vulnerable road users).
- Modifying accident consequences (e.g. by improving rescue, etc. — note that this can effect efficiency and environment as well as safety).
- Effects of combination with other systems.

It is not of particular importance to which of these areas a particular hypotheses is allocated. The six areas are instead to be used as a checklist to ensure consideration of multiple aspects of system impact.

In applying this procedure, it should be noted that:

- Area 1 includes the human-machine interaction aspects of system use.
- The driving task (see Table 4.2) can be defined, following Michon (1985) into the three levels of strategic, tactical and control (operational) aspects.
- Consideration should be given to such mediating factors as user/driver state, experience, journey purpose, etc.

It should also be noted that the effects of system use may be:

- Short-term or long-term in terms of duration and
- Intended or unintended in terms of system design.

This additional step for hypotheses generation assures that very general hypotheses are not forgotten as well as hypotheses on unintended, short term and long term effects. It is intended to serve as a means for crosschecking.

Table 4.2: Levels of the Driving Task by Michon (1985)

Level	Explanation/ example
Strategic	Finding the way through a road network (navigation) including • Modifying modal choice • Modifying route choice • Modifying exposure (frequency and/or length of travel)
Tactical	e.g. changing lanes, keeping the vehicle on the lanes, including modifying speed choice
Control/ Operational	Maintaining speed/ headway and distance to other vehicles

Prioritising the hypotheses

The prioritization among the generated hypotheses is a difficult process. No specific advice can be given on how to proceed, but there are some general guidelines:

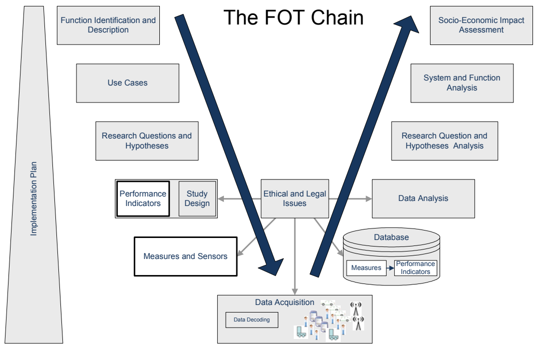
A complete list of the hypotheses that have been developed should be recorded. If it is considered that some are too trivial or too expensive to address in the subsequent study design and data collection, the reasons for not covering them should be recorded. In general, it should be left to the judgement of the experts acting as hypotheses generators which hypotheses are likely to reflect the real driving situation. Those should then be prioritized, keeping in mind that also unintended effects are very important.

4.2.5 Step 5: Link hypotheses with indicators for quantitative analyses

Some of the hypotheses will already incorporate an indicator which needs to be measured, e.g. a very concrete hypothesis like “The function will increase time-to-collision (TTC)”. In this case it is obvious which indicator to choose, while the method to measure TTC might include complicated procedures and/ or costly measurement equipment. Chapter 5 gives an overview about many reasonable indicators. One should consider these indicators when planning the experimental design, since a detailed description how to calculate the indicators from measurements is also provided.

Other hypotheses might be rather unspecific, but still reasonable after rephrasing into testable ones. This rephrasing goes hand in hand with the identification of related reasonable indicators. For example, a hypothesis like “The function will increase lane changing performance” is not directly testable, since “lane change performance” is not an indicator itself. Hence, surrogate measures must be identified to evaluate lane change performance. These surrogate measures or indicators can e.g. be found in publications of corresponding research projects. If appropriate information cannot be found or is not accessible, new performance indicators need to be developed. Those indicators and the measurement methodology must be valid, reliable and sensitive, that is, the measurement must actually measure what it is supposed to measure, they must be reproducible and the measurands must be sensitive to changes of the variable. A sensitivity analysis should be performed beforehand during a pilot study to make sure that the new performance indicator is suitable. When one or more surrogate measures have been identified, the initial hypothesis can be reformulated into one or more testable hypotheses. In the above mentioned example, reasonable indicators associated to “lane change performance” might be: use of turning indicator or the number of lane change warnings. The initial hypothesis will then be reformulated into: “The system will increase the use of the turning indicator.” and “During the system use, the number of lane departure warnings will decrease.”. The next step is then to evaluate how the indicators “use of turning indicator” and “lane departure warnings” can be measured. In this context, Chapter 5 provides useful information.

5 Performance Indicators



5.1 Introduction

During the process of developing hypotheses, it is important to choose appropriate Performance Indicators (PIs) that will allow to answer the hypotheses, but that will also be obtainable within the budget and other limitations of the project. Many different kinds of PIs have been used in previous studies, and they are related to various aspects of driving. Below a definition and description of PI is given. Further, it is explained how the PI is related to measures, and the types of different measures that have been identified are described. In this chapter, examples are provided to illustrate the concepts. An overview of the PI-Measures-Sensors-table, that can be found in the annex of FESTA Deliverable 2.1., is given, and background text related to the different groups of PIs and measures is provided.

5.2 Performance indicators definition

Definition: PIs are quantitative or qualitative measurements, agreed on beforehand, expressed as a percentage, index, rate or other value, which is monitored at regular or irregular intervals and can be compared with one or more criteria.

Further explanations:

- Hypotheses steer the selection of PIs and the criteria against which those should be compared. Hypotheses are seen as questions that can be answered with the help of measurable PI.
- Criteria can be baseline, different experimental conditions, absolute values, etc. This depends on the research questions and hypotheses.
- New PI or combinations can be developed during the course of the study. They will have to be validated in follow-up studies.
- A denominator is necessary for a PI. A denominator makes a measure comparable (per time interval/per distance/in a certain location/…). Therefore “crash” or “near-crash” in themselves should rather be considered to be “events”, because they become comparable only when they get a denominator, like “number of crashes per year per 100.000 inhabitants”. For certain PI either time or distance can be used in the denominator (e.g. number of overtaking manoeuvres, percentage of exceeding the posted speed limit).

For PIs measured via rating scales and questionnaires, focus groups, interviews, etc., the “denominator” would be the time and circumstances of administrating the measuring instruments, for example before the test, after having experienced the system, and so on.

PIs are very diverse in nature. There are global PIs as well as detailed PI, there are observed and self-reported PIs, there are PIs calculated from continuous and from discrete data, and so on. An example for a rather global PI based on continuous log data would be the mean speed on motorways, whereas an example for a PI based on discrete, self-reported data would be the level of perceived usability of a function. Some PIs can be based on either self-reported, discrete measures or on logged data, such as for example the rate of use of a system. The participants can be asked how often they use a function, but the actual function activation and the different settings chosen by the driver can also be logged from the system.

All PIs are based on measures, which are combined and/or aggregated in certain ways, and which are normalised in order to allow comparisons. The measures are described below.

5.3 Measures

Five different types of measures were identified, namely Direct Measures, Indirect Measures, Events, Self-Reported Measures, and Situational Variables, which are described in more detail below. A measure does not have a “denominator”. Therefore it is not in itself comparable to other instances of the same measure or to external criteria. The measure itself, however, can very well be a fraction (like speed). Several PIs can use the same measures as input, and the same measures can be derived from different types of sensors. An example would be speed that can be read from the CAN bus, logged from a GPS receiver, or calculated by an external sensor registering wheel rotations.

5.3.1 Direct (raw) measures

A Direct Measure is logged directly from a sensor, without any processing before saving the data to the log file. Linear transformations like the conversion from m/s to km/h are not considered to be processing. How the sensor arrives at its output is not relevant for the classification. Longitudinal acceleration, for example, is a Direct Measure if logged directly from an accelerometer, but not if derived from the speed and time log. In this case it would be a Derived or Pre-Processed Measure, because it is not directly available from a sensor and has to be calculated from other measures, i.e. pre-processed, before logging. Further examples of Direct Measures are: raw eye movement data, the distance to the lead vehicle as measured by radar and a video film of the forward scene.

5.3.2 Derived (pre-processed) measures

A Pre-Processed Measure is not directly logged from a sensor, but either a variable that has been filtered, for example, or which is a combination of two or several Direct or other Pre-Processed Measures. An example for a Pre-Processed Measure is time to collision (TTC), which is based on the distance between a vehicle and another vehicle or object, divided by the speed difference between the two vehicles or the vehicle and the object. The distance to the vehicle or object on collision course is a Direct Measure from a radar, for example. The speed difference between the own vehicle and the other vehicle or object is another Pre-Processed Measure, based on the own speed as read from the CAN bus, for example, and the calculated speed of the other vehicle or object. Further examples of Pre-Processed Measure based on raw eye movement data and the layout of the vehicle are: pre-defined zones that the driver looks at, like for example the mirror, the windscreen, and the radio.

A special case of Derived Measures are those that are coded by a human observer when data logging is done. Examples might be reduced eye movements, classifications of scenarios or classifications of secondary task engagements. These Measures are considered to be “derived”, because data reduction by a human observer is more than only a linear transformation, and they can be based on more than one Direct Measure. In case of secondary task classification one might use both a video of the driver’s hands and a log file of an eye tracker, and for scenario classification both a road database and a video of the forward view might be used.

5.3.3 Events

Events can be seen as singularities based on Direct Measures and/or Derived Measures or on a combination of those. They can be short in time, like a crash, or extended over a longer period of time, like an overtaking manoeuvre. One or several preconditions must be fulfilled for an Event to be classified as such, that is, one or several “trigger” criteria must be exceeded. For the Event “overtaking manoeuvre”, for example, the non-technical definition might be: A vehicle in a vehicle-following situation changes lanes, accelerates and passes the vehicle in front, then changes lanes back into the lane, in front of the vehicle(s) that have been overtaken. Depending on the infrastructure design, the definition might need to be extended to motorways with more than two lanes in each direction, for example. For a more technical definition that sets the trigger criteria of when exactly an overtaking manoeuvre starts and when it ends, either the literature has to be consulted or an own definition has to be developed. This can possibly be based on previous data, or, if nothing else is available, on the data from the current FOT.

Another example of an Event, based on TTC and possibly other measures like a film of the driving scene or steering wheel angle, is a near miss, where the TTC has to be below a certain trigger value in order for the episode to be considered a near miss Event.

FESTA, however, will not provide trigger values for Events, and neither will the exact measures that have to be included for the definition of a certain Event be provided. The Events listed in the matrix should be seen as examples.

Several PIs can be related to one Event type, for example for overtaking manoeuvre it could be of interest to determine the number of overtakings, the duration of overtaking, the distance/time spent in opposite lane, and so on. For near misses, the number of such Events per distance, time or capita could be counted, and it could be split further into different traffic environments, for example the rate of near misses on motorways, in urban areas, etc. These examples illustrate that PIs can be built by counting Events, or by considering certain aspects of those Events.

5.3.4 Self-reported measures

A number of PIs are based on Self-Reported Measures, which are gleaned from either questionnaires, rating scales, interviews, focus-groups, or other methods requiring introspection on the part of the participant. These measures are typically not logged continuously, but rather only once or a few times during the course of one study. The measures related to Self-Reported PIs could be the answers to each single question or the checks on the rating scales, while the sensors would be the questionnaires or rating scales themselves. It is more difficult to make a meaningful distinction between measure and sensor for semi- and unstructured interviews and especially for focus groups.

In the matrix only a small number of Self-Reported Measures are included, which are those that are necessary for the computation of a PI that is not solely based on self-reported measures, like for example “deviation from intended lane” or “rate of errors”.

5.3.5 Situational variables

Situational Variables can be logged like Direct Measures or computed like Derived Measures., They can also be self-reported and they can correspond to Events. Their commonality is that they can be used as a differentiation basis for other PIs, in order to allow for a more detailed analysis. It might, for example, be of interest to compare certain PIs in different weather or lighting conditions, on different road types, or for different friction conditions. These Situational Variables are included in the PI matrix in the measures table, but they are not linked to any specific PI. In principle, all kinds of measures can be used as Situational Variables, such as when analyses are performed for different speed intervals.

5.4 The PI-Measures-Sensors matrix

A matrix was developed that in one table contains PIs covering different aspects of research questions that might be addressed in an FOT (see the annex of FESTA

Deliverable 2.1). These PIs are described with respect to different categories. For each PI, the measures on which it is based are listed.

All these measures are then described in another table. Different categories are provided for description, where some are reserved for Direct Measures, others for Derived Measures and for Events. Each Direct Measure points to a sensor from which the measure can be read. As mentioned above, for certain measures different sensors can be used. In this case, each of those is described as a separate measure.

A link is made between the PIs and the measures table by indicating for each PI which measure is needed to compute it. In this way, when the hypotheses have been generated, it should be possible to pick the appropriate PI and from there proceed via the pointers to the necessary measures and from there to the sensors. If several sensors can provide the same measures choices can be made due to budget limitations, sensor limitations or other restrictions.

Presently most measures for Self-Reported PI are not included in the matrix. Instead, a direct reference is made to the appropriate questionnaire, rating scale or method needed to obtain this PI. For correct deployment of the recommended method, the user is directed to the instructions for this particular method.

Measures that describe driver characteristics are not included in the matrix itself, but in the annex to the matrix. In this annex, it is explained which instruments could be used to assess different aspects of driver characteristics (FESTA D2.1). The characteristics covered in this document are usually stable over a longer period of time.

This matrix is not meant to be exhaustive; it is only an aid for selecting PIs, measures and sensors. It should by no means be regarded as being limited to the PIs or measures entered now, and users are encouraged to expand the matrix during the course of the FOTs. Further instructions on how to work with the matrix are provided in FESTA Deliverable D2.1.

5.5 Background information from tasks

The PIs are split into different sub-groups, depending on which area of the traffic system they are concerned with.

5.5.1 Indicators of driving performance

Driving performance is discussed and analysed in relation to traffic safety. Given that the accidents are usually multicausal, a set of indicators should cover a number of factors. Otherwise any FOT is likely to miss essential information that is required to produce reliable and valid results.

Traffic safety is regarded as a multiplication of three orthogonal factors, namely exposure, accident risk and injury risk (Nilsson, 2004). The driver decision making and behaviour covers all these aspects. Typically, strategic decisions are highly relevant for exposure, tactical decisions for the risk of a collision, and operational decisions for the risk of injuries (Michon, 1985). Consequently, an FOT should cover all these aspects, because it is essential to cover the driver tasks and driver behaviour widely, and include decisions like whether to use the vehicle at all, route planning before the trip, timing of the trip etc. However, the focus is on driving performance while driving a vehicle. For example, the traveller behaviour in public transport is excluded after the decision to use other modes than passenger cars. Relevant aspects include interaction with other road users, use of controls, use of IT systems, and other activities while driving. In addition, driving conditions should also be taken into consideration.

Another approach to traffic safety is to investigate driver behaviour in terms of how close to an accident the behaviour is: normal driving, incident, conflict, near-crash, or crash. Although crashes may not be regarded at a first sight as driver behaviour, we suggest that road crashes will be included as events because they provide an ultimate measure for road safety. In the wide-scale field experiments even this direct criterion of safety may be relevant. It is more self evident that near-crashes will be included.

The *events* such as crashes and near-crashes indicate a lack of safety rather than safety, and the interpretation is that traffic is safe in the absence of these phenomena. Most of the indicators are derived from situations involving lack of safety. An *indicator of driving performance* is a behavioural variable which *indicates* the quality of the behaviour in respect to road safety. The behaviour is measured directly from driver (e.g. frequency of glance to given object) or indirectly from the vehicle (e.g. speed).

5.5.2 Indicators of system performance and influence on driver behaviour

In this task indicators were developed that describe the actual performance of the system to be tested. These indicators are mostly related to both *safety* and *acceptability*. Here the focus is directed at the question whether the system actually functions the way it is meant to under realistic conditions. False alarms and misses could be obvious indicators of that. Relations exist with indicators of acceptance and trust, which examine the subjective opinion of the participants on how the system worked.

Furthermore, indicators that describe the influence of the system on the driver and the interaction between system and driver are described. They will enable assessing the driver's willingness to use the system in various situational contexts. They will also contribute to the identification of potential misuses of the system leading to incidents or conflicts. In a longitudinal perspective, they will also contribute to an analysis of the learning and appropriation phases.

The intrinsic performance of the system

The first issue is the intrinsic performance of the system studied. It is related to the *precision and the reliability of the system*. Does the system perform as expected? In this case we need indicators signalling any deviations, such as false alarms and misses, but also indicators about the context in which these deviations occur. Ideally, the origin of the deviation should also be identified. The identification of false alarms or misses may be based on automated sensors or may require a video recording of the driving scene. For example, in the French LAVIA (ISA) project, loss of the recommended or target speed were automatically recorded while mismatching between the target speed and the posted speed limit was identified on the basis of a video recording of the driving scene.

The intrinsic performance of the system should be distinguished from the operational envelope of the system (i.e. the use cases for which the system was designed to work). This is important when assessing the opinion on the performance of the system: when asking the driver to assess the system performance, the limits of the system operation should be differentiated from system deviations. Two main indicators related to the operational envelope are: 1) availability of the system over driving time (Percentage of the driving time the system is available, e.g. some system are only available above a certain speed, for special road characteristics, etc.); and 2) frequency of take-over requests (the system is active but not able to provide assistance due to system limits, e.g. for ACC maximum brake rate is limited).

Both intrinsic performance and competence envelope are assumed to play a role for the drivers' opinion on the system.

Modes of drivers' interaction with the system

The second issue is the driver's interaction with the system. This goes beyond the analysis of overall driving performance when using support systems: in fact 1) it is examined how drivers use and interact with the system; and 2) it is examined how this interaction may affect driving behaviour and performance.

How drivers use and interact with the system

Some support systems require/enable the driver to activate/deactivate the system, to override the system, to select one system among other systems available, to select or to register some vehicle-following or speed thresholds, and so on. In other words, using a system implies the application of a number of procedures, and these procedures should be registered and analysed. This is the case for systems such as speed limiters, cruise control, adaptive cruise control or navigation systems for example. These procedures may be classified as the driver's direct or indirect interventions, depending on whether they are applied through vehicle controls (brake or accelerator) or through system controls. As for the indicators of system performance the situational context should be taken into account. This is important for identifying potential misuses of the system leading to incidents or conflicts as described above. In a longitudinal perspective, these indicators will also contribute to an analysis of the evolution of system usage from the learning and appropriation phases to the integration phase. Furthermore, the frequency with which the system "interferes" with the driver's activity has to be assessed. For example, when driving with a speed limiter, how often is the system "active", that is, effectively limiting the vehicle speed.

How this interaction may affect driving behaviour and performance

For analysing the effect of the driver's interaction with the system on driving behaviour and performance various levels of analysis could be employed, depending on the desired level of granularity of analysis. Obviously, this granularity depends on the recording means available as well as on the time required for performing such analyses. For example, studying changes in glance behaviour requires video recordings and is time consuming.

For an analysis of behavioural changes at a more global level, synthetic indicators should be conceived. These indicators are assumed to reflect changes at the tactical or strategic level of the driving task. Indicators such as "lane occupancy" and "frequency of lane change" are often used to assess changes at the tactical level. Changes at the strategic level could be reflected by changes in the itinerary chosen or changes in driving time.

Recommendations:

Classify the support systems by type and level of interaction implied by their use;

Classify the indicators according to the level of granularity of analysis that they permit;

Classify the indicators according to the means and time required for collecting and analysing them.

5.5.3 Indicators of environmental aspects

Exhaust emissions include many different substances like: HC, CO, NO_x, PM, CO_2, CH_4, NMHC, Pb, SO_2, N_2O and NH_3. Greenhouse gases – CO_2, CH_4 and N_2O – represent the same society cost anywhere, while costs for other substances depend on the geographical position.

There are two alternatives for quantifying exhaust emissions: measured exhaust emissions or calculated. For measurements there are still two alternatives: on board or in the laboratory. The laboratory alternative demands use of logged driving patterns. Because of the high complexity and costs for measurements of exhaust emissions, in practice, calculated emissions is in most cases the only reasonable alternative.

Models for exhaust emissions in general include three parts: cold start emissions; hot engine emissions and evaporative emissions. The following formula is a rough description of an exhaust emission model:

Σ(Traffic activity) x (Emission factor)=Total emissions

Traffic activity data include at least: *mileage* and *engine starts*. Hot emission factors for one vehicle are functions of the driving pattern and vehicle parameters. Cold start emission factors are functions of the engine start temperature, trip length and average speed. Evaporative emissions are to a large extent a function of fuel quality and fuel tank temperature variations.

Models on a micro level, including engine simulation, should in principle be able to describe most ICT functions. This is not the case for models on a macro level in general. Micro models are often used for emission factor estimation and macro models for total emission estimations.

The conclusion about what to include as PIs would then be: exhaust emissions or measures with high correlation to exhaust emissions.

5.5.4 Indicators of traffic efficiency

The efficiency of a traffic system can be measured as, for example, traffic flow, speed and density in relation to the optimum levels of these properties given the traffic demand and the physical properties of the road network.

A combination of FOTs and traffic modelling is required to allow estimation of traffic efficiency impacts of the tested technologies. A schematic picture of the proposed methodology is shown in Figure 5.1.

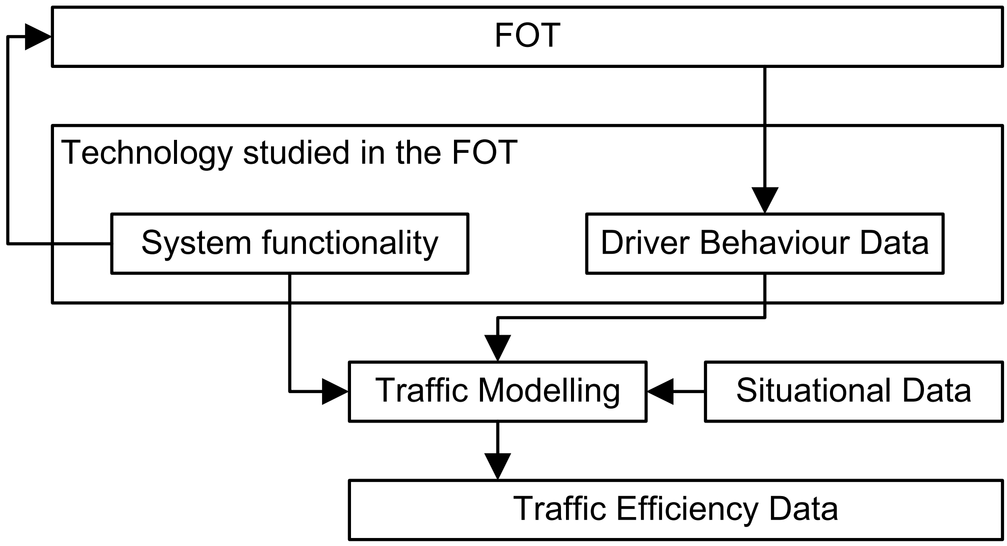


Figure 5.1: FESTA Traffic efficiency estimation based on FOT results

Driver behaviour data are based on the data collected in the FOT. These driver behaviour data will, together with the system functionality[1] of the tested technology, be used as input to traffic modelling in order to aggregate the individual driver/vehicle impact on traffic efficiency effects. This requires that both driver/vehicle data of equipped vehicles and properties of the traffic system that the vehicles have driven in (henceforth referred to as Situational Variables[2]) are collected in the FOT.

[1] System functionality refers to the way in which the tested FOT system works. Information on when and how the system operates can be used to create parameters for the models developed.

[2] Situational Variables are not necessarily directly relevant for Performance Indicators or Derived Measures, but must also be measured or recorded as they provide key background information that complements the driver behaviour data and is sometimes needed to derive the driver behaviour data. Examples include light conditions and road type.

The driver behaviour data required in order to estimate traffic efficiency for any type of FOT system are specified in terms of PIs and Measures and included in the attached matrix. These data (along with the Situational Variables, which can be found in the Measures Table in the Annex of Deliverable D2.1 should be ascertained for the baseline case (non-equipped vehicle) and for the equipped vehicles, so that comparisons can be made between the two.

The appropriate traffic modelling approach will differ depending on which type of driving tasks supported by the considered technology. Michon's (1985) hierarchical driving model can be applied to select a traffic modelling approach. To model systems that support tactical or operational driving tasks, it is appropriate to apply a traffic microsimulation model. A microsimulation model considers individual vehicles in the traffic stream and models vehicle-vehicle interactions and vehicle-infrastructure interactions. To model systems that support strategic and some types of tactical driving tasks it is appropriate to apply a traffic simulation model. A mesoscopic model considers individual vehicles but model their movements and interactions with a lower level of detail than microscopic models.

It is advisable to study traffic efficiency for a series of scenarios with varying levels of traffic penetration of the tested systems. The systems should also be studied in representative traffic volumes. This is achieved straightforwardly by running the traffic simulation model with different inputs. The situational data will also contribute to the differences between the scenarios (both measured and modelled).

Outputs from the traffic models will be used to make comparisons of traffic efficiency for the studied scenarios. Example outputs of interest are traditional quality of service and traffic efficiency indicators such as speed, travel time, and queue length.

5.5.5 Acceptance and trust

Acceptability indicates the degree of approval of a technology by the users. It depends on whether the technology can satisfy the needs and expectations of its users and potential stakeholders. Within the framework of introducing new technologies, acceptability relates to social and individual aspects as well.

Regarding the dimension of "Acceptance and Trust", the following – soft – PIs should be focused on during FOTs:

Ex-ante usefulness (level of usefulness perceived by the user prior to usage): before using a system, what are the dimensions of usefulness that occur to the future user immediately? What are the benefits he expects from using the system? **Ex-post usefulness** (level of usefulness perceived by the user after practice with the system): after a first use of a system, what are the user's impressions regarding the system's benefits. Ex-post usefulness is to be analysed in relation to the statements of the indicator on "ex-ante usefulness". The reactions to both indicators will give useful information for system acceptance. The measurement of these two indicators can be operationalised via self designed questionnaires, based on established methodological approaches (see Nielsen, 1993; Grudin, 1992). A qualitative approach like a Focus Group with a formalized protocol and individual in-depth interviews is also appropriate. The **Observed rate of use** of the system or of specific system parts represents an additional indicator for system acceptance and perceived usefulness. **Perceived system consequences** (perception of positive or negative consequences of system's use) is another key indicator for system performance: the user expresses his/her impressions and attitudes regarding the potential consequences when using the system, which can be positive as well as negative. These impressions can best be collected via an interview and be exploited in Focus Groups, which have the advantage of group dynamics that can provide additional information on the subjective norm. Construction of standardised questionnaires is possible as well (for methodological background on this indicator, see Featherman and Pavlou, 2003). **Motivation** (level of motivation/impetus to use system) should be connected with the indicator **Behavioural intention** (level of intention to use system). Both indicators can be investigated best via self-designed questionnaires based on established methodological findings (see Amstrong, 1999; Ajzen and Fishbein, 1980). The **Response to perceived social control/response to perceived societal expectations** indicates the impact of perceived social control of the user’s behaviour. This indicator is a more sociological one, which should give an indication whether the user feels a social benefit (for example, social recognition) when using the system, or on the contrary, that he hesitates to use the system because he fears social disapproval when using the system (see Castells, 2002; Bahrdt, 1987). **Usability/level of perceived usability** concerns the aspects of the user’s general capacity to interact with the system (including installation and maintenance issues, see Grudin, 1992; Shakel and Richardson, 1991). For these indicators, the combination of in-depth interviews, Focus Groups and self-designed questionnaires based on established methodology is recommended.

5.5.6 Driver characteristics

Even though driver characteristics are not PI in themselves, they are important as Situational Variables, which is why they are included in this section. The focus here is on describing the drivers that participate in the study, as compared to selecting drivers based on certain characteristics, which is treated in (Chapter 6). Drivers differ on a large variety of characteristics, which may all have an influence on how they drive and use different systems and services. These differences may be important to take into account when planning a FOT. Four categories of driver characteristics may be distinguished:

- 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.

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 vehicle, which may influence the driver’s behaviour.

There are several different reasons for considering driver characteristics:

- To make sure that the sample of drivers is representative of the target population.
- To explain the outcomes of the FOT.
- To improve systems and services, taking into account differences between drivers.

Driver characteristics may play different roles in FOTs:

- Characteristics of drivers possessed before the FOT may play a role in how they behave in traffic during the FOT.
- Although some characteristics are stable, other ones may change when using a system or service in the FOT. Attitudes may change radically before and after using a system for a longer period of time.

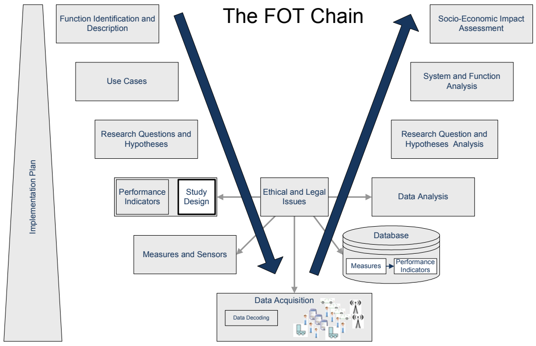
In general it is useful in a FOT to gather as many characteristics of drivers as practically possible. Even if no specific impacts are expected of certain characteristics, some outcomes may be explained better with more knowledge about the participants. A minimum set of data such as age, gender, income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

Next, information is needed about driving experience. Usually this is 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. Some widely used personality tests are the Five Factor Model test and the Traffic Locus of Control test. Special attention may be given to the personality trait of sensation seeking, which is correlated with risky driving. The Sensation Seeking Scale 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.

When evaluating the acceptance and use of new systems in the vehicle, drivers' acceptability of technology is important. Both social and practical aspects play a role. Technology acceptance has different dimensions, such as diffusion of technology in the drivers' reference group, the intention of using the technology, and the context of use (both personal and interpersonal). Measuring acceptability can be realized via (existing) standardized questionnaires, in-depth interviews before and after "use" (driving), and focus groups.

6 Experimental procedures



This section of the handbook provides guidance on the overall experimental design of FOTs in order to ensure experimental rigour and scientific quality. The first section — Participants — provides advice on participant selection, including demographics, driving experience, personality and attitudes, along with consideration to sample size. The second section – Study design – provides guidance of the formulation of hypothesis, experimental design and possible confounds. The last section – Experimental environment– suggests how the road environment (road type, weather conditions etc.) plays a part in the design of an FOT and the subsequent data analysis.

6.1 Participants

6.1.1 Characteristics

Depending upon the research questions, there is often a need to select a particular group of participants for inclusion in the FOT and ensure that this group is in some way representative of those drivers who will ultimately interact with the system.

The types of variables that should be taken into account include:

- Demographics variables, such as age, gender, social economic variables, and permanent or temporary driver impairments
- Driving experience, in general but also experience with various systems, accident history and the usual time of driving and roads used
- Personality and attitudes.

The first of these two variables are relatively easy to measure, using questionnaires. The data are objective and can be verified by the experimenter. Personality and attitudes, however, deserve more attention as there are a number of different ways in which one can evaluate these. FOTs may incorporate a battery of psychometric measures. Such measures are generally included in order to relate psychological factors to driving behaviour. Since drivers exhibiting certain traits or attitudes are known to engage in riskier driving behaviours, it would seem important that systems under investigation in FOTs are trialled amongst a range of drivers to ensure that the systems work for those who need it most.

Personality aspects that may be taken into account are:

- Sensation seekers, who tend to drive more recklessly
- Locus of control: drivers with an internal locus of control will continue to maintain direct involvement with the driving task choosing to rely on their own skills, whilst

those with an external locus of control may be more likely to rely on the system and surrender involvement in the driving task

- Drivers' attitudes towards road safety issues.

Personality and attitudes are known to affect the ways in which drivers interact with systems, and it may therefore be of interest to preselect certain personality types in much the same was as one would sample e.g. young males, or elderly drivers to a particular trial.

Recruiting on a personality/attitude base will ensure that a system is tested on a broad range of drivers who may interact with the system very differently. Recruiting on a personality/attitude base may be appropriate since these are likely to influence behaviour directly. Variations in beliefs are likely to explain differences in driver behaviour and system use. Before beginning recruitment for any FOT, researchers must consider the relationship between individual differences and the behaviour which the system is seeking to influence.

In addition to selecting drivers, personality and attitudes can also be used as covariates in analysis in order to identify several differences in driver behaviour and system use between groups. It is not imperative that FOTs base their recruitment on such measures. However, their inclusion within the experimental design provides useful insight into the manner in which individual characteristics influence behavioural adaptation to new systems.

Before deciding to recruit on a personality/attitudinal base, researchers should consider that, when tiding the inclusion criteria for any study, it is inevitable that there will be a progressive shrinking of the research participant population. 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. For example, since speeding represents a thrill seeking behaviour, high sensation seekers may be less likely to volunteer to participate in an ISA trial. Inevitably selecting participants on additional measures such as these will increase the burden associated with the recruitment phase of any FOT.

6.1.2 Sample size and power analysis

FOT studies should be able to assess the functionality of the ICT systems and their impact on the driver behaviour, traffic safety, environment, etc. When the chosen sample size is too small, it is difficult to statistically prove effects of the system that are actually there. With very large sample sizes the chance of finding an effect increases. However, there are two major drawbacks on just using a very large sample sizes:

- Every driver/participant needs a car equipped with the system and with a data logging system, which is expensive.

- Small effects which are statistically significant might be found, but they might not be relevant when looking at power effect.

The appropriate sample size for an FOT depends on a number of choices that have to be made in the final set-up. [3] These are, for instance, the number of ICT systems that are going to be tested and the choice of a between- (a separate group of drivers without an ICT system, but with data logger) or a within-subjects design (each participant drives a certain amount of time with and without the ICT system).

In order 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 are there, power analysis is needed to calculate the desirable sample size. This power analysis is based on a number of assumptions:

- Suppose an FOT is based on a between subjects design, such that different groups of drivers each drive with a different system. Or at least one group with an ICT system and one group without and ICT system
- The power is 80 %, indicating the chance of statistically proving a difference between the groups when it is there (i.e. a chance of 20 % of failing to prove it)
- The alpha level is 5 % (i.e. the chance of falsely finding a significant effect)
- Two-tailed testing, because we have no reason to assume that either one of the groups performs better/worse than the other
- The effect size is 0.2, which is typical for a small effect that can be expected in a FOT with a lot of disturbing factors compared to more experimental test set-ups. An effect size of 0.5 is typical for a medium size effect.

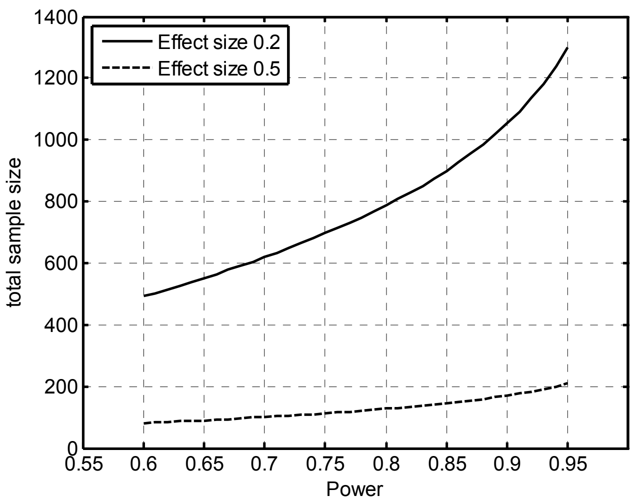


Figure 6.1: Total sample size as a function of the statistical power and the effect size (2-sided test, alpha = 0.05, independent variables).

[3] Further detail on how to choose the sample size is to be found in FESTA Deliverable D.2.4.

Figure 6.1 shows that a total sample size of 800 (i.e. two groups of 400) drivers would be needed to be able to statistically prove small size effects between the two groups. The groups are relatively large to compensate for the relatively high number of disturbing factors when trying to find effects in real traffic. If we expect medium size effects groups of only 75 drivers would be sufficient. If a within subjects design is chosen, one group of 400 drivers would be sufficient to test both the without and with system conditions.

6.2 Study design

6.2.1 Hypothesis formulation

As a general rule, research practice proceeds in the following way:

1. Formulation of the hypothesis
2. Testing the hypothesis
3. Acceptance or rejection of the hypothesis
4. Replication of the results or (in the case of rejection) refinement of the hypothesis

A hypothesis is specific questions which can be tested with statistical means by analysing measures and performance indicators. It is a tentative explanation for certain behaviours, phenomena, or events that will occur. It is essential for an FOT to be designed with clear hypotheses in mind in order to aid the interpretation of the results.

In formulating a hypothesis, consideration should be given to the variables under scrutiny. It is vital that the variables collected in an FOT allow the researcher to accept or reject their hypotheses. To do this, both the independent and dependent variables should be well defined at the start of the FOT. The independent variable is one which can be manipulated by the researcher. As the researcher changes the independent variable, he or she records what happens using dependent variable(s). The resulting value of the dependent variable is caused by and depends on the value of the independent variable. Other variables, known as controlled or constant variables are those which a researcher wants to remain constant and thus should observe them as carefully as the dependent variables. Most studies have more than one controlled variable.

6.2.2 Experimental design

The two basic types of experimental designs are *within subject design* (this is sometimes also referred to as crossed design) and *between subject design* (this is sometimes also referred to as nested design). FOTs also need to contain a control condition, in which subjects do not get any treatment. This condition is meant to serve as the baseline: This is how drivers behave in case there is no treatment or no experimental manipulation at all.

Within subject design

In a within subject design, each subject encounters every level of treatment or experiences all experimental manipulations. For example in a FOT evaluating navigation systems, every subject drives for some time with (experimental condition) and for some time without (control condition) the system. In this specific case, one half of the subjects would start with the control condition and then switch to the navigation (experimental) condition and half of the subjects would do this vice versa.

This type of design has two advantages: (1) fewer subjects are needed compared to a between subject design, and (2) is more likely to find a significant effect, given the effects are real. The power of a within subject design is higher than in a between subject design. This is related to the reduction in error variance, since there are no individual differences connected to differences in treatment measures. A disadvantage is the risk for carry-over effects, which means that if a subject experiences one condition, this may affect driving in the other condition.

Between subject design

In a between subjects design, each subject participates in one experimental (or control) condition. The major distinguishing feature is that each subject has a single score (with or without the system). Note that the single score can still consist of driving on various types of roads, during long periods of time or different types of driving behaviour, workload and comfort.

The advantage here is that carry-over effects are not a problem, as individuals are measured only once in every condition. The total number of subjects needed to discover effects is greater than with within subject designs. The more treatments in a between subject design, the more subjects are needed altogether. In order to limit the confounding effects due to individual differences in a between subject design, one should either use random assignment, in which the assignment of what subject is exposed to what treatment is done randomly or use matching groups (also called matched pairs), in which one also has to make sure that different groups are comparable with respect to pre-selected characteristics, such as gender and age. In order to do this, one needs to identify the variables that one wants to match across the groups, and measure the matching variable for each participant and one needs to assign the participants to groups by means of a restricted random assignment to ensure a balance between groups. Also, one needs to keep the variable constant or restrict its range. This will reduce differences within each group and therefore reduce within treatment variability.

The main drawback with the matched pairs design is in the sampling process. As the number of characteristics that require matching increases, so a correspondingly large sample pool will be required to allow adequate matching to be possible. A further problem is that this design assumes that the researcher actually knows what extraneous factors need to be controlled for, i.e. matched — and in some circumstances this may not always be the case.

Longitudinal and Cross-Sectional Designs

One question a FOT may have to answer is whether an effect of a treatment (e.g. driving with a system) change over time? To investigate this, longitudinal or cross-sectional designs can be employed. While longitudinal surveys of this type can be very useful they do not provide an answer to the questions concerning why the changes may or may not have occurred. If things like that are measured in FOTs, one should already have a clear idea why a positive effect may disappear after a while. This could for instance be such factors as risk compensation (because the systems warn you, you can drive until you are warned).

One of the difficulties with longitudinal studies is that it is hard to keep subjects motivated during the entire study period, or people may move, or become ill. Because of these difficulties other methods for investigating changes over time have been developed and the cross-sectional design offers an alternative.

The cross-sectional design looks at changes over time by taking a number of cross-sections of the population at the same instant in time. This is obviously quicker and less costly than a longitudinal study, and there is a lower chance of actually 'losing' participants during the run of the experiment. On the other hand, a main drawback with the cross-sectional study is related to the previous experiences of the participants and how this might have an impact on the findings.

6.2.3 Threats to validity: confounds and other interfering effects

As a general rule, the results of an empirical 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 because the study has been designed in a way that alternative explanations for the effects are implausible or can be excluded. The internal validity of a study increases to the extent to which such alternative explanations can be ruled out. In the literature these factors are also described as confounded variables which need to be controlled by appropriate measures right from the beginning of a study.

In the literature several interfering effects have been described which interfere with the effect of an independent variable on a dependent variable and contribute to a decrease of internal validity if they are not controlled by measures implemented in the experimental design. The following effects which constitute threats for internal validity of FOTs:

- **History:** Unplanned events unrelated to the study might have an effect on the correlation between independent and dependent variables. For example, during the performance of a FOT an important paragraph of the road code might be changed (e.g. new speed limits for certain road categories) which is accompanied by increased police surveillance activities.

- **Maturation:** Mainly effects due to experience and learning which affect the dependent variable and are (in long-term studies) erroneously attributed to the independent variable.

- **Testing:** If the behaviour of interest is sampled at different times there might be a biasing effect from the number of times, e.g. by becoming more familiar with the test situation. For FOTs this might become relevant if subjects are tested at different times over the course of the study but not if their behaviour is sampled continuously and more or less unobtrusively.

- **Selection:** In general, the participation in an FOT is voluntary which means that the strategy of recruiting subjects can have a biasing effect. For example, to offer a certain amount of money (e.g. 500 Euros) as compensation for the effort caused by completely finalising the study might be an incentive for participants with a low income whereas it might insult people with a very high income.

- **Drop-out:** During the run of an FOT one has to take into account that not all subjects will finalise their participation as planned. However, this drop-out can have a biasing effect on the results of an FOT if the subjects who quit early differ systematically from those who finalise as planned with regard to relevant characteristics (e.g. socio-economic status, age, gender etc.).

- **Experimenter-bias:** Effects on the dependent variable which result from the social interaction between the experimenter and the subjects which might occur, for example, if at the beginning of an FOT the experimenter explains the system functions very carefully to some subjects due to sympathy whereas he is careless with this at some others.

6.3 Experimental environment

The experimental environment is a critical element within an FOT, since it will determine the data that is collected and the ability to fulfil the objectives of the FOT. In general, environmental factors can be treated in several different ways, including

- Explicitly included in a FOT because you are particularly interested in data connected to that environmental factor (e.g. motorway routes for lane departure warnings)

- Explicitly included in a FOT because these environmental factors are part of the range occurring within a normal driving scenario (e.g. night time driving)

- Measured scientifically so that the data relating to that environmental factor can be included within post trial data analysis (e.g. vehicle headways)

- Recorded (in varying levels of detail), so that portions of data can be excluded from analysis (e.g. heavy rain, where all or some of the data from a particular day may be discarded; or overtaking manoeuvres where short periods of data within a larger set are discounted during a study of steady following behaviour).

6.3.1 Geographical location

In line with above, the geographic location can be chosen because it is representative of the intended area of use for a vehicle/system (e.g. predominantly motorway environments). Alternatively, the geographic area can be chosen because it displays the characteristics needed to collect the specific data you are interested in during the FOT (e.g. the choice of mountainous and/or northern European environments in order to collect data on the use of systems in cold environments).

The population within a particular geographical location may affect the running of the FOT. For example, certain cultural issues, population characteristics, car ownership, use of new technologies, and language issues may be apparent. In addition the characteristics pertaining to the road and prevailing traffic may be of importance, including:

- Road type and localities present
- Traffic patterns, such as types of journeys (e.g. commuter or tourist travel), traffic flow, traffic density, vehicle types, and frequency and sophistication of journeys
- Other transport options, the availability and costs and the inducement or penalties to encourage particular transport mode choices
- Legal regulatory and enforcement environment, such as speed limits, levels of enforcement of traffic regulations (e.g. speed cameras), penalties for traffic or other violations, standardisation (e.g. compliance of road signs with international standards).

The geographical location may also have implications with regards to technical and other study issues, including infrastructure and data communication issues such as:

- Network/beacon infrastructure for vehicle-infrastructure communication
- Network coverage/reliability for telecommunications, especially if automatic over-the-air data transmission is used instead of manual data download
- Localised GPS coverage issues (e.g. urban canyons, foliage cover)
- Logistical issues, such as safe and secure access to target vehicles for data download (if data is not being transmitted over the air), and for maintenance
- The availability and quality (resolution, scope and depth of content) of electronic maps that can integrate vehicle location
- Availability of other data, e.g. from the police, highway authorities, fleet operators, maintenance personnel.

The most important point in relation to the geographical area is that it must be chosen based specifically on the *objectives* of the particular FOT, and in particular, in relation to the validity of the data that is being collected. 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 generalised to the wider ‘population’ of interest (i.e. external validity)?

The starting point is to consider the overall objectives of the FOT, including the types of cars and systems that will be incorporated into the trial. The second major consideration is that of generalisation of the results. In particular it is necessary to ensure that geographical aspects are included to ensure that the data collected during a specific FOT can be generalised to the wider population of interest. 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 the presence or absence of a particular factor, then the geographical environment(s) must include that factor (and possibly variation thereof).

6.3.2 Road type

The road type is the environmental factor that perhaps has greatest dynamic influence on individual and collective driver behaviour, and hence impact on safety, mobility, traffic efficiency and the environment within an FOT. It is highly dependent on the geographic area, as discussed above.

The road type will encompass a number of variables which will influence driver use of systems, driver attitudes, driver behaviour, and driver outcomes. The FOT may want to include roads with specific characteristics, including:

- Surfaced or unsurfaced roads
- Minimum, average and maximum speeds of traffic
- Number of lanes, and presence of lane marking
- Visibility (of the environment and other traffic)
- The types of manoeuvres that a driver will need to undertake (e.g. stopping at traffic lights, or overtaking manoeuvres)
- Typical vehicular headways
- Presence of safety features such as rumble strips or speed cameras.

Three main categories of road should be differentiated:

- Urban
- Rural
- Motorway

Note that road classifications differ in different countries and there is no standard European classification. Ideally a map and a database of the region of deployment of the FOT 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.

6.3.3 Traffic conditions and interactions with other road users

Traffic conditions and interactions with other road users are important considerations. A distinction needs to be made between:

1. 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
2. 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;
- Stability: this can be within a traffic stream (in which case it is expressed in terms of the frequency of speed variations on a traffic lane in a given unit of time) or between different traffic streams (in which case it is expressed in terms of the frequency of lane changes in a given unit of time)
- Speed: the average speed of traffic
- Composition: 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:

- The category to which they belong (light vehicle, heavy vehicle, van, motorcycle, pedestrians)
- Their speed and acceleration (direction and rate)
- Their manoeuvres and behaviour (merging into the subject's lane or pulling out into a lane, merging from an entry slip road, braking, etc.).

Other characteristics to be taken into account are:

- Route choice
- Temporary road/traffic variables
- The traffic encountered
- Impact of road measures on driver behaviour
- Static and dynamic variables associated with the road

6.3.4 *Roads to include in an FOT*

When setting up and running an FOT, it is necessary to consider the extent to which specific road types need to be incorporated into the trial and hence which participants need to be selected. The basic questions to consider are:

- Are specific road types needed to answer the research questions for that sample?
- Would any system of interest be used on a range of different road types?
- Do you expect driver behaviour (in terms of safety, mobility, traffic efficiency and environmental impact) to differ according to the road type they are travelling along?
- Do you need to be able to compare results according to different road types?
- Do you need to include specific road types in order to generalise the results to a wider population?
- Are interactions with other road users to be included in the analysis? If so video equipment needs to be installed.

By considering the above questions, one can determine whether a range of different road types are needed, or whether the FOT can concentrate on collecting data based on specific road types. In a FOT, the objective is usually to study the normal driver behaviour. This means that drivers should not be encouraged to change their normal routes.

6.3.5 *Weather conditions*

Weather conditions are hard to predict, control for, or measure accurately in an FOT. However, weather conditions and associated factors such as ambient lighting are relevant aspects for all FOTs, irrespective of the overall purpose of the study. A well designed FOT must consider a range of weather-related issues, with a view to including, targeting or excluding particular weather conditions. In order to include weather as an experimental variable within analysis, or to specifically include or exclude data for analysis, it is necessary to use a consistent taxonomy and definition of weather conditions.

Related to how weather factor are measured, is the level of accuracy that you employ in the measurement of weather factors, including location and time attributes. A further complication with weather factors is that it is often combinations of weather and other dynamic and static factors that have a practical impact on an individual driver or general traffic conditions within an FOT. Extreme weather conditions present a risk to FOTs because they often can't be predicted, and can make journeys impossible, prevent access to vehicles, or in the worst case can destroy equipment.

Data may be confounded due to abnormal weather, for example snowfall increasing driver headways and reducing traffic speed or bright sunshine causing glare on screens in vehicles, or momentary distraction to drivers.

There are several ways of potentially measuring weather conditions:

- In real-time using direct measurement of the factor, e.g. vehicle sensor to measure ambient temperature (which could then be used to link the use of features to outside temperature)
- Indirect real-time measurement using a surrogate sensor (e.g. recording the use of the windscreen wipers to indicate when it is raining)
- Subjective rating scales (completed by the driver or other) – e.g. a driver assessment of the degree of rainfall
- Post-hoc data mapping – the use of weather records to estimate the weather conditions
- Post-hoc analysis of video data by a trained data coder

At a general level, there are four main considerations with regard to weather:

- Which weather conditions are relevant?
- Should they be 'designed in' or 'designed out' of the study?
- Do weather conditions of interest have a macro (e.g. a rainy day) or micro (e.g. reflected glare) level impact?
- What level of data is needed, and how is this obtained?

6.3.6 Time of day and seasonal effects

Temporal factors such as time of day, and seasonal effects have a considerable impact on the planning of FOTs, and the analysis of data. In contrast to the weather effects outlined above, the temporal factors can usually be predicted, and so it is usually easier to deal with the issues successfully. The main issues to do with the time of day, week, and seasonal variations are:

- Influence on driver state (e.g. sleepiness)
- Disruption caused by external events, for example school opening times
- Influence on traffic levels
- Other temporal influences on traffic
- Impact on vehicle occupants
- Glare
- Ambient light levels
- Seasonal confounding of data collection
- Influence on route choice

- Pragmatics to do with drivers work and life schedules
- Using time of day as a surrogate, for example, time of day can be used to specify or control for traffic levels or ambient light levels.

Time of day and seasonal effects are different to weather issues in several ways, including:

- Time of day and seasonal effects are much more predictable than weather conditions
- They are often proxies – i.e. not important in themselves, but important because they result in variation of a factor of interest (e.g. traffic levels, or level of the sun above the horizon)

These two factors mean that a greater emphasis should be placed on planning around relatively predictable time of day and seasonal effects, and considering their impact on the FOT.

6.4 Conducting a pilot study to test the evaluation process

Conducting a pilot study is necessary to prepare the deployment of the FOT and to support the design of the relevant tools for the evaluation process (Saad, 1997; Saad and Dionisio, 2007). This task should be performed early in the evaluation process. It represents an important step for the mobilisation and the dialogue between the various teams involved in the FOT and for promoting a common framework and consensus for the evaluation process. These preliminary field tests have to deal with three main levels of analysis with specific objectives.

1. Obviously, the first preliminary field tests have to check *the technical functioning of the data collection systems in real driving situations*. They should enable to identify potential problems of sensor calibration or drift and thus to establish the periodicity of maintenance procedures during the FOT. They should also permit to validate the data collection procedure from data acquisition, data transmission to data storage.

   The technical teams involved in the FOT should be in charge of these field tests.

2. The second level of preliminary field test deals mainly with the issue of *assessing the usability and usage of the systems under study and of identifying the main critical issues associated with their use in real driving situations*. This is particularly relevant for:

   - Structuring the familiarisation phase of the drivers before their participation to the FOT;
   - Contributing to the design of the questionnaires for the subjective assessment of the systems;
   - Testing and/or improving the various tools developed for data processing, such as automatic identification of critical “use cases” and “scenarios” and video based identification of triggering events or categorisation of road and traffic contexts.

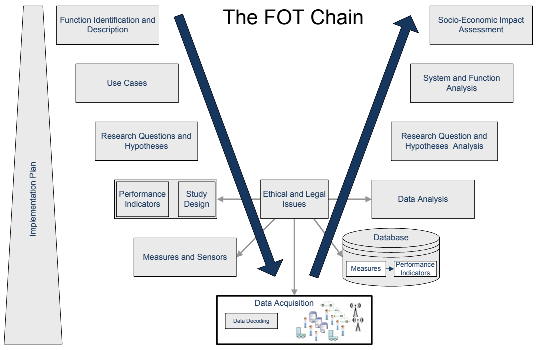
- Identifying a number of critical scenarios when using the systems, scenarios that could be investigated more extensively when the data gathered from the FOT are processed and analysed.

This test requires the participation of a sufficient number of drivers (depending on the target population in the FOT) and should be performed in real driving situations. An experimental journey on the road could be designed for that purpose (depending on the hypotheses formulated). This level of analysis provides useful data for designing the relevant tools for the evaluation process as mentioned above, for estimating the time required for data processing and data analysis and thus calibrating these phases in the FOT. It may be seen also as an opportunity for training the team (s) in charge of data processing. Finally, it represents an important step for testing some of the hypotheses formulated in the FOT and/or for refining them.

Psychologists, ergonomists, and human factors experts should perform these tests in close cooperation with the team in charge of statistical analysis as well as the team in charge of developing data processing tools.

3. The third level consists of *testing the feasibility of the overall evaluation process* from the selection of the participants through to data collection. It is a kind of final rehearsal before the deployment of the FOT. It enables in particular a check of the communication process between the various teams involved in the practical deployment of the FOT and of the robustness of the technical tools designed for data collection and transmission.

7 Guidelines for Data Acquisition



This section aims to provide guidelines and recommendations for how to handle data in an FOT study. Data acquisition, data storage, and data analysis tools will be covered here.

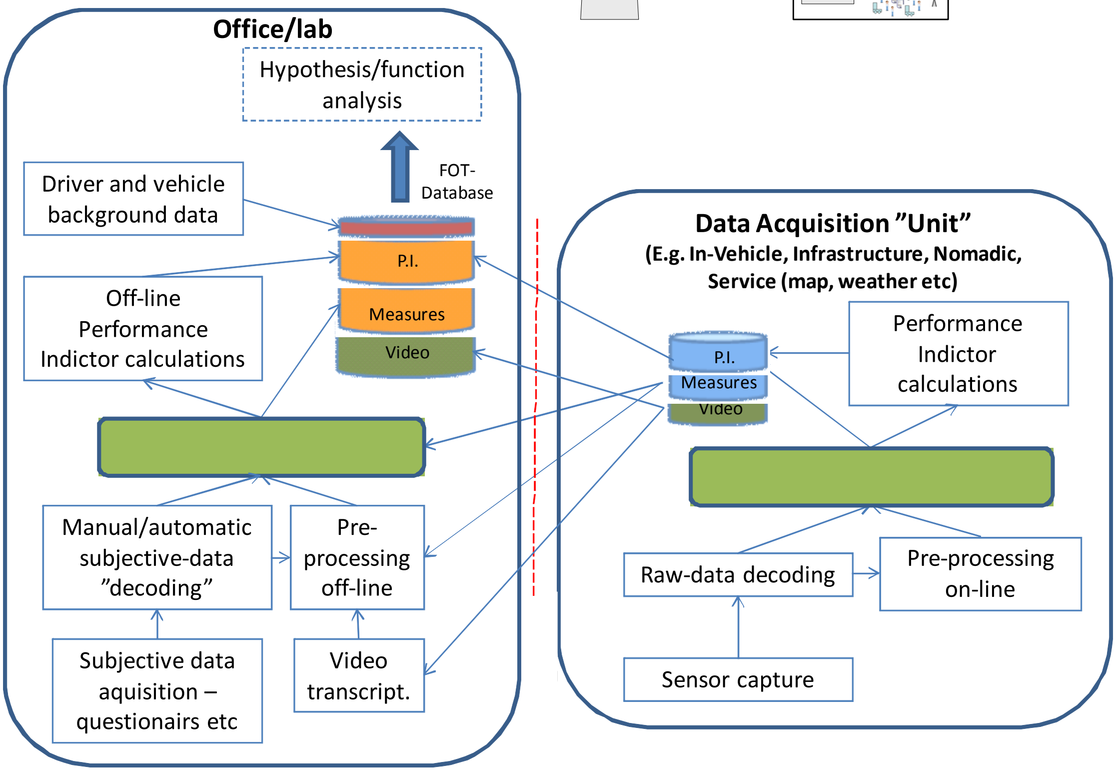


Figure 7.1: Data structuring

Please refer to Figure 7.1 for an overview of a data handling structure for an FOT, and for the naming conventions used in this document. The example data structure above includes data from an electronic data acquisition system, as well as collected subjective data. The Data Acquisition Unit (on the right) comprises sensor systems requiring raw data decoding. The raw data may then be pre-processed, in this case by low-level data processing such as simple filtering or calculation of directly derived results. Both raw data and pre-processed data (derived from raw data) are stored in the same format locally and can be kept locally, until moved from the Data Acquisition System (DAS) to the main storage, and used for analysis.

Acquisition of subjective data may also be performed. *Subjective data is also considered as acquired from a sensor.* This data can similarly be subject to manual or automatic decoding, database storage (pre-processed or not), and subsequent use in performance indicator calculations.

7.1 The measures and sensors tables

FESTA recommends use of the FESTA matrix, a spread sheet in Excel format containing three tables: “Performance Indicators”, “Measures”, and “Sensors”, that may in a later stage be utilised to create a relational database. The sensor table should be used in preparation of, during, and after an FOT project, in conjunction with the other two related tables (see Appendix 1 in Deliverable D2.2). Properly handled and thoroughly implemented, the tables are valuable tools for data structuring, for data requirements specifications and for identification of connections between sensors, measures and performance indicators.

The FESTA matrix is intended as an aid and gives examples of commonly used sensors, together with specifications. It is not exhaustive and future users are encouraged to modify, expand, or limit the matrix to suit the particular FOT.

7.2 Data acquisition

Methods of data acquisition in FOTs include methods to collect background data, digitally acquire data from sensors, and subjective data (such as data acquired from questionnaires). In addition, data in the form of manually or automatically transcribed data and reductions of collected data is also considered sensor acquired data (but with a manual sensor – the analyst). In FESTA *all the data sources mentioned above are considered sensors*. Subsequently can all data be acquired, stored, and processed in a generalised way.

All of these different data types are used to support the hypotheses defined for the specific FOT. The data to be collected should be defined and based on research questions and hypotheses.

7.2.1 Background data acquisition

The background data about the driver is crucial and needs to be collected integrated in the driver interaction procedure. Due to privacy issues different parts of the background data may or may not be suitable for storage in a database, or be used in statistical and other forms of analyses.

Data could be gathered by interviews and/or questionnaires, by different tests, or by specific instruments. The driver background information should be considered as acquired from a sensor, and preferably be added into the database and to the sensor matrix.

7.2.2 In-vehicle data acquisition

An in-vehicle Data Acquisition System (DAS) is needed in FOTs where the focus is either to study in-vehicle systems by collecting data from the systems in the vehicle. A suitable DAS can differ from study to study and a specific solution cannot be recommended for all types of FOTs. See section 3.1.2 in D2.2 for a list of different DAS solutions.

The guidelines and requirements in this document are based on experiences from FOTs using some kind of in-vehicle data acquisition.

7.2.3 Nomadic devices

A nomadic device (ND) or an aftermarket device (AD) could be either part of the function/system under test, or it could be part of the data acquisition system, acquiring specific FOT data.

Nomadic devices can also be used as data storage tools as they are easy to install and use on different kind of vehicles. If the vehicle has a dedicated gateway for ND, this option can be used for capture of further vehicle related data.

Using the local wireless connections, the storage capacity of ND could be extended with large capacity hard disks. A possible drawback of a ND, when used as a DAS in itself, is that test subjects must remember to bring the ND to the vehicle every time he/she uses the vehicle.

7.2.4 Subjective data acquisition

As explained before, also subjective data are considered as “sensor” data in the scope of the FOT methodology. All subjective data should therefore be stored and handled logically as if it were collected from a “real” sensor. Subjective data may include data acquired from the test subjects in different ways. Results from interviews and questionnaires are typically subjective data.

The result from the subjective data acquisition should preferably be stored in an electronic format. Electronic compilation of the questionnaire may be considered to reduce the overall manual work and cost, maybe by using web based tools.

For subjective data to be stored, the following related information is required:

- Date and time (hh:mm) of test start
- Date and time (hh:mm) of test end
- Subject ID code
- If present, reference to objective data (file name, location)

7.2.5 Real time observation

In this context, real time observation data is data collected by an observer that directly or indirectly (in real-time or afterwards – for example on video) is observing the drivers and systems to be evaluated. The data acquisition process is usually relatively manual but the results should be transferred to digital format and uploaded to the FOT database for further analysis.

Real time observation data help provide a more detailed picture of a driver’s behaviour, as well as verifying the information gathered by other instruments. As the overall purpose of an FOT is to collect information on as natural driving as possible, with an observer physically in the car there is always the risk of the driver not acting the way he or she would otherwise.

Direct real time observations must therefore be carried out with great care and as unobtrusively as possible, or avoided completely.

7.2.6 Acquisition of infrastructure data and other services

General aspects

The infrastructure can be equipped with sensors to detect e.g. traffic or weather conditions. Data from such systems can be collected in raw format or in an aggregated form. If data is collected both on the vehicles and on the infrastructure separately, it is necessary to synchronise the two sets of data. It is recommended that *GPS time* is used as the synchronisation source.

Infrastructure

It is in many countries required to contact local road authorities before the installation of equipment close to a road. Working close to or on roads may (depending on country) require special training or licence. In some countries it is even required to use a special company or local road authorities for any installation work close to or on roads.

Services

When using such sources it is recommended for traceability (during and after the project ends) to record information about for example version of service, update rates and resolution/precision of the information they have during the duration of the study. It is also recommended to invite the service providers for discussions and possibly partnership in the FOT.

7.3 Specific sensors in FOTs

7.3.1 In-vehicle video

Video is by most state-of-the-art FOT experts considered a very important data source for the identification of driver behaviour and reactions, as well as for the process of analysis to understand the underlying context with regards to the surrounding. When a certain situation or event has been identified for evaluation of a particular system, the video provides the analyst with information about the context of both driver behavioural aspects and the interaction with the environment (if external video is used).

Video can be used in mainly two ways:

- Driver monitoring: Firstly the driver eye/head movements in relation to the vehicle/environment/context, and secondly the driver interaction with the vehicle and other driver actions (pedal, gear shift, steering wheel handling, mobile phone, eating, etc.)
- Environment/contextual monitoring: Helps understanding the driving contexts by collecting information about the surrounding traffic.

For a specific FOT the number of cameras needed, the needed resolution, the views captured by the cameras, etc., should be *defined by what is needed to address the hypotheses*.

To find the correct requirements thorough investigation is needed. Evaluate the results using calculations of predicted storage needs, as well as evaluation of video image quality and set requirements accordingly. In this evaluation experts from both study design/analysis planning as well as the analysis team should be included.

Generally the following parameters affect video quality, and need to be considered: picture resolution, frame rate, colour settings, “regions of interest”, bit rate strategy, bit rate limitations, “quality” strategy settings, and other video compression method related settings.

For more exhaustive methods for defining the video requirements, for information on potential camera related quality problems involving interlacing, and further information on video compression, see section 3.2.1 in D2.2.

Recommendations

Make a thorough analysis of video acquisition requirements. Set requirements necessary for each individual view, to possibly achieve a first limitation of video data. Choose a well-tested hardware video frame grabber/compression solution, and choose compression suitable for the FOT. The MPEG-4 part 2 and part 10 (H.264) compression feature sets have been used successfully in other FOTs, as well as MJPEG compression.

7.3.2 Internal vehicle bus data

Most modern vehicle manufacturers features one or several internal vehicle networks such as CAN, LIN, MOST or FlexRay. An internal network may carry large amounts of useful information for the FOT. However, there are several concerns with accessing and ascertaining quality on data from an internal vehicle bus.

Accessing the vehicle bus

Accessing information from an internal vehicle bus can be highly complex and even void warranty if it is done without the OEM’s permission and supervision.

OEM cooperation

A description of the entire vehicle network will often contain proprietary information, and may reveal detailed information on specific functions and the vehicle system architecture. Thus, non-disclosure agreements are typically required and can be hard to attain.

An option to using the entire vehicle bus description is that the OEM only provides the description of a selection of signals which enable access to the most important data. Still, however limited the access to the vehicle bus is, there may still be proprietary issues.

By using a bus gateway, the OEM will be able to extract data from the bus without providing any information about how the bus actually works. The gateway is programmed by the OEM to read certain information, decode it and then pass it on to the FOT logger equipment. An illustration of the process is shown in Figure 7.2.

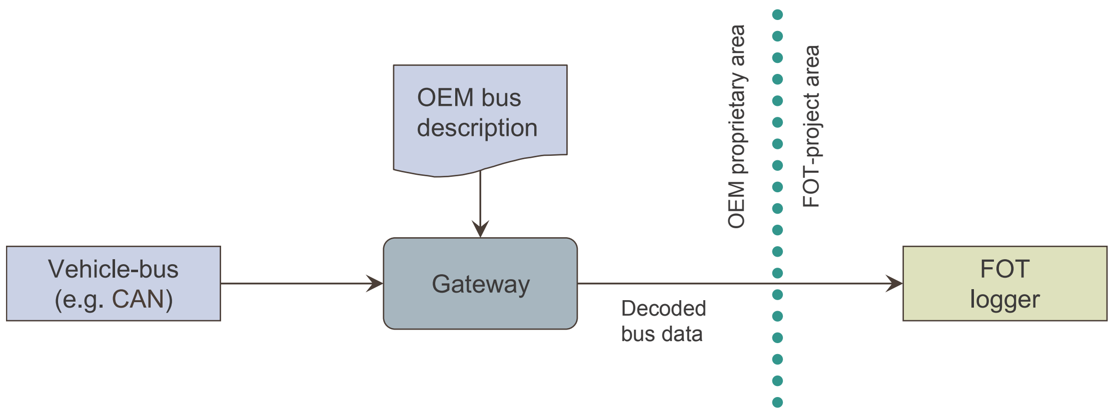


Figure 7.2: Illustration of process for securing proprietary vehicle-bus data using a gateway.

When physically connecting to the vehicle bus, it is important to follow applicable standards of the bus in order to prevent interference that may reduce network functionality (and thus warranty). It should also be stressed that every bus implemented by an OEM might not necessarily fulfil the standards in every detail.

Sensor specifications and details

When acquiring sensor data from a vehicle bus, the information is passed through several stages before it can be read from the bus (see section 3.2.2 in D2.2). These stages are likely to affect the signal value both in terms of amplitude and frequency and thus need to be closely observed.

To ascertain that qualitative measures are attained, plenty of contribution is required from the OEM. A thorough description of a vehicle bus and its ECUs will in many cases require involvement by subcontractors as well. A list of required details for successful data acquisition from the frequently used CAN bus is provided in section 3.2.2 in D2.2.

7.3.3 Automatic in-vehicle driver monitoring

Head/eye tracker

Two of the main issues for many systems are visual distraction and the effects the system has on the driver attending to traffic versus to the0020system. By using eye or head trackers in the vehicles, continuous data of some driver state/attention measures can be obtained. The problem with head and eye trackers is mainly the risk of significant data dropouts due to limitations in driver head and gaze tracking.

State-of-the-art head/eye-tracker technology today is relatively expensive, but the benefit of using such a system should not be underestimated. It is recommended that head or eye tracking systems are employed in FOTs where driver state is an issue.

Using an unobtrusive system is a requirement for head/eye-tracking systems for FOTs. The driver should not have to initiate the tracking system or wear any device; the system should be as inconspicuous as possible to the driver. Also, the system should not require any manual calibration in the field.

Other

It is strongly recommended that adding sensors that the driver has to put on and wear should be *avoided* to assure as natural driver behaviour as possible. (For further examples see section 3.2.3 in D2.2.)

7.3.4 Extra analogue/digital data sources

Access to certain information that is not available on vehicle bus systems is often needed. An analogue/digital I/O device for data acquisition is thus required. Also, anti-aliasing issues need to be addressed. The requirements for the different kinds of extra analogue and digital sensing need to be defined in the study design and will not be covered here.

7.3.5 Radar and other non-video environment sensing

Sensors already integrated (by OEM/Road Administration)

If a required environment sensor is integrated into the vehicle or the infrastructure, great effort should be spent on trying to add the sensor data to the used data acquisition system. In this integration several issues have to be considered (for details see section 3.2.5 in D2.2):

- OEM allowing/disallowing access to data.
- System interfaces: Some low level sensing information (e.g. object tracks from radar) may require special interfaces to be implemented both in hardware and software;
- System comparability: If the studied vehicles have different OEM integrated systems, they will provide different quality of data as well as different resolution, range, field of view etc.

Add-on environment sensing

If additional sensing needs arise, please consider the recommended process for adding sensing:

- Use the sensor matrix for sensing needs and identification;
- Requirements on the extra sensing need must come from hypotheses and performance indicator requirements;
- Additional considerations: Does the extra sensing require additional interfaces? Does it require information from vehicle buses? How can the sensor be integrated without significant effort? Does it require repeated calibration?

General guidelines for specifications are difficult to define. However, for example radar and other object tracking sensing systems, required field of view, radial and angular resolution/precision is important to define based on the hypothesis. For further suggestions, see section 3.2.5 in D2.2.

7.3.6 GPS

A GPS device can provide a *GPS Time* reference time stamp and a difference between the GPS time and UTC. It is highly recommended that this is used for synchronisation within a data acquisition unit as well as between systems (in-vehicle, ND, infrastructure and services). Information about the present local time zone can be useful for subsequent analyses and synchronisation with non-UTC devices.

Multipath propagation of the GPS signal is a dominant source of error in GPS positioning, since they depend on local reflection geometry near each receiver antenna. In most cases these can be corrected to a large extent with DGPS solutions. The errors depend on time of day, and satellite positioning (in zenith or low orbits) as well as other atmospheric disturbances.

7.3.7 Audio and driver annotation

Continuous audio recording is potentially a significant privacy issue and is not recommended. In state-of-the-art FOTs drivers have had access to a comment button on the dashboard, which, when pressed, would start recording any verbal comments during a pre-set number of seconds (usually around 20-60 seconds). Use of the button could be encouraged if the driver feels that it is warranted or at agreed (critical on non-critical) events. Be sure to inform the drivers consistently about the annotation possibility and provide some simple guidelines on how to offer the comment.

7.3.8 System function/status

A system under evaluation, such as an LDW, ACC or FCW, needs to be continuously monitored to ensure that it is operating properly. The system status signal will thus form a measure to be recorded in the data acquisition system. The status signal will depend on the specific system and needs to be provided by the system manufacturer or vehicle manufacturer, thus requiring strong collaboration with the actual provider.

7.3.9 Vehicle metadata

Information about the studied vehicle is important for analysis and study design. The recommendation is that for each type of study, systems, functions and specifications that may act as confounding parameters in a specific analysis should be stored. During analyses, these parameters may have confounding effects if an inhomogeneous test fleet is used. Examples:

- FOTs with focus on safety: Information should be stored about integrated systems that may contribute to driver distraction,

- FOTs with focus on environmental issues: A more powerful engine, automatic gear shift or four-wheel-drive, are most likely to be confounding parameters in an environmental analysis.

7.3.10 Coding/classification/transcription

As part of the data reduction and analysis process, described elsewhere in this handbook, sections of time will often be labelled with classifications according to a coding scheme or syntax. Depending on the study, sections of time can be given categories such as "crash", "near-crash", "incident", "Curve speed warning", "lane change", "crash avoidance by steering", etc. When classifications are made they are often saved and thus become a new data source which is added to the database. For example, an index indicating all instances of lane changes in the dataset can be created and saved. Regardless if the classification of data is performed by a human analyst transcribing video or by an algorithm applying kinematic trigger values to the data, this process of classification should be seen as a type of sensor providing a new data source. Thus, it is comparable to other types of off-line performance indicator calculations (see Figure 7.1). It is recommended to plan that these new data sources or measures will be created during the data reduction phase after the data has been uploaded to the FOT database.

7.3.11 Geographical Information System (GIS)

One of the lessons learned in state-of-the-art FOTs in the US is that geographical information, such as road curvature, roadside embankments and other on-and off-road information has been underestimated as a valuable source of data in the analysis. Contextual indicators of events and situations identified in the analysis provided added insight into both behavioural aspects and how the infrastructure influences system performance. (See section 3.2.12 in D2.2 for more information.)

GIS derived information about the current road (road type, speed limit, rural/urban, banking, curve radii, etc.) could be used directly in the vehicles as separate measures. By doing this on-line, the necessary post-processing is reduced, but requires additional software and possibly hardware. One advantage of performing this on-line can be that the absolute position (e.g. GPS) does not have to be stored, thus reducing the some privacy concerns. An advantage of using commercial navigation software as part of the on-line map-matching and information extraction can be that there is no need for potential in-project development of map-matching algorithms. This technique has been used in some state-of-the-art FOTs. It is important to validate that map-matching and other GIS data extraction is done in a proper way.

7.4 Mechanical requirements

The following mechanical guidelines and requirements are primarily applicable for FOTs with in-vehicle Data Acquisition Systems.

7.4.1 Size and weight

The system should preferably have negligible effect on the driver's use of the vehicle – including limitations in trunk space.

7.4.2 Connectors and interfaces

State-of-the-art FOTs state that from experience as much as 80 % of the DAS hardware problems can be deduced to physical connector issues (to the DAS and to peripherals). It is recommended that connectors with some locking between connector genders are used. Cable pull-relief should be used when possible.

7.4.3 DAS mechanical cover and ease of access

It is recommended that a layman without tools is able to find and have visual access to any indicator LEDs on the DAS. Also, having the possibility to connect interface devices without having to remove covers is preferable.

7.4.4 Crashworthiness and vibration resistance

For all FOTs the minimum requirements for ruggedness is that the entire system should operate under the normal driving conditions for the specific FOT, including the harsher situations of normal driving.

7.4.5 DAS environmental requirements

Environmental requirements for the DAS mainly concerns temperature. If the DAS is placed in a shielded location, the need for water resistance may be negligible, although the DAS internal parts should be able to withstand reasonable levels of condensation. If applicable, it is recommended that a simple dust/particle filter is placed by the main air intake of the DAS. Important considerations include sufficient cooling of the system due to internal heat generation and for ambient temperature. The cooling needed depends on the processor used and the workload. If the FOT is using video compression done in software, the need for processing power may be drastically higher than if only simple signals are recorded. High processor load generates significant levels of heat. The need for forced ventilation is depending on the DAS mounting position.

Too high and too low temperatures (both static and transient) do affect the DAS. Components with moving parts need special attention. Hard drives are some of the most sensitive components. Automotive grade hard drives are available, although somewhat more expensive than normal consumer hard drives, and with limited storage capacity. Flash memory cards or solid state drives (SSD) are available in increasing capacities and are clear alternatives for operating system hard drives and for storage in lower data volume FOTs.

7.5 Electrical requirements

7.5.1 Power management

The main requirement for the DAS setup is that the installation should never affect or impair normal vehicle function. *This requirement should be enforced in all environmental conditions* and is one of the most critical issues for drivers to accept equipment to be installed in the vehicles. *If FOT installations are rumoured to impede on the trustfulness on vehicle operation, few people will volunteer in any subsequent study, and the study at hand as well as subsequent studies may be compromised.*

The entire DAS installation may not draw power so that the vehicle battery charge falls below the level of being able to start the vehicle. Care should be taken so that the system does not draw power (or only minimum power) if the ignition has been turned on but the engine did not start, or if the engine stops without the ignition key being removed. State machine evaluation should be applied.

7.5.2 Interference with in-vehicle equipment

In all in-vehicle installations the aim should be to minimise or manage without the use of AC powered devices. The DC/AC and DC/DC converters are sources of electromagnetic noise and may affect both standard in-vehicle equipment (such as the FM radio), and the FOT installation itself.

Attaching any equipment to the in-vehicle bus systems has to be done very carefully. Transmitting data on vehicle buses should in most cases *not be needed or not done at all* in an FOT implementation. Failure to adhere to this might be dangerous and result in vehicle operational malfunction that may result in significant cost, injury or death, or produce other very unwanted results. Even adding only listening/eavesdropping devices should be approved by the vehicle manufacturer before being implemented.

7.5.3 Laws and regulations

Depending on the FOT study at hand different regulations will apply. In some cases CE certification of the FOT equipment may be necessary, and each project must verify what is applicable to the specific study. If the vehicles are to be driven in non-EU countries the specific regulations for each region should be verified.

For wireless communication and for some sensing systems there are regulatory restrictions on transmitting electromagnetic radiation. A few example of sensing systems that have direct regulatory restrictions are lidar and radar. The restrictions may be based on electronic interference or harm. This has to be taken into account for each individual FOT sensor setup. Still, for each instrument and jurisdiction, care should be taken to investigate the applicable regulations.

Several countries have regulations for equipment employing radar or laser technology on public roads, as these can affect effectiveness of for example authority speed surveillance instruments.

7.6 DAS data storage

7.6.1 Storage capacity estimation

The main aim of the *storage capacity estimation* is to guarantee the availability of free space for recording the vehicle data. Storage capacity depends on the following factors: number of recorded signals, sample rates of the recorded signals, sample size of the recorded signals, data collection method, driving hours, data size reduction (filtering/compression), and data deletion procedure.

Ideally, the sample rate for each signal should be the lowest possible able to guarantee no information, relevant for answering the FOT hypotheses, is lost in the sampling process. Some sample values are reported in the FESTA D2.1 for CAN and sensor data.

Using a dynamic sample rate and a high sample-rate buffer would make possible adapting the resolution of the recorded data depending on the event. The drawbacks of using dynamic sample rate are that the recording system complexity (and probability of faults) increases, more post-processing on the data will be necessary to handle the different sample rates. Database design and search becomes more complicated.

The data collection method can be continuous, or limited to specific events of interest, for example time intervals in which the lateral acceleration is above a certain threshold or in which the vehicles enters a curve with speed above a certain threshold (see section 3.5.1 in D2.2). In FOTs where only triggered storage is used and where the driver subjectively triggers in some way, it is important that past data also is recorded by using for example a ring buffer (buffering one to five minutes in the past). The level of pre-trig time will differ between projects, and should be defined based on the hypotheses for each study.

Note that events such as occurrences of safety systems warnings can be considered. The probability of such events should be part of the knowledge of the FOT designer and should be used for the estimation of the storage capacity.

Driving hours depend on the nature of the driver and the vehicle. See section 3.5.1 in D2.2 for examples.

Compression algorithms can help to reduce the data size. A lossless compression (such as zip) can be used for CAN and sensor data whereas a lossy compression (such as MPEG-4 and MP3) is normally acceptable for voice and video, respectively. The drawback of compression is that the complexity of the system increases and new possible sources of error and malfunctioning are introduced.

A safe data deletion procedure implies that no data is deleted in the vehicle until a copy of the data has been backed-up, verified, and stored in a safe place.

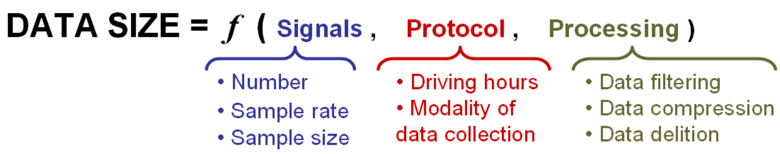


Figure 7.3: Factors influencing data size

Please refer to section 3.5.1 in D2.2 for storage capacity estimation equations, as well as data size estimation examples.

Storage capacity may be depleted and the intervals for retrieval and uploading may present some variability. These factors should be taken into account by guaranteeing enough tolerance on the final storage size. Since no space to record data would result in data loss a 20 % to 50 % on storage size tolerance is recommended.

In some studies where the levels of allowed data loss are very small it is recommended to use direct on vehicle data backups. This can be done for example by using several storage media with a data mirroring solution (e.g. RAID).

When triggered data acquisition has been chosen as the data collection method, great care has to be taken to define the trigger definitions. Even if triggered logging is used for the evaluation of effects, most studies will require baseline data. It should be possible configure and acquire baseline data, using the same DAS, also for these cases.

FOT activities that use any type of triggered data acquisition and have high data-rate data sources will need significant amounts of main memory to handle the necessary ring-buffer for pre-triggering storage. Moreover, estimations of needed storage space and data retrieval/upload frequency are affected (See section 3.5.2 in D2.2.)

7.6.2 Data retrieval/uploading procedure

Data retrieval/uploading procedures are needed to make sure that all collected data is backed-up and stored in a safe place in order to minimize data loss. The aim is to prevent data loss, verify data completeness, and to prevent data storage waste (caused by double storage).

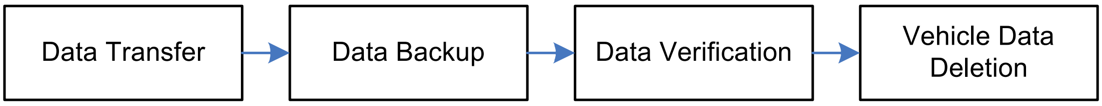


Figure 7.4: Data retrieval /uploading steps

Data transfer

Data transfer is aimed to assure that a copy of the data collected in the vehicle is stored in a safe location. *Data back-up* is aimed to prevent data loss by having a multiple set of the data stored in different safe places. *Data verification* is aimed to assure that no data was lost during *data transfer* and *data back-up*. *Vehicle data deletion* is aimed to ensure that storage space is newly available in the vehicle once the data has been safely transferred and backed-up.

Depending on the support used to record the data in the vehicle and the data size, different data transfer modes can be implemented. For a list of potential modes, transfer rates, and ranges please see section 3.5.4 in D2.2.

Generally data transfer poses two main problems: It may be time-consuming and data can be lost during the transfer. The following table presents an assessment of different transfer modes.

Table 7.1: Pros and cons for different data retrieval/uploading modes

Data transfer mode	Time efficiency	Cost efficiency	Technical complexity	Data security	Driver comfort
Data is "picked up".	+ Very fast from a vehicle point of view.	- May require to pay someone to pick up the data. - May require an extra vehicle chasing the FOT vehicles.	+ Very reliable because simple.	- Data may be misused between the "pick-up" and the final storing.	- May require the driver to go somewhere or to move the data.
Data is transmitted with wired transmission.	- May be slow for big amount of data.	- May require to pay someone to do the downloading.	- The equipment needs to be safely accessed.	+ Can be very secure if the driver is not able to access the data.	- May require the driver to go somewhere or to move the data.
Data is transmitted with wireless transmission.	- May be slow also for relatively small amount of data.	+ Download can be automatic. - May be expensive depending on the network providers.	- Automatic wireless download is very complex. - Needs hot spots.	+ Can be very secure.	+ The driver does not need to do anything.

Data back-up

Data should be backed-up and stored in a safe place as soon as it is available. Ideally, the backed-up data and the main copy of the data should be in two different safe places. In some studies where the levels of allowed data loss are very small, it is recommended to use direct on vehicle data backups. This can for example be implemented by using several storage medias with a data mirroring solution (e.g. RAID).

Data verification

Due to the potentially huge amounts of data handled, data verification is important since the probability of errors during the copying process is high.

Vehicle data deletion

Data from the vehicle should be deleted *only* once the data has been backed-up and verified.

Data loss

Experience from previous FOTs tells that data loss at the retrieval/upload stage is common even if it could be almost totally avoided with a robust and well-tested procedure. To prevent data loss during the data upload/retrieval procedure it is important to verify that data is consistent before deleting it from the vehicle. In case data is picked up and the data is not consistent, the vehicle data logger should be checked as soon as possible. Monitor the state of the data logger so that any issues can be recognised and fixed as soon as possible.

7.7 System configuration

Although DAS system configuration needs can be different, one basic requirement that apply to all FOTs is that it should be possible to find and the configuration a specific DAS after the study has been finished.

7.7.1 DAS inventory management

A system for basic inventory management is recommended for FOTs with more than a few vehicles in use. For such a system to be efficient, sensors, DAS units, vehicles and all other equipment need to be included, as well as relevant supporting procedures developed. For any one point in time it should be possible to deduce the exact hardware and software configuration of a particular installation.

7.7.2 Configuration tools and traceability

In addition to an inventory management system, it is appropriate to employ a system (and supporting management procedures) for configuration management. Software versions and additional information such as (but not limited to) software configuration files, start-up scripts, device calibration files, and other file based data needed for proper operation of the DAS should be stored together with information about the present DAS hardware configuration.

Similarly to above, it is important that traceability of software configuration changes is maintained. In this way, the exact software configuration of a particular installation at any particular time can be reproduced and reviewed.

7.7.3 Switching between configurations

For some FOTs switching between configurations in each vehicle may be necessary. In addition to having to change DriverID when a new subject is introduced to a vehicle there are a few other situations where it may be important to know who was driving or what was done with the vehicle. If this is not separated from the subjects, these trips may be classified erroneously. Examples of situations to consider are: the vehicle driven back for overhaul/maintenance, validation testing just prior to vehicle delivery, and other engineering testing.

If possible, the use of remote desktop tools over wireless (e.g. WLAN or 3G) or wired networking (Ethernet) is recommended for remote administration.

7.8 Acquisition of data

When controlling the power supply to the DAS the start-up and shutdown speeds must be optimised to reduce loss of data. Loss of data can occur both during hardware initiation when no software is started and during hardware termination when no software is able to trigger on a vehicle restart.

7.8.1 Start-up

Normally the data acquisition will start as the vehicle ignition is turned on. In order to minimise the data lost during the start-up procedure, the hardware and software must load and initiate as quick as possible. The start-up time (or the duration where data is lost) should be well monitored and documented, preferably as a property associated with each recorded trip (since it might differ with temperature etc.). Start-up of the DAS hardware shall not be done if the voltage is too low.

7.8.2 Acquisition of data

The DAS hardware should be kept powered on and running during the entire trip. To ensure that the host power system is not overexerted, the power management unit must continuously monitor the power supply and initiate shutdown if a permanent voltage fall is detected. The system must not shut down on temporary variations such as the drop during engine crank. For such circumstances, an energy reserve such as a battery may be required.

7.8.3 Shutdown

When the DAS recording has stopped, the DAS should be kept running for a short time (typically a few minutes) in case the vehicle is started again. Otherwise the DAS hardware should shut down itself as fast as possible. The power management unit should then, after a short interval, cut the power supply to the DAS, regardless of it is properly shut down or not.

7.9 Synchronisation

FOTs are considered to be used for logging of data and not as part of a hardware-in-the-loop component (for example prototype system development). This means that data from the different data sources do not necessarily have to be available for storage close to the real-world event, as long as they are individually time stamped for off-line recreation of the time-line.

7.9.1 Time stamping versus real world event

Latency is the time between when the actual real world event takes place and when the data from each respective sensor is time stamped in the logger. In most FOTs there is a need to specify the requirements for allowed levels of latency, based on the hypothesis. It is recommended to explicitly evaluate what the latency is for each data source (and for some sources individual measures) and compare this with the defined requirements (based on hypothesis; needs to be done as part of the hypothesis and performance indicator generation). If latency has been measured properly and there is limited jitter, the time stamp time can be corrected by offsetting with the latency.

Note that for data sources that are not controlled by the DAS implementers, such as vehicle CAN, it may be even more difficult to obtain the necessary information for the latency from real world event until time stamping.

7.9.2 Integrated sensing synchronisation

Depending on the methods of analysis and the implementation in the database, the needed level of synchronisation as well as the importance of measuring latency between different integrated data sources in a vehicle will differ. In Table 7.2, issues and methods for calculating the latency for some in-vehicle data sources are shown.

Table 7.2: Methods and issues in calculating latency for in-vehicle data sources.

Data source	Methods and issues in calculating latency
Video	May produce significant jitter/fluctuation on itself and other data sources. Approximate latency can be ascertained using a synchronised LED light (measured by digital I/O). Preferably hardware synchronised cameras should be used.
GPS	The latency can be calculated very precise since the GPS time is the actual acquisition time
CAN	Difficult to establish latencies for internal systems, but with reference sensors it is possible to get the latency for some measures.
Acceleration/ Yaw rate	Latency can easily be measured using a reference sensor.
radar	A lab setup with reference sensors of tracked object motion can be used

7.9.3 Synchronisation with nomadic devices

The recommended method to realise synchronisation between the different sources of data is to use GPS UTC time. Thus, all acquired data should be associated with a time stamp that is represented as absolute GPS derived time. It is recommended to use nomadic devices that have easy interface with GPS, so that the absolute time information is available.

7.9.4 Synchronisation of infrastructure systems

Also in the case of infrastructure, the UTC time derived from GPS if useful. Another possibility is to use an on-line time synchronisation service like NTP/SNTP. In the case of a traffic data and safety related FOT, data may have to be stored as raw data, accurately synchronised in time, to allow the reconstruction of the scenario in the following data analysis phases.

7.9.5 Synchronisation of cooperative systems

Synchronisation of systems with communication between vehicles can also be realised without a central infrastructure. Also here, it is recommended to use the common reference time provided by GPS. This should be easily done because many of these systems are using GPS localisation. If not, one should be added to the DAS.

7.9.6 Synchronisation with interviews and other subjective sensors

In many cases it is enough that the interviewers write date and time (hours and minutes) of the interview or questionnaire. If the subject is requested to indicate and/or comment events during the driving (for example by pushing a button), this should be time stamped when logged if possible, to enable synchronise of the event/comment with other data. The accuracy needed is in most cases less than 5 seconds. For post-hoc structured comments or questionnaires on video or events, it is important to define a process of linking the events to absolute time.

7.10 DAS status and malfunction management

7.10.1 Self diagnostics and layman feedback

To simplify laymen feedback on system status to the responsible technicians at times of system problems, LEDs or similarly externally viewable information about the system status can help and are recommended.

7.10.2 System status uploads

In order for the people responsible for DAS and sensing in the project to be able to continuously monitor the status of the test vehicles while on the road, a remote (wireless) transfer of the system and sensing status is preferred. (See section 7.14.1 for details.)

7.10.3 Malfunction management

A process for identification of problems in the vehicle, contacting the driver and exchanging sub-systems or the entire vehicle should be developed.

7.10.4 Spare system management

For quick and efficient problem solving it is recommended to keep spare parts for all components pertaining to the DAS. It has been suggested that a 10 % extra number of spare parts should be kept within the FOT, and that this number be added specifically when estimating the total project cost. All spare parts are to be managed by the inventory management system. Testing and calibration of spare parts should preferably be planned and performed within the process. Supporting management procedures need to be developed as well. At least one fully equipped DAS system should be kept “on the shelf”; prepared and calibrated for immediate use.

7.11 System installation

7.11.1 Installation procedures

Before initiating the installation procedures, an installation specification document shall be prepared. The installation specification must in detail describe how each component of the system shall be installed. Specifically, the installation specification shall provide solutions to the following topics:

- Mounting: positioning, means of attachment, accessibility, safety and security;
- Cabling: dimensions, shielding, drawing, mounting, tolerance and labelling;
- Connectors: soldering/pressing, robustness, impedance and labelling to avoid mix up;
- Power supply: consumption, fuse, voltage, source and switching;

- Environmental endurance: effects on electromagnetic disturbances (EMC), temperature, humidity, vibration, shock, electric safety and dirt.

It is of great importance that the FOT system installation is adapted to the requirements set by all other systems in the vehicle. If this is not done properly the installed system could generate disturbances that might void warranty of the original systems – or even an entire vehicle. All systems that possibly may be in conflict with the installed system need to be identified. An adaptation plan must then be developed for each system to ensure that they will be able to operate properly after installation.

The actual installation work needs to be done by operators that are authorised to work on the actual host system, and during the installation work, all changes to the host system (if any) must be documented in detail. Depending on the study and region, the authorising body could for example be the OEMs, insurance companies (voiding warranty etc), the project and/or a legal entity.

If several vehicles are to be installed, it is recommended to select one vehicle as prototype. The prototype installation will then revise the installation specification continuously during the work.

7.11.2 Installation verification and calibration

When the system is installed it needs to be verified and calibrated before the data acquisition starts. The verification will refer to the installation specification and verify that all requirements are met. Monitoring of all potential sources of interference so that no conflicts are caused is important.

To ensure data validity and quality a calibration and verification scheme is recommended. For data quality aspects it is important that all installed systems of the same category are calibrated and verified using the same procedures. During the verification process a full dataset should be recorded for the analysts and quality management team, in order for them to verify that the installation adheres to the analysis requirements.

7.11.3 Dismounting the system

When the data acquisition is finished and the system is to be uninstalled, the installation documentation shall be used to ascertain that the host system is restored to its former condition. Finally, all proprietary data need to be removed from the FOT system before it is disposed or reused in future projects. Remember to include dismounting costs in the FOT budget.

7.12 Proprietary data in FOTs

The concerns regarding proprietary data are to keep the CAN/LIN/MOST specifications OEM-confidential and to hide the actual system design to prohibit reverse engineering based on data collected within an FOT project. Regarding the first issue there are two cases to be distinguished.

- When the OEM is strongly involved in the data acquisition process during the FOT execution, the confidentiality of the CAN/LIN/MOST specification is not an issue within the project.

- When the OEM does not handle the data collection by himself, the usage of CAN gateways is proposed. The CAN gateway has to be programmed by the OEM to provide the data from the CAN/LIN/MOST bus according to the agreed logger specification.

The second issue – reverse engineering of functions and systems – is also an issue within the FOT project. Each project will have to handle this and define what is needed. In some cases it may be necessary for the OEMs to handle detailed low level data and aggregate it on a certain level before it is provided to the project partners responsible for data analysis. A general recommendation to future FOT projects is to define in advance, *what level of system data is needed to answer a specific research question and whether the involved OEMs are able to provide this data to the project*.

In some FOTs OEMs might be interested in the acquisition of additional data, which is not directly related to the project and proprietary to the OEM. This should be allowed. The OEM could separate the additional data from the project data before the data is provided to the further project for analysis.

7.13 Personal integrity and privacy issues in data acquisition and analysis

Recommendations for the definition of necessary legal arrangements depending on the specific FOT are not covered here. See chapter 3 for further information.

Different levels of data security should be implemented in order to cover personal and privacy issues properly. The data access right of a project partner should depend on his specific role in the project.

Several different levels of data security should be implemented. Driver data allowing direct conclusions about the identity of the driver is considered to have highest requirements regarding data security. Video data of the driver’s face and GPS data are typical examples of the data that belongs to this category. Some types of metadata like the car serial number also belong to this category of data. For this kind of data anonymisation via monikers is required before upload into a database.

It might also be necessary to implement a GPS data based control, which deactivates the video recording, when required (e.g. when driving in countries with specific legal requirements).

7.14 On-line quality management procedures

In all FOTs assuring data quality in the data collection and data management process is very important. The procedures for data quality assurance before, during and after the data collection should be well defined. Specifications and plans should be written for each individual FOT

It is recommended that a quality management team is appointed for each individual FOT with roles such as: daily quality overview, OEM contact person, subject contact person, DAS and sensor maintenance person, and vehicle maintenance person.

7.14.1 Remote automatic upload

In most state-of-the-art FOT wireless transfer of vehicle and data status has been used, in order to assess the status of vehicle, DAS and data without having to physically access the vehicle when the vehicles in the study are on the road. Different transfer techniques such as simple text messages (SMS) or GSM/3G, have been used for status uploads. A maximum delay from the time of actually collecting the data until it has been analysed for quality and status should be defined. Otherwise the project risks that the vehicles on the field are potentially not collecting the required data. The maximum delay should be set based on the accepted levels of data loss and the length of the study.

For a thorough listing of example variables/measures that may be of interest to store in the vehicle DAS summary files for per trip data upload, see section 4.1.1 in D2.2.

When the data has been uploaded and put into a database trip statistics can be calculated per vehicle or driver. It is recommended to use this to identify extreme/abnormal driver/usage behaviour early in the study so that if necessary drivers can be exchanged. This driver monitoring early in the study is highly recommended so that the study schema of the specific FOT is kept.

If an FOT is to be executed across country boarders and include roaming for the wireless services, investigation of the cost/benefit of using the quality data upload systems outside of the “home country”, should be made.

7.14.2 Automatic and manual quality checks

It may be tempting to do the quality assessment fully automatic, but state-of-the-art FOTs have indicated that by doing this you risk contacting the driver in cases where the error or anomaly is not significant for the study. Due to this, it is recommended that the quality assessment should be set up in different steps and that before (if) employing a fully automatic system, the algorithms for the assessment should be thoroughly validated. An automatic thresholds based warning system can be applied for some hard and very important measures. A tool for maintaining the warning system should preferably be checked by a one responsible person each day.

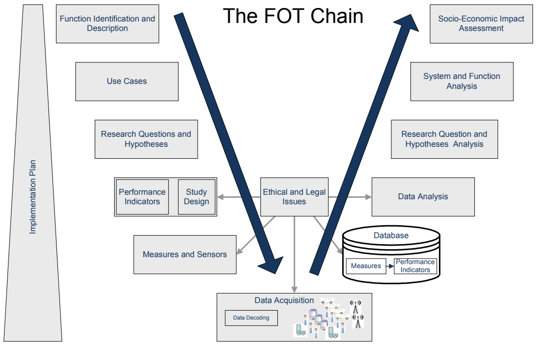
For a description of a process for on-line data quality checking, see section 4.1.2 in D2.2.

7.14.3 DAS and sensor maintenance

It is recommended that a specific subject/driver only has one contact person throughout the study. The process for contacting the driver should be clear and preferably there should be a list available within the project with contact information for the contact-responsible for each vehicle. All communication with the driver should then be at least initiated via this person.

For a description of a process for fixing DAS problems/maintenance, see section 4.1.3 in D2.2.

8 Guidelines for FOT Databases and Analysis Tools



As state-of-the-art FOTs has proven that various types of studies demand different data models and hardware specifications, this section will not describe a generic solution for all types of FOTs.

The following sections will focus on database for large FOTs (where thousands of hours of raw data are collected). Such a considerable amount of data, especially if video data is also collected, will test hardware to the limit. A smaller FOT might still use the guidelines in this chapter and apply them to a less complex database.

8.1 Database design and implementation

8.1.1 Possible database models

Since the FOT specific demands make it difficult to recommend one generic database model, two main possibilities are here recommended. (See Appendix 2 in D2.2 for example database schemas and for further information).

Database model 1

Strategy: A measure equals a table column in the database. To avoid keeping all columns within the same table, splitting the data into several main table(s) and sub-tables is necessary.

Pros: Since measures are split into different tables there will not be too much data in each table which makes it easier from a database performance point of view.

Cons: The design may end up with lots of database tables and it may be difficult to keep track of where measures are stored. In addition, it will not be possible to activate/deactivate measures during the study.

Database model 2

Strategy: Instead of storing data in different tables, a *value table* can be created to store all values. Instead of using a column as the reference, measures are stored as rows in a measure table.

Pros: The measures as can be used as references for other tools (such as a quality assurance system), and that measures can be added/removed or activated/deactivated easily.

Cons: The value table can become huge limiting the database performance. For example, in order to store for 1000 hours of data, 200 measures at 10 Hz, 7,200,000,000 database rows will be needed. Different databases deal with this issue in different ways.

8.1.2 Data filling in the database

Beside the sensor and video data collected by the DAS, other data needs to be transcribed in the database in order to prepare the database for data analysis.

Trip ID and time data transcription into the database

Trip ID and *time* are cornerstone indices in the database designs. This means that sensor data must be time stamped when inserted in the database. State-of-the-art FOT relational databases use a common sample rate to ensure the validity of trip ID and time. If different sample rates are needed within the same FOT database, the different datasets should be organised into different tables. Furthermore, data with frequency differing from the default one (e.g. 10 Hz) should be clearly marked as potentially incompatible with the main data. If the need to join this data with the main data arises, it is suggested that data from the deviant table is extracted, resampled and inserted into a table with a common sample rate.

Transitional data transcription into the database

Transitional data can be stored separately into tables that only contain data when transitions occur. Despite the potentially huge storage overhead, the trend is to handle transition data the same way as measure data to simplify analysis. *Events data* can be described as shortcuts or pointers to specific events within the database. It is up to each individual FOT to define what an event is and the algorithm that defines it. *Manual annotations* are another way to create pointers to events in the database.

Background data transcription into the database

There are two types of *background data* which should be stored in the database model: 1) the driver's and 2) the vehicle background data. Driver data should be stored in the *driver* table but any data to identify the driver should be kept securely and separately (see Chapter 3).

Subjective data transcription into the database

To reduce errors, automatic transcription of subjective data is always preferable. Transcription of audio/voice messages is recommended.

Time history data transcription into the database

Most of the data in a typical FOT are stored in time history tables. For the *database model 1* (see description above) it is important not to create tables with too many columns. When using the *database model 2* (see description above)*,* specific database tools and functions could be considered. Examples of these database tools and functions are: table partitioning, block compression, index pre-definition.

Events/classifications transcription in the database

An FOT database can consist entirely of events if a triggered data collection approach is adopted (as opposed to continuous data collection; see Chapter 6). In other FOTs, where data collection is continuous, the ability to find and classify events of interest is of central importance. Classification and use of "events" (classified time periods) is an important aspect of FOT analyses (see Section 7.3.10). Some events are straightforward and simple to identify, for example hard braking defined as peak deceleration > 0.7 g, and may not need to be saved as a discrete or transition variable. However, many events involve a considerable amount of effort to find and validate and are worth saving into a discrete variable database or index to facilitate data query and analysis. Event pointers should be saved to speed up analyses and can be used in combination to describe more complex situations with multiple events (see chapter 6).

Adding tables to the database

Tables are initially created manually based on information in the sensor matrix (see chapter 5). Any change to the database design must be documented thoroughly.

8.1.3 User data spaces and data sharing

It is vital to keep track of where data is created and manipulated. A copy of the collected raw data should be kept as original data in a read only space to prevent accidental data loss. In addition, it might only be the person responsible for data uploading who can insert new data and the FOT database owner who can delete data. As the analysis work begins, there will be the need for the analyst to store new data (coming from combination or processing of the raw data; see Chapter 6) in a private user space. If this new data is also relevant for other users of the database, a solution to share this data in a project internal space should be implemented. The approval process to share data should be described and basic meta description of the data is needed (as data origin and function/method/algorithm applied to the data). Some of the data could be of public interest and therefore exported or accessible via web interface. Although sharing this data must be approved from all stakeholders in the FOT and/or on an aggregated level.

8.1.4 Hardware and storage

It is important to consider that the database server will rest inactive during data collection and then run at 100 % of its capacity during analysis. If the supporting organisation can provide flexible solutions, such as server virtualisation and/or clustering, the FOT study access to the database when running analysis on the data can be prioritised. When the project ends, machine usage can be set to a minimum until a subsequent study needs to use the data.

Very fast and reliable disks can be used even with a limited budget. In most cases storage at some kind of disk cabinet, NAS (Network Attached Storage) or SAN (Storage Area Network) is most appropriate. A storage setup with some kind RAID configuration should be considered, in order to be better prepared if a disk crashes or some data blocks are corrupted. The database should use faster disks than the file server and using disk cache is recommended to increase the performance of the system.

8.1.5 Risk management

An FOT study can generate huge amounts of data; especially when video is used, and the management must decide on the need for backup and acceptable downtime for recovery of the FOT database. It is up to the steering committee of the study to have a documented backup policy and crash recovery strategy. Further, the backup strategy might need to vary during the lifecycle of the study (collecting phase, analysis phase). If so, each phase and strategy should be documented. Disaster recovery (when local database and backup hardware are destroyed) strategy must be taken care of and there should be an offsite backup of the data.

8.1.6 Database and data storage implementation

Database

Storage of all data but video should be stored in a relational database, supporting ANSI SQL.

This implementation must consider what to do with data loss from a sensor. Various strategies can be employed: if a sensor gives no data, a NULL value can be inserted. State-of-the-art FOTs suggest that using the last known sensor value makes analysis easier. The problem with data that is actually not valid has to be dealt with (see Chapter 6).

Video data storage

By looking at the state-of-the-art for FOTs, two ways of storing video data have been identified: either 1) video files can be stored on a file server, or 2) the video can be stored directly in the database. In the latter case each frame can be stored as a JPEG image as a BLOB (Binary Large Objects). Another option is to store the complete video as a binary file as a BLOB.

Hardware

If video is not stored within the database, it is recommended to separate the database and video file server in order to configure the hardware individually. Outsourcing system operations is possible; however, the costs for network bandwidth, backup, and administration can be very high.

Distributed system at various locations

It is strongly recommended that the database is not distributed. For the database, use a single common database. For video storage, also other options can be considered (see section 5.2.4 in D2.2).

Physical access

Physical access as well as the approval process for access to the hardware must be documented.

Logical access

Logical access as well as approval for access to the database must be documented. A role-based access is advised when any user to a certain role of the database obtain certain access. This also applies for the supporting operating system. Any FOT must define the roles and permissions of the database. These roles can be:

- *Database administrator (DBA):* Unrestricted access to the database.

- *FOT database owner:* Unrestricted access to FOT database data and permitted distribute role access to users.
- *Uploader:* Allowed to insert and update data in the FOT database.
- *Analyst:* Allowed to read data from the FOT database and to manipulate data in private user space.
- *Publisher:* Permitted to insert/update/delete data in shared user space.
- *Web application:* Permitted to read data from specific user space containing aggregated data.

Personal integrity and sensitive data

Driver data must be stored according to the access restrictions defined by the steering committee. In a collaborative study, some data may be classified as sensitive by one partner or even by a supplier of measurement equipment.

Private vs. public data

Private data should be kept in a private "user space" (database or schema), in order not to risk inadvertent confusion with original project data.

Backup

An FOT database backup strategy should be based on "acceptable downtime". Off site backups are mandatory for managing a disaster scenario. The majority of the data is never edited (video and raw data in the database) and data mirroring should be sufficient. For data created by private, organisation, or public user spaces, a daily backup strategy should be applied.

Video data (file server)

Please refer to section 5.2.9 in D2.2 for a list of potential standard backup solutions.

Database data

The backup policy must be based on the time it takes to recover data and the acceptable loss of data. Even though some studies may use the original logger data as backup, any private or published data created afterwards must have valid continuous backups.

Database acceptance

Before an FOT is launched the FOT database architecture should be reviewed by a system evaluator to ensure that all requirements are fulfilled and to verify policy documents.

8.2 Off-line quality management procedures

8.2.1 *Quality assurance of objective data*

Quality assurance before data is uploaded to database

Before uploading objective data from a vehicle, a well-defined algorithm should be applied to all the data in order to verify data consistency and validity.

Quality assurance of video data

To catch problems with camera failure or other video related problems, a video checking strategy should be implemented. A tool for viewing one or several images per trip can be useful. Moreover, a function to verify at least the size of video files is necessary — the size is somewhat proportional to recording duration.

Driver ID verification/input

Again, it may be necessary to have a process that allows the analysts to view, for example, one image per trip and match this with the IDs of the drivers allowed to drive a specific vehicle. If a driver is unknown, then the data for a particular trip may have to be neglected. A software tool for doing this manual identification of drivers is preferable. Be advised that some eye trackers (if available) provide DriverID functionality.

8.2.2 *Quality assurance of subjective data*

In order to address the validity of the data, the formulation of the questions (and possible answers) is a key issue, especially when designing a questionnaire to be distributed to respondents. Questions must evidently be formulated in a clear and unambiguous way. In addition, questions must also, e.g.: be specific, not too complicated, be formulated in simple terms that can be understood by the interviewee. Hypothetical questions are the most difficult questions and should be avoided.

Regardless of data source, missing data is a threat to the quality (see Chapter 9). In the case of a missing questionnaire, efforts must be made to ensure that data collection is as complete as possible and reminders must be administered. Furthermore, the number of questions should be thought through, in order to limit the number of questions. In addition, the number of open questions should be as few as possible in order to reduce the effort of the respondents. The interviewer plays an important role in collecting data in an interview situation. Interviewer bias, that is the influence of the interviewer on the respondents' response, can be avoided by administering a questionnaire. However the interviewer may also increase the quality of the data collected by, for instance, answering to questions and using probing questions.

8.2.3 *Measures naming guidelines*

It is recommended that the FOT project decides on and adheres to a set of naming conventions for measurements. The strategy used should be well documented and thoroughly enforced. Motivations for a clear naming convention include: 1) project-wide consistency, 2) clarity for direct understanding of used measures in analysis, 3) differentiation of non-comparable measurements, and 3) avoidance of confusion.

When specific measurements are named, references to the following measure attributes are recommended: *indicative name*, *associated source*, *sample rate*, and any other *FOT specific descriptor*. The compounds should be joined consistently to create a single word. Possible strategies are: "camel case" (SomeSignal), underlines (some_signal), or hyphens (some-signal). Depending on context and FOT specific requirements all or only a subset of the compounds can be used.

Examples: [GroundSpeed_GPS_1Hz], [GroundSpeedGps1Hz]

The aim is to clearly understand what a measurement "is", where it comes from, and how it relates to other measurements. To avoid the risk of making faulty comparisons, measurements that are *non-comparable* should be named *differently*.

8.2.4 Automatic pre-processing and performance indicator calculations during data upload

It is recommended to define procedures and implementation schemes on *how* to add calculation of pre-processing and performance indicators in the upload process (see Chapter 9). These calculations should preferably be read-only for the users. The actual algorithms for the pre-processing and performance indicator calculations in this step have to be well defined and tested (on for example pilot test data), or based on previous experience. Since the estimation of some specific performance indicator may set specific requirements on the raw data (see Chapter 5), these constrains have to be taken into account when implementing the automatic pre-processing.

8.3 Data analysis tools

The focus of this section is to describe analysis tools, not to describe analysis procedures or methods.

8.3.1 Data classification/transcription

The exact coding scheme/syntax for events (time segments) will vary widely across FOTs. However, the following features have been identified as important software functionalities:

- organising or categorising subjects into groups and subgroups;
- defining any set and structure of codes, and associating software buttons and keyboard keys to each category;
- editing or updating the coding scheme;
- defining events as either a *state* event (e.g. glance left, glance right) or a *point* event (e.g. stop light);
- defining if state events are mutually exclusive or start/stop and set a default state;
- defining if codes are nominal (e.g. road types) or rating scales (e.g. observer ratings of drowsiness);

- defining if codes are compulsory or optional, logging freely written comments created by human analyst (no coding scheme); and
- support for inter- and intra-rater reliability analyses.

8.3.2 Time history and event analysis with video

This section describes the basic functionality of tools for viewing numerical time history data and the associated environment sensing data which includes video data, map data (e.g. GPS), and traffic state data (e.g. radar).

Recommended functionalities for visualisation and interaction with data

- replay single-participant data (numerical time-history data, video data, map data, and traffic state data) simultaneously. Multiple windows for different plots and illustrations provide maximal flexibility to arrange and resize is often spread out on multiple computer screens. Recommended visualisation functionality:
    - Video recordings synchronised with other raw data plots;
    - Continuous variables and performance indicators which can be plotted (and zoomed) on graphs;
    - General information (FOT reference, subject ID, event lists, etc.);
- aggregate and visualise multiple participants' data at once to compare flows of events.

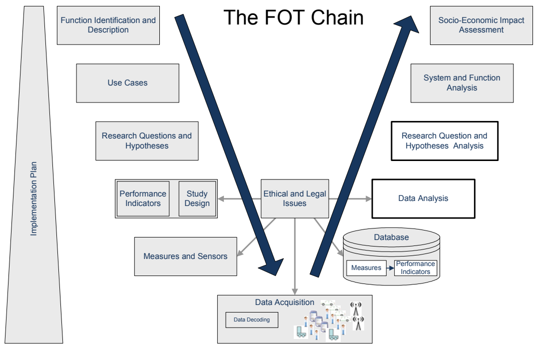
Recommended functionality to support data analysis

- database query functionality (e.g. SQL)
- signal processing of numerical data (see also Chapter 8)
- fully customizable mathematical computation, analysis, and algorithm development functionality, automatic or semi-automatic calculation of performance indicators, and application of trigger algorithms to find events of interest (e.g. lane changes, near crashes, jerks)
- image processing of video data (e.g. machine vision algorithms to detect traffic signal status)
- grouped analysis of data (e.g. scripts)
- export results function to tabular format or statistical packages.

A general recommendation for an analysis package is to use SQL software for database queries, mathematical analysis software for computation (such as Matlab), and common statistical software packages (such as SPSS). If huge datasets have to be analysed or more specific requirements exist, then more specialised or proprietary solutions may be necessary. SQL and some software tools may require a fairly high level of knowledge to use, so it may be advantageous to develop proprietary easy-to-use graphical user interfaces.

9 Data Analysis and Modelling

9.1 Introduction



The strategy and the steps of data analysis need to be planned in order to provide an overall assessment of the impact of a system from the experimental data. Data analysis is not an automatic task limited to some calculations algorithms. It is the place where hypothesis, data and models are confronted. There are three main difficulties:

- the huge and complex amount of data coming from different sensors including questionnaires and video, that needs to be processed;
- the potential bias about the impact of the system(s) on behaviour which may arise coming from sampling issues including location of the study, the selection of a relatively small sample of drivers, etc.;
- the resort to auxiliary models such as simulation models to extrapolate from the behavioural effects estimated and tested within the sample to effects at the level of the whole transport system.

To be confident of the robustness of the outputs of the data analysis, one has to follow some strategic rules in the process of data analysis and apply to the whole chain and to its five links (Figure 9.1) the required techniques such as applying appropriate statistical tests or using data mining to uncover hidden patterns in the data. [4]

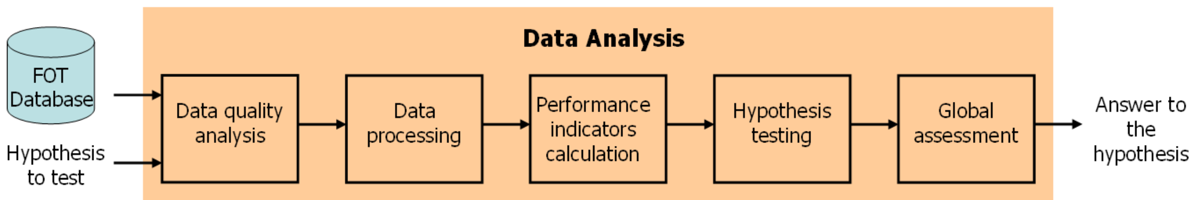


Figure 9.1: Block diagram for the data analysis

Some specific actions are required to tackle the difficulties mentioned above and to ensure the quality and robustness of the data analysis:

1. A pilot study is a prerequisite to check the feasibility of the chain of data collection and treatment and to achieve a pre-evaluation of the usefulness of the system.

[4] For more detailed information, the reader should refer to FESTA deliverable D2.4.

2. The data flow has to be monitored in detail but also overall. One of the strategic rules to follow is to ensure local and global consistency in the data processing and data handling and analysis.

3. The sources of variability and bias in the PIs have to be identified, where feasible, in order to control for them in the data analysis.

4. There is a crucial need for an integrative assessment process which should ideally combine within a meta-model information gathered on the usability, usefulness and acceptability of the system with the observed impacts of the system on behaviour. The estimated effects obtained from the sample of drivers and data have to be extrapolated using auxiliary models to scale them up.

5. Appropriate techniques have to be applied for each link of the chain — data quality; data processing, data mining and video analysis; PI calculation; hypothesis testing; and global assessment. The techniques come from two set of statistical and informatics tools belonging to two main kinds of data analysis: exploratory (data mining) and confirmatory or inferential (statistical testing).

9.2 Consistency of the chain of data treatment

There will be a lot of computations and data flows starting from the measurements collected into the database through estimation of PIs to the testing of hypotheses and on to the global assessment. This process, in the form of a chain of operations, has to be monitored in detail but also overall. There are five operations linked together in terms of data treatment. In addition, three kinds of models are needed as support to carry out the three top operations: probability models for justifying the calculations of the PIs, integration models to interpret in a systemic way the results of the test, and auxiliary models to assess the effects on a larger scale (scaling up). Moving from the data to an overall assessment is not only a bottom-up process; it also has to include some feedbacks (Figure 9.2). There are two movements along this chain: a data flow going up and a control feedback loop from the top which concerns the consistency of the evaluation process and which mainly depends on control of uncertainty.

In moving up the chain, the consistency of each operation can be checked locally according to the specifications which are governed by the nature of the PIs which correspond to a set of hypothesis related to the use cases of the system. For each PI, there are some rules which ensure the validity of the calculation procedures. For example, it is important to sample data which can change rapidly at a high data rate. The sampling rate must fit the variability of the variable. From a database design point of view, however, it may be easier to collect relatively static data at a high frequency.

As a complement to local consistency, a global criterion is to have sufficient sample size to get enough power to carry out the test of a hypothesis or to make an overall assessment with enough precision. This is a feedback loop coming from the top to control the uncertainty of the estimations. The precision required for measurements depends on the uncertainty of the auxiliary models, of the regression models and of the probability models.

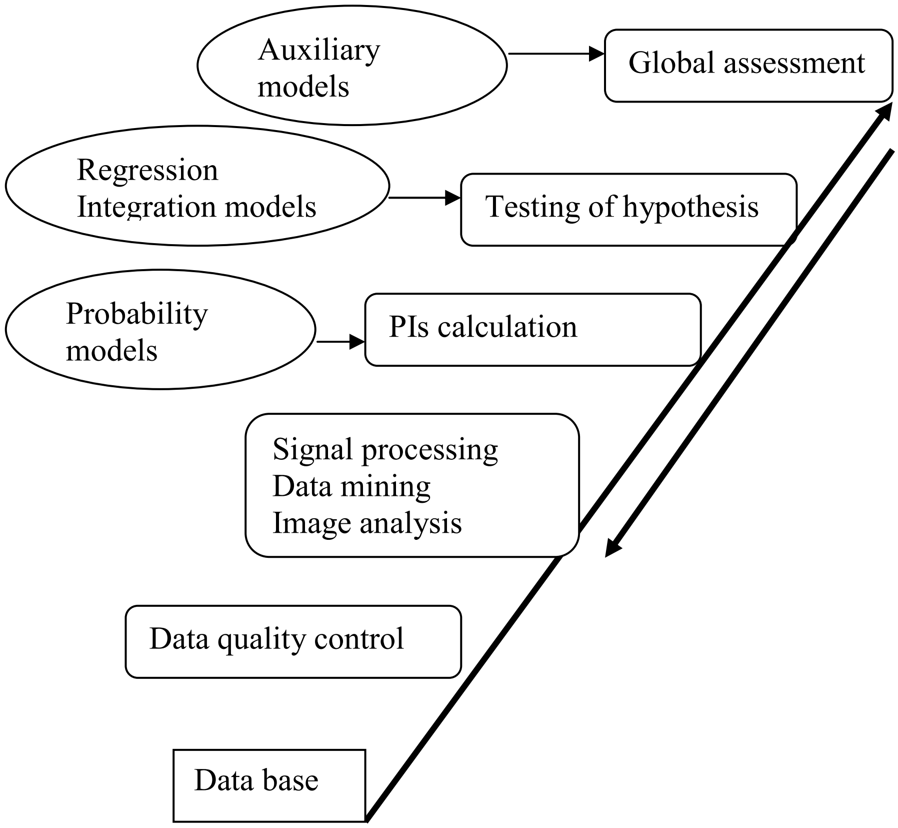


Figure 9.2: Deployment of the chain with feedbacks and additional models

9.3 Precision in sampling

The aim is to measure the effect of an intervention or treatment — which in the case of an FOT is the use of a system or systems — on a sample of subjects and in various driving situations while controlling for external conditions. From the sample, we have to infer the effect on the population by aggregating the values obtained through the sensors without and with the system to get an estimate on the effect on the chosen PI. How to insure that this inference is valid, in other words that the estimation is very near the true effect in the population? The precision of the estimate depends on the bias and variance which could be combined to get a measure of the sampling error (Wannacott and Wannacott, 1990). To control the bias and variance, one has to rely on a well defined sampling plan using appropriate randomisation at the different levels of sampling: driver, driving situation and measurement.

Consideration should be given to identifying the possible sources of (unintended) bias and variance in the sample and either attempt to minimise or account for these in the data analysis. This is one of the most fundamental principles of statistical methods.

1. Driver variation. The simple fact of the matter is that drivers vary. The range of behaviours that drivers exhibit (in terms of speed selection, headway preference, overtaking behaviour) is immense, but fortunately the variation obeys some probability laws and models. Strict randomisation procedures ensure that only the

outcome that is being varied (or the outcome whose variation we are observing) is working systematically. However, strict randomisation is not usually possible or desirable[5] in an FOT, particularly when the sample sizes are relatively small. The theoretical best method is to stratify the population of drivers according to some variables or factors related to the outcome and to sample proportionally to the size of the sub-population and to the a priori variance of the outcome (e.g. speed choice). For practical reasons, a different sampling or selection procedure may be followed. In either case, it is important to be able to compare the sample to the overall driver population in order to identify what are the main discrepancies and to assess possible sources of bias.

2. Driving situation variation. There will be variation within and between the journeys and the driving situations within these journeys. For example a particular journey may be affected by congestion part-way through, or weather conditions may change from day to day. This type of variation cannot be controlled and is considered to be random. The observation period should be sufficiently long to allow for these random effects. One example here is that seasonal effects should be considered.

3. Measurement variation. Once in a driving situation, by means of the sensors, we get a series of measurements at a certain frequency. Their size is not fixed but varies. Each set of measurements within a driving situation constitutes a sample of units taken from a cluster, according to sampling theory. Usually, there is a correlation between the measured outcomes. The information coming from this sample of measurements is not as rich as expected from an independent sample. One such cluster is at the driver level — the data collected from one driver is not independent.

How to quantify the variance of the estimate of an outcome from the experimentation taking into account theses three sources of variations? The total variance of the average of the indicator on the sample breaks down into an inter-individual, intra-individual and infra-situational variance. If the inter-individual variance is strong, an increase in number of situations observed and in the measurement points per situation will not bring any precision gains (Särndahl et al., 1992). However, it may help to ensure a reduction in bias from, for example, seasonality.

9.4 Requirements for integration and scaling up

Having treated and aggregated the data by means of statistical models, there are two kinds of problems to solve related to first the synthesis of the outputs and second to the scaling up of the results from the sample to a larger population. Integration of the outputs of the different analysis and hypothesis testing requires a kind of meta-model and the

[5] It may not be desirable, for example, to waste sample size by recruiting drivers who only drive small amounts each week. Many FOTs have for good reasons used a quota sampling procedure, in which equal numbers of (say) males and females are recruited. This can create bias when scaling up the observed data to estimates of effects at a national or European scale.

competences of a multidisciplinary evaluation team (Saad, 2006). Scaling-up relies upon the potential to extrapolate from the PIs to estimates of impact at an aggregate level.

It is often necessary to employ quantitative models from previous studies to estimate the effect of indicator in question. It is, however, important to note that individual models have usually been developed for particular purposes, from particular data and with specific assumptions. However, in the absence of appropriate models available for the purpose of study, it is usually necessary to apply the “least bad” model available with appropriate weighting or adjustment.

It is also important to consider the constraints, assumptions, and implications behind the design of the study in mind when interpreting the analysis results. Behavioural adaptation may lead to side effects (i.e. indirect effects) and also result in a prolonged learning process. However, the study period may, for practical reasons, not be sufficiently long to fully explore this.

Extrapolating from the sample to the population depends on the external validity of the experiment. The power of generalisation to the population of the estimates of impact is related to their precision which is composed of two parts — bias and variance. We can use three approches:

1. If the required performance indicator is available in the sample (e.g. if journey time is an impact of choice for efficiency and journey time has been collected), the impact at the population level can be calculated directly, although sometimes a correction factor or other form of extraplotaion adjustnment may have to be introduced (Cochran, 1977).

2. If neither a performance indicator nor a proxy indicator is available, then it is necessary to adopt an indirect approach through models which provide an estimate of the output from the behavioural PI estimated from the sample. Speed changes can be translated into changes in crash risk by applying statisticaly derived models from the literature which have investigated the relationship between mean speed, speed variance or individual speed and crash risk. Emissions models can be used to calculate the instantaneous emission of a car as a function of its recorded speed and gear selected.

3. Finally a macroscopic or microscopic traffic simulation model can be applied to translate the effects observed in the sample to a network or traffic populations effect. The outputs from such a simulation can for example, be used to calculate journey time effects or fuel consumptions effects at the network level.

9.5 Appropriate techniques at the five links of data analysis

The five links follow the right branch of the development process of a FOT from data quality control to global assessment. Different techniques of data analysis and modelling which could be used at each step are presented here.

Step 1: data quality analysis

Data quality analysis is aimed at making sure that data is consistent and appropriate for addressing the hypothesis of interest (FESTA D3, section 4.5). Data quality analysis starts from the FOT database and determines whether the specific analysis that the experimenter intends to perform on the data to address a specific hypothesis is feasible. Data quality analysis can be performed by following the four sub-steps reported below (and shown in Figure 9.3). A report detailing the quality of the data to be used to test the hypothesis of interest should perhaps be created.

The sub-steps for data quality analysis are:

1. **Assessing and quantifying missing data** (e.g. percentage of data actually collected compared to the potential total amount of data which it was possible to collect).

2. **Ensuring that data values are reasonable and units of measure are correct** (e.g. a mean speed value of 6 may be unreasonable unless speed was actually recorded in m/s instead of km/h).

3. **Checking that the data dynamic over time is appropriate for each kind of measure** (e.g. if the minimum speed and the maximum speed of a journey are the same, then the data may not have been correctly sampled).

4. **Guaranteeing that measures features satisfy the requirements for the specific data analysis** (e.g. in order to calculate a reliable value of standard deviation of lane offset, the lane offset measure should be at least 10s long; additionally, this time length may depend on the sampling rate — see AIDE D2.2.5, section 3.2.4).

Please, notice that the first three sub-steps refer to general quality checks; thus, if any of these fails, data analysis cannot proceed. If a failure is encountered, it should then be reported to those responsible for the database responsible so that the possible technical error behind can be tracked down and solved. However, the last sub-step is different, and is related to the specific analysis or to a specific performance indicator to be used in the subsequent data analysis. As a consequence, if step 4 fails, it may not be due to a technical issue that needs to be solved, but to intrinsic limitations in the collected data.

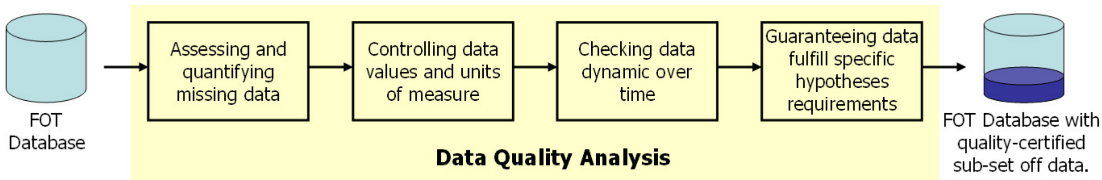


Figure 9.3: Block diagram of data quality analysis

Data quality analysis is handled differently with regard to data from in-vehicle sensors (generally CAN data and video data) and subjective data (generally from questionnaires). Subjective data, once collected, is hard to verify unless the problem stems from transcription errors.

Step 2: data processing

Once data quality has been established, the next step in data analysis is data processing. Data processing aims to "prepare" the data for addressing specific hypothesis which will be tested in the following steps of data analysis. Data processing includes the following sub-steps: filtering, deriving new signals from the raw data, event annotation, and reorganization of the data according to different time scale (Figure 9.4). Not all the above-mentioned sub-steps of signal processing are necessarily needed for all analyses. However, at least some of them are normally crucial.

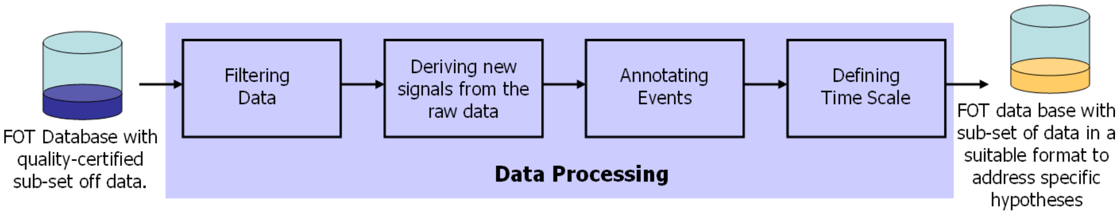


Figure 9.4: Block diagram for the procedure of data processing

Data filtering can involve a simple frequency filter, e.g. a low-pass filter to eliminate noise, but also any kind of algorithm aimed at selecting specific parts of the signals. Very often a new signal more suitable for the hypothesis to be tested has to be elaborated by combining one or more signals. Marking specific time indexes in the data, so that event of interest has been recognized, is fundamental to individuate the part of data which should be analyzed. Ideally, an algorithm should be used to go through all FOT data and mark the event of interest. However, especially when the data to be annotated is from a video and requires the understanding of the traffic situation, writing a robust algorithm can be very challenging even with advanced image analysis techniques and manual annotation from an operator may be preferable. Re-organizing data into the most suitable time scale for the specific hypothesis to be addressed has to be considered in the following steps of the data analysis.

Step 3: PI calculation

There are five kinds of data which provide the performance indicators: Direct Measures, Indirect Measures, Events, Self-Reported Measures and Situational Variables. The scale of the dataset and the uncontrolled variation in driving situations that occurs from driving freely with vehicles become a seriously limiting factor unless efficient calculation methodology is implemented. The choice of which PIs and hypotheses to calculate is clearly dependent on the amount of effort required. Efficient calculation methods need to anticipate that (a) PIs will be calculated on imperfect data – there is a strong need to create special solutions for "exceptions to perfect data", and (b) PI calculation requires situation or context identification — a "denominator" or exposure measures to make a measure comparable is required to determine how often a certain event occurs per something (e.g. km, road type, manoeuvre). The fact that test exposure is largely uncontrolled (not tightly controlled as in experiments) means that analysis is largely conducted by first identifying the important contextual influences, and then performing the analyses to create a "controlled" subset of data to compare with.

The ability to find and classify crash-relevant events (crashes, near-crashes, incidents) is a unique possibility enabled by FOTs to study direct safety measures. This possibility should be exploited by using a process of identification of critical events from review of kinematic trigger conditions (e.g. lateral acceleration >0.20 g). The definition of these trigger values and the associated processes to filter out irrelevant events are of particular importance for enabling efficient analyses.

Care should be taken to use appropriate statistical methods to analyse the PIs. The methods used must consider the type of data and the probability distribution governing the process. Categorical or ordinal data, such as that from questionnaires, needs to be analysed appropriately. Data on the degree of acceptance of a system (e.g. positive, neutral, negative) can be applied in multivariate analysis to link it to behavioural indicators so as to create new performance indicators.

Step 4: hypothesis testing

Hypothesis testing in a FOT generally takes the form of a null hypothesis: no effect of the system on a performance indicator such as 85th percentile speed, against an alternative such as a decrease of x % of the performance indicator. To carry out the test, one relies on two samples of data with/without the system from which the performance indicator is estimated with its variance. Comparing the performance indicators between the two samples with/without intervention is done using standard techniques such as a t-test on normally distributed data. Here the assumption is that there is an immediate and constant difference between the use and non-use of the system, i.e. there is no learning function, no drifting process and no erosion of the effect.

However, the assumption of a constant effect is often inappropriate. To get a complete view of the sources of variability and to handle the problem of serially correlated data, multi-level models are recommended (Goldstein, 2003).

With such models, drivers or situations with missing data have generally to be included. Elimination of drivers or situations because of missing data in order to keep complete data set may cause bias in the estimation of the impact.

It is assumed that data will have been cleaned up in the data quality control phase. Nevertheless, to be sure that the estimation will be influenced minimally by outliers, one can use either robust estimates such as trimmed mean and variance or non-parametric tests such as a Wilcoxon rank test or a robust MM (minimum mean) regression (Gibbons, 2003; Wasserman, 2007; Lecoutre and Tassi, 1987). Such tests provide protection against violation of the assumption of a normal distribution of the performance indicator.

Additional Step 4: data mining

Data mining techniques allow the uncovering of patterns in the data that may not be revealed with the more traditional hypothesis testing approach. Such techniques can therefore be extremely useful as a means of exploratory data analysis and for revealing relationships that have not been anticipated. The data collected in a FOT is a huge resource for subsequent analysis, which may well continue long after the formal conclusion of the

FOT. One relatively simple technique for pattern recognition is to categorise a dataset into groups. Cluster analysis tries to identify homogeneous groups of observations in a set of data according to a set of variables (e.g. demographic variables or performance indicators), where homogeneity refers to the minimisation of within-group variance but the maximisation of between-group variance. The most commonly used methods for cluster analysis are k-means, two-step, and hierarchical clusters (Lebart et al., 1997; Everitt, 2000).

Step 5: global assessment

This section deals with the issue of identification of models and methodologies to generalise results from a certain FOT to a global level in terms of traffic safety, environmental effects and traffic flow. One problem when generalizing results from a FOT is to known how close the participants in the FOT represent the target population. It is often necessary to control for: usage, market penetration and compliance (the system might be switched off by the driver) and reliability of the system. The process of how to go from the FOT data to safety effects, traffic flow and environmental effects is illustrated in Figure 9.5. In this process two steps need to be taken. One is scaling up the FOT results, for example to higher penetration levels or larger regions. The other is to translate the results from the level of performance indicators (for example, time headway distribution) to the level of effects (for example, effect on the number of fatalities). For each type of effect there are (at least) two different ways to generalize the results: through microsimulation or directly.

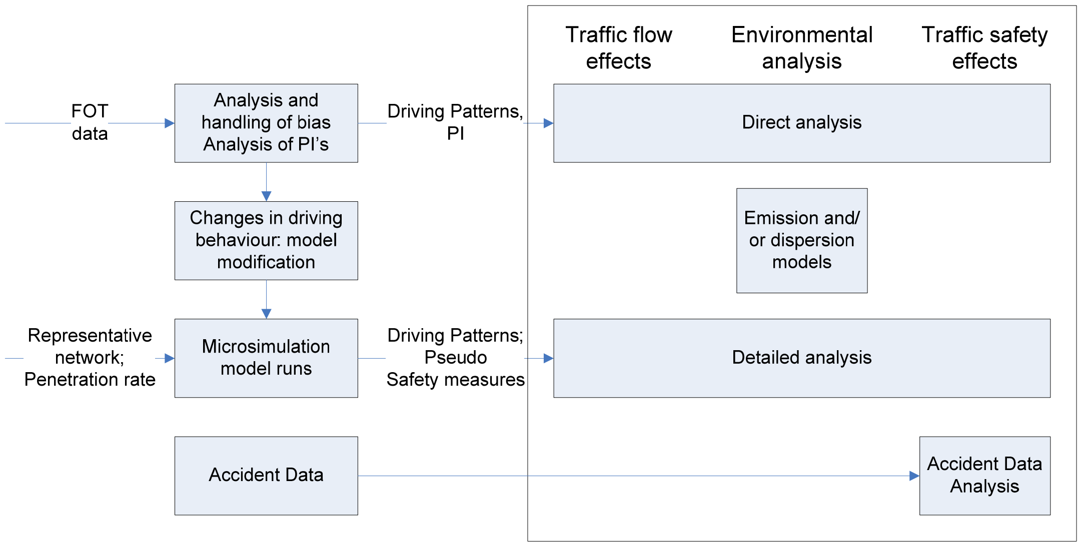


Figure 9.5: Block diagram of scaling-up process

The direct route includes both estimation directly from the sample itself and estimation through individual or aggregated models. Some advantages of the direct route are that it is rather cheap and quick. The alternative is to use a traffic microsimulation model which represents the behaviour of individual driver/vehicle units. The advantages of microsimulation are that they can be more reliable and precise and can incorporate indirect effects (such as congestion in the network at peak times).

Since traffic microsimulation models consider individual vehicles in the traffic stream, there is consequently the potential to incorporate FOT results in the driver/vehicle models of the simulation. Impacts on the traffic system level can then be estimated through traffic simulations including varying levels of system penetration into the vehicle population.

Microsimulation does not necessarily yield the impact variable that is of interest. Various aggregated and individual models are necessary to convert for instance speed to safety effects (e.g. via the Power Model which considers the relationship between driving speed and the risk of an accident at different levels of severity). In addition, the modelling detail of traffic microsimulation places restrictions on the practical size of the simulated road network. Macroscopic or mesoscopic traffic models combine the possibility to study larger networks with reasonable calibration efforts. These models are commonly based on speed-flow or speed-density relationships. Large area impacts of FOT results can therefore be estimated by applying speed-flow relationships obtained from microsimulation for macro- or mesoscopic traffic modelling.

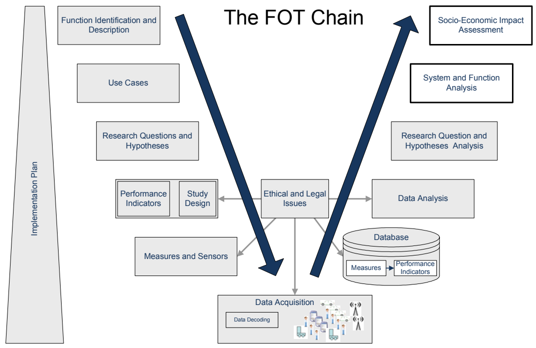
Exhaust emission from road traffic is a complex process to describe. Models for exhaust emissions in general include three parts: Cold start emissions, hot engine emissions and evaporative emissions. An exhaust emission model can roughly be described as:

$$\Sigma(\text{Traffic activity}) \times (\text{Emission factor}) = \text{Total emissions}$$

Of course traffic activity data then has a high correlation to total emissions. Traffic activity data includes: mileage, engine starts and parking. In addition to traffic activity data one needs data for: the vehicle fleet; road network; meteorological conditions; fuel quality etc. If the driving pattern is influenced by the traffic situation, such data for the FOT vehicles are directly available. In order to estimate driving pattern changes for all vehicles by traffic situation, microsimulation models could be used. In order to estimate emission factors for these alternative driving patterns there is need for exhaust emission measurements or exhaust emission models on an individual level. The recorded speed traces from the FOT vehicles can also be post-processed through a fuel consumption and emissions model to produce data on environmental effects.

Speed has a close relation to safety. The speed of a vehicle will influence not only the likelihood of a crash occurring, but will also be a critical factor in determining the severity of a crash outcome. This double risk factor is unique for speed. The relationship between speed and safety can be estimated by various models such as the Power Model (Nilsson, 2004; Elvik et al, 2004), that estimates the effects of changes in mean speed on traffic crashes and the severity of those crashes. The Power Model suggests that a 5 % increase in mean speed leads to approximately a 10 % increase in crashes involving injury and a 20 % increase in those involving fatalities. More examples of models for speed-safety relationships are reviewed in Aarts and van Schagen (2006). In general it is important to consider under which assumptions the models are valid. The Power Model, for example, is valid under the assumption that mean speed is the only factor that has changed in the system. Therefore these models are more suitable for FOTs with systems mainly dealing with speed, and even then they fail to consider changes in the distribution of speed (shape of the speed distribution and changes in speed variance).

10 Socio-Economic Impact



10.1 Introduction

FOTs supported by funding from the European Commission require a socio-economic evaluation at a European level. As a result this chapter concentrates on the methods for providing such an analysis. Many parts of the chapter will also be relevant to FOTs conducted at a national or regional level within Europe or to FOTs conducted outside Europe.

A consistent methodology for carrying out socio-economic evaluation for EC-funded FOTs will maximise the comparability of the results across regions, ICT systems and FOTs. These FOTs will test ICT-based systems for mobility (from now on referred to as “ICT systems”).

In the past, the impact assessment of FOTs focussed 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 goal of this chapter is to provide concise advice on how to carry out a socio-economic impact assessment. It will address the possible breadth of impacts that can be considered and the available resources for carrying out the assessment. This advice contains references to examples of good practice in existing (web) documents.

Our advice will be useful for a variety of parties: the organisations conducting the FOTs, including the socio-economic impact assessment specialist; the client commissioning the FOTs; and the consortia drawing up proposals for the FOTs. This chapter assumes that a “professional” in the area of socio-economic impact assessment will carry out the analysis. This individual we will refer to as the “analyst”. This information on socio-economic assessment is not meant as a “tutorial”.

Furthermore, the choices made in other parts in setting up the FOT are linked to the choices made for (parts of) the Cost Benefit Analysis (CBA). For example, choices about performance measures and scaling up are directly linked to what can be analysed in terms of scope and impacts in the CBA.

Our advice is based on the review of about 20 studies (see Table 10.1 below) reporting the socio-economic impacts of ICT systems. Broadly, there are three research angles where the evidence comes from:

- **Socio-economic impact assessment studies** typically investigate the impacts of a system for a future time horizon. These prospective studies make use of an ex-ante impact assessment, often based on literature review, simulation work and expert estimation. They are often comprehensive in scope but they do not involve, or only to a limited extent, data from real-life conditions.

- **Transport appraisal guidelines** or **scoping studies** in this area are very much focused on the appraisal part of the impacts. They dig deep into the methodology and practise of appraisal. They also involve proposals for standardisation of appraisal. Their detriment is that they are not developed for specific use in the field of safety evaluation.

- **Field Operational Test assessment studies** typically assess the impacts of one or more system functions. FOT evidence can lead to a quantum leap in the impact assessment because FOT produce measured data about effects. Therefore, the assessment can rely on ex-post measurement data.

All different research angles have their specific strengths and weaknesses. In preparing this guidance document, we found it useful to combine the strengths of the different perspectives.

Table 10.1: Summary of reviewed socio-economic impact assessment studies

Study			Main Focus				Impacts				
Name	Year	Geograph. Scope	Methodology	Case Studies	Application	FOT Assessment	Safety	Mobility – Direct	Mobility – Indirect	Environmental	System Costs
eIMPACT	2006-08	EU	X		X		X	X	X	X	X
SEiSS	2005	EU	X	X			X	(X)	X	X	X
ECORYS / COWI	2006	EU			X		X		X	X	X
ROSEBUD	2005	EU	X	X			X	(X)	(X)	(X)	X
ADVISORS	2000	EU			X		X	X	X	X	
HEATCO	2006	EU	X				X		X	X	
RAILPAG	2005	EU	X	X							
FUNDING	2007	EU			X						
NATA	2003-08	UK	X								
Full Traffic	2008	NL			X	X	X	X	X	X	
TAC Safe Car	2006	AUS			X	X	X	X		X	

Study			Main Focus				Impacts				
Name	Year	Geograph. Scope	Methodology	Case Studies	Application	FOT Assessment	Safety	Mobility – Direct	Mobility – Indirect	Environmental	System Costs
Freightliner FOT	2003	USA			X	X	X		X		X
Mack FOT	2006	USA			X	X	X		X		X
Volvo FOT	2007	USA			X	X	X		X		X
IVBSS	2007	USA			X	[X]	[X]				
RDCW FOT	2006	USA			X	X	X				
ACAS FOT	2006	USA			X	X	X				
ICCS FOT	1999	USA			X	X	X				X
CAS Benefits	1996	USA			X		X				

Annotation: (…) … addressed as an option

[…] … subject to future reports

10.2 Issues

The socio-economic impact assessment investigates the impacts of a technology on society. Ideally a socio-economic impact assessment provides the decision maker with relevant information in a concise format. The relevant comparison is between the benefits and costs between a base case, e. g, a scenario without the ICT system ("without case") compared to those of the scenario with the ICT system ("with case"). In preparing to carry out a socio-economic impact assessment, the analyst is faced with making choices about the impacts to be investigated in the analyses, the geographical scope of the assessment and the analyses to be carried out. This section will go deeper into the issues surrounding these choices. The chapter will conclude with guidance on how to make the choices and carry out the cost-benefit analysis.

10.2.1 Assessment scope and process implication

At the start of the socio-economic assessment, a view will need to be taken on the scope of the analysis. Ideally the assessment would include all impacts of the system no matter how small that impact is: safety, mobility, efficiency and productivity, environmental, user acceptance and human factors, performance and capability, legal and implementation issues, and costs. However setting an unlimited such a broad scope for a socio-economic assessment will result in excessive data collection and analysis in terms of expense and time. Given that the purpose of the assessment is to firstly ensure that the implementation of the system is economically beneficial and secondly to aid the choice between alternatives, the scope of the assessment often can be narrowed by excluding minor or insignificant impacts as long as the exclusion of these impacts will not bias the appraisal. An *impact table* such as in Batelle Memorial Institute (2003, p. 45) is extremely useful at the start to clarify which impacts have been considered and which — if any — have been ruled out as negligible or impossible to assess.

10.2.2 Geographical scope of assessment

The issue related to geographical scope is the ability to translate the findings of the FOT to a "higher" geographical level. The FOT is usually carried out at one or more locations, on a regional or national scale. However, the number of equipped vehicles and, if relevant, equipped roads, as well as the number of "equipped" kilometres driven, is usually a small percentage of the total vehicle fleet and the kilometres of roads. Therefore, in order to draw conclusions about the impacts and effectiveness of the system tested, a "scaling up" of the results is needed in order to draw conclusions and in order to ensure transferability of the results. Chapter 9.5 addresses the scaling up issues, which is to the national or European level. The availability of data plays a role in the decision to what level to scale up the results. The Guidance section (10.3) goes into more detail to explain how to deal with this issue.

10.2.3 Analysis of impacts

The analysis of impacts represents the most sophisticated part of the assessment. Figure 10.1 provides an overview over the most common effects (safety, mobility, environment, costs) which are considered in a FOT assessment. This assessment framework involves the distinction between direct and indirect effects (in safety mechanisms but also with respect to mobility effects, see below). It also implies the distinction between effects on internal and external costs. Mobility effects typically lead to lower internal costs of transport (i.e. time, fuel consumption) and also external costs (e.g. pollution, CO_2). The reduction of external costs is flagged out separately under environmental benefits because of its importance on the political agenda. The assessment can of course also consider wider economic effects (e.g. growth and employment effects of new technologies). However, given limited time and budget, it is useful to concentrate on the main impacts.

Safety benefits

The assessment of safety impacts has to consider several effects which can be combined to predict the overall safety benefit. Ideally, accidents and their consequences) have to be used in order to estimate accident risks. More commonly, information collected on incidents (conflicts) and on driver behaviour more generally has to be used to estimate changes in risk and therefore should also be integrated into the assessment plan. As an example and representing best practice, the Mack FOT puts the goals of the safety analysis as follows:

1. Determine if driving conflict and crash probabilities will be reduced for drivers using the system,

2. Determine if drivers drive more safely using the system,

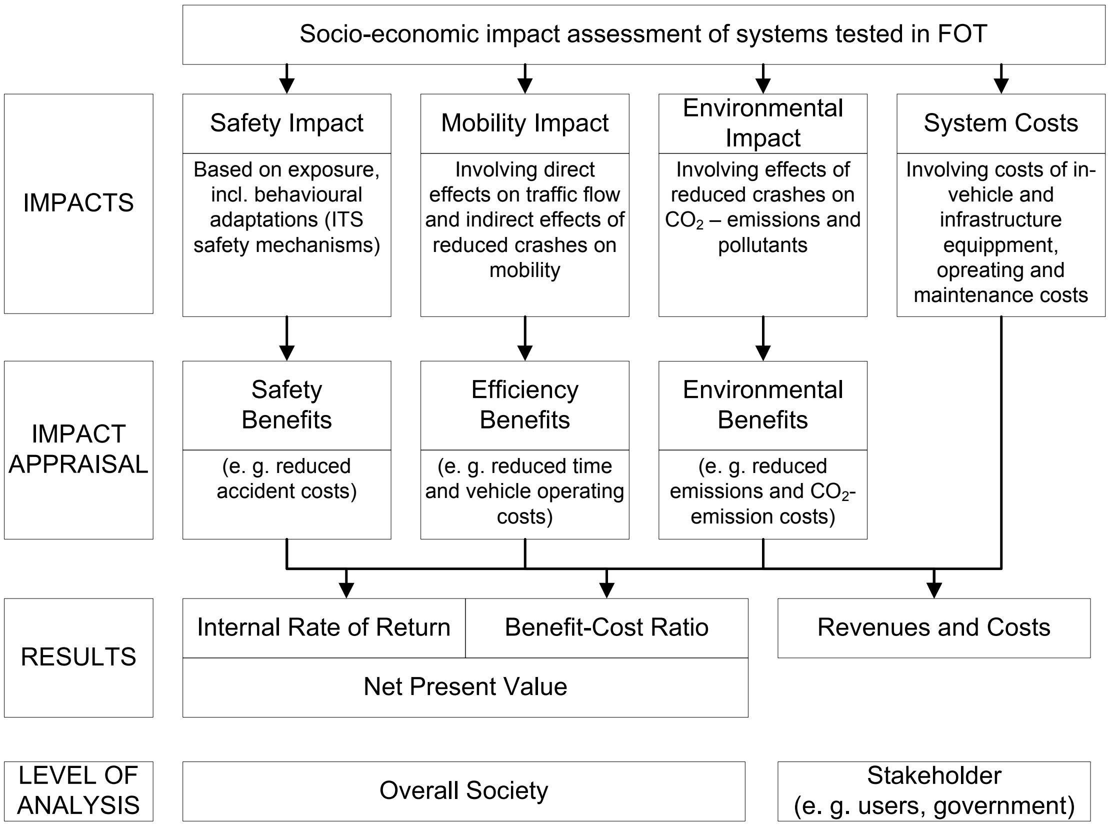


Figure 10.1: Scope of the impacts within socio-economic impact assessment

3. Determine reduction in crashes, injuries, fatalities if all fleets operating in the observed area were equipped with the system,
4. Determine if drivers using the system have less severe crashes than drivers without the system.

The first step collects sensor data from each vehicle within the FOT (e.g. brake force, steering angle). Based on earlier definitions the number of driving conflicts can be determined. Thus, two numbers for the driving conflicts – reflecting the with and the without case – are available to calculate the exposure ratio. This ratio reflects the number of driving conflicts in the with case compared to the without case. To provide an example: given a system which maintains the safe distance to a predecessor vehicle, the number of driving conflicts due to close following will be reduced from 10 conflicts per 1000 km to 5 conflicts per 1000 km. Thus, the exposure ratio equals 0.5 which indicates that driving with the system is safer than without the system. In general, an exposure ratio below 1 indicates a safety benefit.

The benefit of lower exposure to accident risk will likely be modified based on adaptations of individual behaviour due to psychological reasons (second step). Behavioural adaptations can comprise e.g. adapting the following distance, adapting the speed variance, adapting the lane change behaviour (risky cut-ins or changing the lane without signalling it in advance). Examples for such behavioural changes can be found in the ITS safety mechanisms (eIMPACT). In this project, nine mechanisms have been introduced which lead to positive or negative safety effects. In most cases, the motivation for behavioural adaptation is that the driver wants to avoid “public” warnings (noticeable to all passengers) and “education” by the system.

The third step deals with scaling-up from the FOT to a wider area (EU, country, region). This process is subject to the procedure proposed in scaling up.

The last step leads to the prevention ratio. In-depth information on accidents is used to calculate the mitigation effects of using the system. Maybe the system cannot avoid the accident but it can mitigate the accident consequences. This issue has to be considered in determining the effects for casualties. For systems affecting speed, the Power Model can be applied to calculate changes in severity.

Combining steps 2 to 4, it is possible to calculate the prevention ratio. For this ratio the probability of having a crash (casualty) when having a driving conflict in the with case is compared to the same probability in the without case. In the above example the number of driving conflicts in the with case was 5 and 10 in the without case. Let us assume that out of the 5 driving conflicts 1 accident occurs and out of the 10 driving conflicts 3 accidents occur. Thus, the probability of having an accident due to a driving conflict is 0.2 in the with case and 0.3 in the without case. These values reflect the prevention ratios.

Efficiency benefits

Efficiency benefits are typically composed of two effects. They involve:

- Direct mobility effects resulting from a smoother traffic flow, e.g. where the system allows traffic to re-route to avoid current congestion, or improves mean speeds by encouraging safe following behaviour,
- Indirect mobility effects resulting from reduced crashes e.g. reduced delays at incidents and accidents.

Direct mobility effects can play an important role in the socio-economic impact assessment. On the appraisal level, direct mobility effects are reflected in changes of time costs, fuel consumption costs and reliability changes. Because socio-economic impact assessment identifies quite commonly reductions of time costs as a major driver of the results, direct mobility effects are generally worthwhile to explore.

The investigation of direct mobility effects typically involves microscopic traffic flow simulation. A number of models (e.g. ITS Modeller, VISSIM, Paramics, DRACULA) have been applied to assess these impacts. Best practise, including on cross-validation of models, can be found in eIMPACT D4 (Wilmink et al., 2008) and Full Traffic (Technische Universiteit Delft, 2008). Typically, when traffic flow becomes more homogeneous, the standard deviation of the vehicle speed becomes lower. As a result, the average vehicle speed may increase or the infrastructure capacity improves. As a consequence, time costs and vehicle operating costs will decrease.

However, the realisation of those benefits is closely related to the likely market penetration. Mature ICT systems typically can produce such effects, ICT systems in the phase of market introduction typically can not. For internal efficiency it is therefore important to figure out at the beginning of the FOT assessment (when the scope is defined) whether direct mobility effects will be likely to appear or not.

Compared to the direct mobility effects, experience suggests indirect mobility effects are not restricted by conditions of market penetration. They can be realised in any case, as an add-on to the safety benefits. Indirect effects occur when the number — as well as the severity — of crashes is reduced. The benefits result from less congestion, therefore reducing journey times and fuel consumption. Typically, indirect traffic effects add up to about 10 % of the safety benefits.

Given the state of the art in traffic modelling, indirect mobility effects are assessed more frequently than direct mobility effects. Good practise on the appraisal of indirect mobility effects can be found, however, in recent European scale assessment studies (eIMPACT; COWI, 2006) and US American FOT assessments (Batelle Memorial Institute, 2003; Volvo Trucks North America Inc, 2007). Some countries have methods specifically to address these effects (e.g. INCA in the UK) (see http://www.dft.gov.uk/pgr/economics/rdg/jtv/inca/).

Environmental benefits

Environmental benefits comprise lower CO_2 and air pollutants emissions. Noise also fits into this category but we would caution that noise should only be analysed where ICT systems are expected to make a significant difference between the two scenarios (with/without the system). CO_2 and pollutants emissions are both speed dependent, with CO_2 emissions directly linked to fuel consumption. Hence, there is a close relation to the mobility effects discussed above. The impact of CO_2 emissions is on a global scale, and is not linked to the particular country or area type where the CO_2 is emitted. The impact does, however, vary according the year in which the reduction (or increase) in emissions takes place — the impact becoming greater further into the future. Actually, mobility effects have impacts on both efficiency and environmental benefits. However, because they are transmitted through the environment, and because they are largely externalities (i.e. their incidence is mostly on individuals other than the emitter) environmental benefits fall into a special category.

System costs

System cost estimation is an element within FOTs which is quite often neglected. System promoters may not see costs as an impact. However from a socio-economic point of view, they are a (negative) part of the impact of systems. Cost estimation should take care of the following aspects:

- **Cost elements to include:** The system costs comprise the costs of in-vehicle, roadside infrastructure equipment and nomadic devices. Besides that, operating and maintenance costs have also to be considered. Examples of good practise for system costs can be found in US American FOT assessments (Freightliner FOT, Mack FOT, Volvo FOT).

- **Relevant size of costs:** CBA applies a resource based view. This means looking at potential savings of productive resources and on the other hand at the resources necessary to achieve this effect. The implication for cost estimation is that only the input of productive resources is relevant and not potential market prices. The convention proposed e.g. by eIMPACT is to use the cost price (the price of the ICT system paid by the manufacturer to its supplier) plus a mark-up which is allowed for in-vehicle implementation. However, the contrary, market prices are relevant for user-centred analyses. Generally, in the face of limited evidence it is useful to apply the “Factor 3” rule of thumb, which means that in the automotive industry market prices for ICT systems differ from the cost prices by a factor of 3.

- **Process of cost estimation:** Typically, cost estimation will be carried out by an expert group comprising of FOT internal staff and external industry experts. To avoid conflicts with confidentiality and the like, it appears sometimes helpful to introduce rough estimations to the group instead of working from blank sheets. Guidance to rough estimations for investment and OEM costs can be applied from a US-American database on ITS costs and benefits (www.itscosts.its.dot.gov).

10.2.4 Classification of assessment methods

Figure 10.2 gives a classification of socio-economic assessment methods, based on which of the elements are included, and in particular:

- Whether a full set of impacts is addressed — for example, if a significant CO_2 reduction can be anticipated, has it been included;
- Whether the assessment is from the social perspective only, or whether financial and stakeholder analyses are also provided.

The recommendation is that the FOTs should be designed to be as complete as possible, both in terms of impacts and stakeholder views. The assessments in the FOTs reviewed are examples of good practice. However, they differ in the types of analyses carried out, as well as in the scope of the effects examined, with the exception of safety impacts.

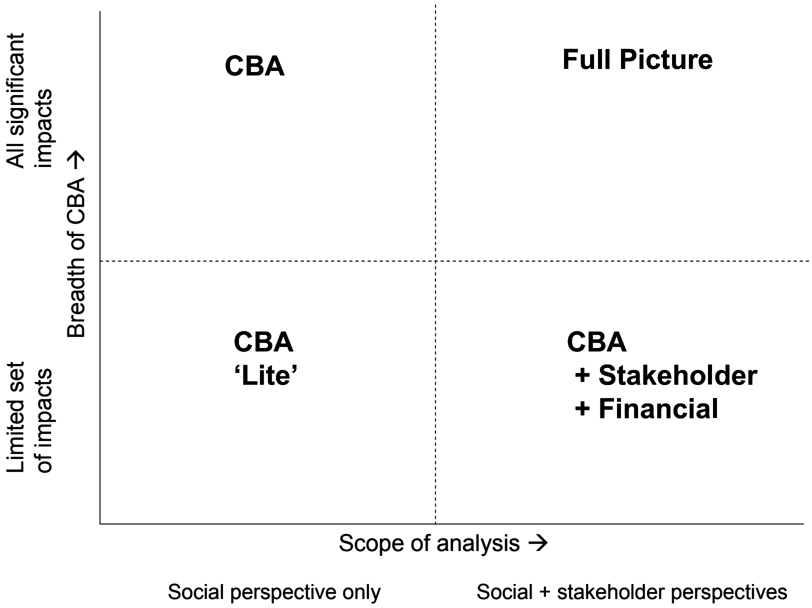


Figure 10.2: Classification of Socio-Economic Assessments

Figure 10.2 highlights another dimension in which assessment methods can be classified, namely whether or not they make use of case-specific Willingness-to-Pay (WTP) evidence. In the design of future FOTs, we recommend that clients and analysts consider WTP studies as a way of getting better evidence on the users' likely demand for the products. WTP can provide uniquely useful evidence on the value of the ICT system to consumers and producers. In absence of this, FOTs can refer to evidence in the literature (market-based). WTP studies will, however, add to the cost and skill set required for FOTs, so the advantages and disadvantages will need to be weighed in each case.

We note that past FOTs have generally relied on market-based values (e.g. the U. S. CAS and Mack FOTs), although the U. S. ICC FOT did make use of specific WTP evidence, and as such is a useful reference. Also, we note that most previous assessment guidelines, including eIMPACT, assume that literature-based values will be used. Here, we leave the option open and recommend that clients and analysts decide at the inception phase of the FOT whether or not to go down the WTP route.

10.3 Guidance

The analyst faces choices in setting up and carrying out the analyses. The choices will be influenced by the priorities identified by those setting up the FOT, as well as budget and time constraints. The list below summarises these choices.

Methods:

- The basic choice is CBA, which summarises benefits and costs at a societal level
- Stakeholder perspectives: Makes use of the same input data as the CBA, but considers stakeholder-specific benefits, costs and financial analyses.

Identification of impacts:

- The basic choice is to use the costs incurred and the main expected benefit(s), as identified by use of the impact table.
- Other impacts — both direct and indirect — can also be included, depending on the stakeholder perspective as well as the choices made elsewhere in the project, for example in hypothesis formulation, measurement methods and equipment and modelling capability.
- Willingness-to-Pay evidence, if also collected during the FOT, can be used to supplement the analysis methods above.

Scope of geographical assessment:

- The basic choice is the country level. In this case, the generic data needs (see section 10.3.2) are limited to the country in question.
- EU-level analyses are preferred. These require substantially more general data from individual countries. Extra challenges in execution can be encountered due to differences in definitions or classifications.

10.3.1 How to carry out the assessment

Carrying out a CBA

The socio-economic impact assessment of a system within an FOT should be based on a CBA, since it is the most widespread, commonly accepted and practised method for analysing socio-economic impacts. It is clear that CBA accounts for all benefits and all costs on a society level, including benefits and costs to all groups. CBA follow a four-step-process involving framework and preparatory work, measuring impacts, appraising impacts in a common monetary value and confronting the discounted society benefits with the costs of the policy measure. However, this process leaves also some room for shaping the individual steps of the process. We recommend considering the following issues.

1. CBA framework:

Definition of the cases to be compared: Looked at is the with-case (ICT-system equipped) against the without-case (without the system).

Base year and time horizon of the assessment: CBA can be performed for the whole life cycle of the considered system or only for selected target years. This decision depends on information needs.

Geographical scope: Because of data availability the geographical scope should be congruent to existing statistical reporting ICT systems. Reference only to the local area where the FOT takes place is insufficient for this reason and because the results of different FOTs need to be compared. This implies however that the socio-economic impact assessment has to undergo a scaling up procedure before the CBA in order to project the impacts from the FOT itself on to a larger area. The most practical appears to be assessment at the national level (assuming “nationwide deployment”). However, it is even more useful to provide results on a European level. The European perspective is important when the effects of FOTs in different member states should be compared or when policy measures are planned or considered to ensure a European scale deployment (e.g. eCall).

Discount rate: The discount rate ensures that benefits and costs are expressed for a common base year. A discount rate of 3 % (real) is recommended as a default (see 'Other economic parameters').

Deployment scenario: It has to be estimated which share of new vehicles or which share of the total vehicle fleet will be equipped with the system in the target years and over the assessment period as a whole (depends on answer to 'Base year and time horizon' issue above). For life cycle assessment it is also necessary to estimate the development of the equipment (technical capabilities, costs).

Impact table: The impact table serves as an instrument to expedite identification of impacts. It is aimed to ensure that the FOT team and the group responsible for the socio-economic impact assessment are fully aware of the complete impacts of the system. For efficiency reasons and likely budget constraints (competing FOTs and competing assessment issues within a FOT) it is necessary to concentrate the analysis on the significant impacts — impacts expected to be negligible, or impossible to analyse within the resources available, should be flagged as such in the impact table. Concerning the system, safety is the relevant impact by definition. Direct and indirect mobility impacts and environmental impacts are typically also addressed. System costs will always be relevant.

2. Inputs for impact assessment (including cost estimation)

Impact measurements represent an essential input to the cost-benefit assessment. We would normally expect most of these to feed through from the FOT experiment to the scaling-up procedure (Chapter 9) to the CBA inputs. In particular, accident prevention and system costs at the national / EU levels should be delivered this way. Impacts on mobility and environment will typically require additional analysis at the CBA stage (although in a well designed FOT experiment, it may be possible to gather data specifically on any expected sources of benefit, e.g. reduced variability of traffic speeds or reduced fuel consumption; see the TAC Safe Car FOT). The analysis of different FOT assessment has revealed some evidence on best practise for impact measurement. The requirements for CBA can be provided as a sort of output specification. This makes sure that the socio-economic impact assessment will be provided with the appropriate input data for carrying out the assessment. In terms of an output specification the following elements have to be put in place:

Accident and traffic performance database: See section 10.3.2.

Effectiveness of the system: These values represent key output of the FOT which have to be provided to the socio-economic impact assessment.

Procedure for *scaling up the effects* to nationwide/European level

Cost estimations: See section 10.2.3 on system costs.

3. Impact valuation

Methodological base for impact valuation: The general objective of this step is to provide unit values for the physical impacts. Several methods compete in the field of impact appraisal. They can be subdivided in objective approaches (e.g. damage costs, avoidance costs) and subjective approaches (e.g. willingness-to-pay). In European member states, different practises and preferences exist for impact appraisal. A lot of surveying and standardisation efforts have been made by projects like HEATCO (Bickel et al., 2006) to come to common European base. As a general recommendation, it can be stated that unit values for CBA should be based on objective approaches. However, willingness-to-pay information can largely contribute to a higher quality of the assessment when analyses for the users are carried out.

Good practise on unit values: See eIMPACT (Assing et al., 2006), HEATCO and the handbook on external costs of transport.

National or European unit values: This decision corresponds with the geographical scope. Assessment on national level will typically make use of national cost unit rates. For European scale assessment we recommend using the harmonized values contained in HEATCO — note that these are still differentiated by country, but are on a harmonised theoretical basis.

4. Results

Cost-benefit analyses can produce different summary measures of performance. It represents good practise to calculate the Net Present Value (NPV) by summing up all discounted values of benefits (plus sign) and costs (minus sign). Moreover, Benefit-Cost Ratios (BCR) are a very common expression of system profitability which can be calculated by dividing the total benefits by the total costs. It is also practical (see "Base year and time horizon") to calculate "snapshot" BCR for target years. In this case, the costs will be transformed to annual values (using the discount rate) and will be compared to the target year benefits. For FOTs, we recommend the calculation of calculate both figures, NPV and BCR.

For the **social CBA**, we recommend reporting:

- safety benefit (€M);
- other benefits to road users (€M) — mainly time savings, operating cost savings and reliability gains;
- environmental benefits (€M) — including climate change, regional and local air quality effects; noise; and other impacts;
- revenue to operators (€M) — there may be multiple operators, including infrastructure and service operators – each will want to know the impact on themselves (financial), although for the social CBA these revenues may be aggregated;

- costs to operators (€M) — including capital, maintenance and operating costs;
- revenue to automotive OEMs (€M);
- costs to automotive OEMs (€M);
- revenue to government (€M) — including tax revenue changes;
- costs to government (€M) — including investments in R&D and in implementation of ICT systems.

Tabulation of the social CBA is shown in Table 10.2. All entries are at Present Values. [6] A common base year (for prices and discounting) aids comparison across different technology options. RAILPAG (EIB, 2005) has a more detailed breakdown by stakeholders (an 'SE Matrix'), which some analysts may find helpful in presenting the social CBA.

In cases where the public sector expects to contribute to the development or implementation of the system, we recommend also presenting a Benefit:Cost Ratio with respect to public sector support, which HEATCO (Bickel et al., 2006: 41-2) identifies in use by the EC, UK and Switzerland:

$$BCR = \frac{NPV}{PV(PublicSectorSupport)}$$

The calculation of the Benefit-Cost ratio (BCR) is delicate issue in CBA. On one hand the BCR is a very powerful measure, because it applies to the common situation where investment budgets are limited and maximum value for money is required (making best use of a scarce resource). On the other hand the definition of 'costs' (the denominator) can be problematic. As a general rule, the BCR is useful when the denominator is defined in the same way for all options being compared — for example, NPV per unit of central government budget (which would be a BCR of interest to central government). Our recommendation of a BCR with respect to Public Sector Support broadens this to the budget for public expenditure as a whole. This avoids creating an incentive to manipulate the BCR by shifting costs to local and regional government.

In the example shown in Table 10.2, the BCR with respect to public sector support will be 4240/(379-34) = 10.3, which indicates a high social return from each € of public funds contributed.

[6] For more detailed information please refer to FESTA deliverable D2.6.

Table 10.2: Social CBA tabulation

Group	Impact	€M (2008 base)			
		2015	2025	Present Value (Total)	
Consumers	Safety benefits	289	299	3715	a
	Other road user benefits	574	606	603	b
	Environmental benefits	63	66	58	c
	...				d
Producers	Revenue	723	780	8520	e
	Costs	248	233	−8311	f
Government	Revenue	3	4	34	g
	Costs	12	14	−379	h

Net Present Value (NPV) = Σa..h	4240

Notes: sign — all negative impacts on the Group affected are shown with a negative sign, thus Costs appear with a negative sign; 2008 base — indicates appraisal at constant general prices using 2008 CPI, and with 2008 as the base year for discounting in the Present Value column.

Carrying out a stakeholder analysis

In contrast to CBA, only particular benefits and costs are relevant for particular stakeholders. The reduction of exhaust and CO_2 emissions are not benefits to users, unless they are charged for it (through vehicle-taxes or tolls). The costs of in-vehicle equipment do not represent costs to the government, unless the government agrees to pay for a share of this. The consequence is that ICT systems which are profitable on society level (NPV, BCR) will not be deployed when a relevant stakeholder group is economically impaired. Hence, it is necessary to include stakeholder perspectives in the FOT socio-economic assessment.

Practically, stakeholder analyses make also use of accounting costs and benefits, but on the level of the individual stakeholder group. This implies the following for *users*, but also in general:

- Cost and benefits must be investigated according to their stakeholder relevance. Safety benefits (reduced accident and casualty risks) for instance are relevant to *users* (and to *insurance companies* as well).
- The appraisal of the impacts can be different. *Users* face market prices when considering the investment in a system (see factor 3 rule of thumb). For benefit evaluation the implication is to use market values if available (e.g. fuel consumption: station prices (incl. taxes) instead of net prices). Otherwise, willingness-to-pay approaches have their justification here because they are better suited to reflect individual preferences.

Further adaptations to the CBA approach involve the use of a different discount rate (reflecting private sector interest rates) and the use of a different result measure (fair market price for a pre-defined annual vehicle mileage or the critical (break-even) mileage for a given market price).

The stakeholder analysis reporting will vary with the analytical methods used. For example, in the TAC Safe Car Project, Monash University used subjective questionnaire methods to investigate users' acceptance of several ICT systems including ISA (see Regan et al., 2006).

Another useful form of stakeholder analysis from the User perspective is Willingness-to-Pay evidence, as shown in eIMPACT D3 (Assing et al*.,* 2006 p. 119).

For the *vehicle OEMs* and both *infrastructure and service operators*:

- Where they are commercial bodies, a *financial analysis* will provide the most important stakeholder information;
- Where they are *public sector* agencies, a *financial analysis* may need to be combined with an assessment against their public service objectives – however, in some cases the overall social CBA will serve this purpose, depending on the approach taken by the agencies involved.

Carrying out a financial analysis

The internal rate of return (IRR) of a project is the interest rate that will generate an NPV of zero. In an equation, this is:

$$\sum_{t=0}^{T} \frac{B_t - C_t}{(1 + i_{IRR})^t} = 0$$

where IRR is internal rate of return.

The stakeholder for whom the IRR is calculated compares the IRR with a target rate. This target rate depends for each stakeholder. For public authorities as a stakeholder the target rate will be less than for private investors as stakeholders.

In any case, a calculated BCR or IRR should be accompanied by an NPV. We recommend that financial IRRs are reported for all FOTs.

The IRR concept can be modified for comparison reasons. For his approach, the cash flow streams are subtracted. With the new cash flows the modified IRR is calculated. If the IRR is above the trigger rate, the project with the larger cash flow is the better project.

Of key interest will be the IRR from the point of view of *specific stakeholders* (or stakeholder types)*.* The IRR for vehicle OEMs will influence their decision about investing in the technology. Similarly, the IRR for infrastructure operators and service operators will influence their decisions – particularly where these are commercial operations.

Hence the key information will be in the form:

IRR_{OEMs} *= … %*

$IRR_{RoadAuthorities} = ... \%$

Further IRRs should be reported where there are other stakeholders with a commercial interest, for whom significant impacts are expected. Tables such as those used by WebTAG (DfT, 2005) also provide a useful series of snapshots of the financial impact. In this case, in order to be meaningful the tables should relate to specific stakeholders or stakeholder types, e.g. vehicle OEMs or road authorities.

The financial results can be taken a stage further by reporting the breakeven point in terms of sales or market penetration, or the target price, down to which the system must be engineered in order to achieve financial viability. Graphical presentations may be useful in these cases.

10.3.2 Data needs

The data needed to carry out a socio-economic assessment for an FOT are extensive, and fall into two broad categories:

- FOT-specific data which will be gathered during the FOT itself
- Generic data, which play a role in:
  - Scaling up the results from the experimental situation of the FOT to the national or EU level
  - Reaching a socio-economic assessment, based on the FOT data scaled-up to National or EU level.

The following section outlines the FOT-specific and generic data likely to be needed. Thereafter recommendations on ensuring data quality and validity are given. Management of the data for socio-economic assessment is after this.

FOT-Specific Data

The key items of FOT-specific data likely to be needed are:

Accident rates (or risks) with and without the ICT system in place for the FOT sample. These will need to be differentiated by all the key drivers of accident rates (risks) in the FOT sample (e.g. road type; driver type; traffic conditions) so that accurate extrapolations can be made to the whole network. Accident rates (or risks) will be needed with and without the ICT system in place for the FOT sample. These may need to be derived from data on unsafe behaviours if the sample is too small to contain a significant number of actual accidents, although this is likely to be done as part of the Performance Indicators in any case. See section 10.2.3 on Safety Benefits.

One approach to estimating the impact on accident rates uses the effectiveness rate (% of relevant crash type avoided) as in the Collision Avoidance Systems (CAS) Benefits Study (NHTSA Benefits Working Group, 1996).

A more sophisticated approach can produce data on accident severity as well as accident rates. Since accident severity is determined by the severity of the most serious casualty only, a complementary item of data would be any expected change in the number of casualties per accident. Regan et al. (2006) measured time spent buckled-up and time before buckling-up to produce injury severity estimates.

Examples of how data is produced for accident severity and accident rates can be found in Regan et al. (2006), UMTRI et al. (2006), Volvo Trucks North America et al. (2007) and USDoT (1999).

Whichever approach is used to estimate accident rates and accident severity, the analysis will need to take account of any options in the implementation path. For example, in the Freightliner FOT study (Batelle Memorial Institute, 2003) there were four possible deployment groups (HazMat tankers; all tankers; tractor trailers; all large trucks) — input data will be required for each of these options.

Multiple scenarios may also be needed to enable sensitivity testing. That is, where there is uncertainty over accident rates/severity or other key variables, this can be handled through 'what if' scenarios based on combinations of the possible outcomes (Batelle Memorial Institute, 2003).

There may also be some value in having spatially differentiated data, and being able to link behaviour to traffic conditions, e.g. urban / nonurban and traffic congestions (Technische Universiteit Delft; 2008, Volvo Trucks North America Inc., 2007).

Market penetration forecasts: In the literature, SEiSS (VDI/VDE-IT, 2005 and Baum et al., 2006) gives particular attention to market penetration.

Usage, reliability and compliance: Although the CAS Benefits Study (NHTSA, 1996) made assumptions about usage, reliability and compliance rather than gathering data, it did draw attention to these important factors in the out-turn effectiveness of ICT systems. Usage refers to the percentage of drivers (or of driving time) for which ICT systems installed on the vehicle will be switched-on and active. Reliability refers to the likelihood that that ICT systems will operate without failure, technically. Compliance refers to the percentage of occasions on which the driver's behaviour complies with warning or indication provided by the system.

Attitudinal and acceptance data: Many FOTs gathered attitudinal and acceptance data (Regan et al., 2006; UMTRI et al., 2006; USDoT, 1999; NHTSA, 2006; and Volvo Trucks North America Inc, 2007).

Costs of the ICT systems: See section 10.2.3. In some FOTs, data has been gathered which inputs directly into the maintenance and operating cost calculations (Volvo Truck FOT). Where the assessment period is longer than the expected service life of the equipment, replacement costs should be included (e.g. in the Freightliner FOT one round of replacement was included since the service life was 10 years and the assessment period 20 years; Battelle Memorial Institute, 2003).

Generic Data

The key items of generic data likely to be needed are:

National and EU-level network, fleet and traffic data, which are used in scaling-up the findings from the FOT to the level of political interest: The International Road Traffic and Accident Database, IRTAD (ITF, 2008) contains traffic data for the EU27. This includes vehicle kilometres on the total road network, vehicle kilometres on motorways, and vehicle kilometres on urban roads. Vehicle kilometres on rural roads can be derived; some data are missing.

ProgTrans European Transport Report (latest version: 2007/08) can be used as a valid source for forecasts. It contains vehicle stock and vehicle kilometres for 1) cars, 2) buses and coaches and 3) goods vehicles. Generally, the report covers past, present (incl. short-term forecasts for the next years) and future (longer term for selected target years). In the 2007/08 report the following years are covered: 1995, 2000, 2005, 2006, 2007, 2008, 2015 and 2020. Geographically, they cover EU-27 by member state plus some more (Norway, Switzerland, Croatia, Turkey, Belarus, Russia, Ukraine plus China, Japan and the USA).

Accident data (accidents, fatalities, severe and slight injuries) for base scenario: National databases are available. At the EU-level, the collection and compilation fo accident data as a basis for the safety impact assessment is a challenge, especially when specific target accidents are going to be explored. Several EU-projects are dedicated to harmonising accident databases, See TRACE (www.trace-project.org) or SafetyNet (http://www.erso.eu/safetynet/content/ safetynet.htm) for more information. Forecasts of road safety are needed. An example can be found in eIMPACT "Impact Assessment of Intelligent Vehicle Safety Ssytems", in which road safety predictions for 2010 and 2020 for the EU-25 are presented.

More detailed network specifications (e.g. infrastructure equipment) may be required for some systems: the presence/absence of beacons, signalisation. Basic figures (e.g. share of Trans-European Road Network equipped with dynamic traffic management) are available from the eSafety Forum Implementation Roadmap Working Group (2005).

Speed-flow relationships or network models, which allow journey times and costs to be derived from changes in flows: Although these are strictly much more than just 'data', it is worth highlighting the key role they play in socio-economic assessment of transport ICT systems. Many of the effects of new ICT systems will be mediated through changes in traffic flow on the network – for example, advanced warning ICT systems allow drivers to change route to avoid hazards, but the net effect on travel times and costs is dependent not only on the behaviour of the individual, but on the behaviour of large numbers of individuals and the interaction with the limited capacity of the network. Hence, at the very least, knowledge of speed-flow relationships is needed to understand the consequences of shifting traffic across the network.

HCM (2000) and FGSV (2001) are sources of speed flow relationships. Network models or strategic transport models incorporate this data and have much wider functionality. The fact that these models are very expensive to develop and maintain means that they tend not to be developed for one socio-economic assessment in isolation. Instead, part of the socio-economic assessment process is usually to identify models already existing which can provide the necessary functionality.

Evidence on accident costs, used to measure the benefits of accident reduction and changes in accident severity: The HEATCO project (Bickel et al., 2006) was designed specifically to provide harmonised cost estimates for socio-economic assessment in Europe. We recommend that the HEATCO accident cost values are used in the FOTs, and we provide one additional piece of evidence to fill a gap in HEATCO which is a generic dataset on the costs of 'damage only' accidents.

Two of the main issues in this field are:

- An apparent inconsistency between 'willingness-to-pay' (WTP) methods and 'cost of damage' or 'human capital' methods as a basis for values — empirically, WTP methods can produce significantly higher values for fatalities in particular (see Assing et al., 2006: Table 16);
- Double counting of casualties' lost future consumption, which is included in both lost future output and WTP to reduce accident risk.

HEATCO addresses these issues by specifying a common framework in which the different elements of accident costs measured by each method can be reconciled. For example, 'human capital' methods do not capture people's full valuation of safety risk, whilst WTP-based values do not capture the external resource costs of accidents (e.g. healthcare costs borne by the state) but often do double-count lost future consumption, as already noted. The HEATCO framework includes:

- property damage;
- medical costs;
- administration costs;
- lost output;
- welfare losses due to casualty reduction.

As a result, the HEATCO values for fatalities are neither as high as the US NHTSA's willingness-to-pay values cited by Assing et al., nor are they as low as the NHTSA's cost-of-damage values. They are broadly in line with 'best practice' European values used in cost-benefit analysis, and the differences can generally be understood by examining the differences in the underlying measurement methods.

Other important functions of the HEATCO values are to provide:

- A common unit of account in the face of taxes and subsidies – HEATCO values are provided at the factor cost unit of account (Bickel et al., 2006: 52);
- A common price base year;

- A common currency, €, for European-level assessments.

In HEATCO accident values are listed in a table, expressed as values per casualty saved. These values do include the full set of accident costs, per casualty.

To apply these values, analysts will require further data:

1. Forecasts of accidents with and without the technology in place – based on the FOT findings and the results of the scaling up process.

2. If these forecasts do not address unreported accidents, then factors for the number of unreported accidents given the number of reported accidents can be found in HEATCO (Bickel et al., 2006, Table 5.1).

3. Growth in the values over time – an elasticity of 1.0 with respect to GDP per capita, thus a 2.0 % annual increase in GDP per capita would imply a 2.0 % annual increase in the values of accident reduction.

4. Damage only accident values. As Baum et al (2007) shows, savings in damage only accidents can make up a large proportion of the benefits from ICT safety systems. Damage only accident costs may be approximated at 17 % of the cost of a Slight Casualty (Nellthorp et al., 1998).

National level assessments may wish to take advantage of the most recent safety valuation evidence at the national level. For multi-national assessments it will be important to ensure that any national evidence is checked for consistency across boundaries, and conversions made if necessary (e.g. in terms of base year, unit of account, cost elements included, measurement methodology, etc).

For EU-level assessments, consistency across countries and comparability between assessments will be important, which makes the use of a harmonised set of values (as above) more attractive. If the harmonised values are found not to provide the detail which the analyst wants — e.g. if differentiated accident costs by road type or user type are expected to be a key requirement for a particular assessment — then it may be appropriate to vary the values above, based on more detailed information (for example, the accident cost data included in national level assessment guidelines).

Evidence on values of time savings and vehicle operating cost savings, used to measure the benefits of changes in traffic flow: Values of travel time savings will be needed to assess the benefits of improved traffic flow due to the ICT systems. HEATCO Tables 4.6-4.8 provide suitable values for working and non-working passenger trips, and for freight transport (Bickel et al., 2006: 73-75). These values increase with GDP per capita, at an inter-temporal elasticity of 0.7.

Sometimes there will be an impact on reliability, not only expected (mean) travel times, and in these cases we recommend using the reliability ratios set out in HEATCO Table 4.3 to value changes in the standard deviation of journey time.

Vehicle operating cost savings are also likely to arise from changes in traffic flows. The traffic models used to predict traffic flow responses to ICT systems will typically be capable of predicting changes in Vehicle Operating Costs, and the fine network detail in these models usually makes it more logical to calculate these cost savings within the model, rather than attempting to do so based on model outputs. As a result, standard values are not offered for these impacts by HEATCO (see Bickel et al., 2006: 135-140).

Emissions factors and values for the damage caused by emissions of greenhouse gases, air pollutants and noise: HEATCO provides values for both sources of emissions (Bickel et al., 2006, Tables 6.2, 6.4). Values for particulate (smoke) emissions are differentiated between urban and non-urban locations, due to their much localised impact pathway. Other air pollutants are valued uniformly at country level. HEATCO provides a shadow price of CO_2 by year of emissions (Bickel et al., 2006, Table 6.12) which should be applied to all forecast changes. The impact of noise changes may be quantified using the HEATCO values for road, rail and aircraft noise in each member state (Bickel et al., 2006, Table 6.9).

Other economic parameters such as the social discount rate: Discount rates are required for socio-economic assessment. In line with HEATCO, we recommend using a risk-free social time preference rated for the countries to which the assessment would apply. If a default discount rate at the EU level is required, we would recommend using 3 % per annum (real). GDP growth data for the members of the EU27 (required for updating values of accidents, etc, over time) is available from Eurostat (ec.europa.eu/eurostat).

Data quality and validity

The EC ROSEBUD project provided the following guidance as part of a "professional code for analysts" (BASt et al., 2005, p. 46):

"Data has to be attributed correctly to its sources, especially when different data sources like national or international accident databases or in-depth databases are used. Where and how estimations were made to fill data gaps needs to be documented. Regression models should be used to generate future time series; trend extrapolations can replace them where available data are insufficient for regressions".

In addition, we would recommend that:

- The principles of statistics apply — statistical tests should be used wherever possible to determine if hypotheses about ICT system impacts are supported by the FOT evidence, and sample sizes should be chosen to obtain statistically significant results;

- When scaling-up from the FOT to the national or EU27 level, a methodical approach based on the key drivers of safety/other significant outcomes identified in the FOT should be used (cross reference);

- Confidence intervals as well as mean data should be recorded for key variables – note that confidence intervals are given in HEATCO for the various economic parameters recommended;

- We have noted the need to recognise the uncertainty in the data using sensitivity analysis – if analysts wish to take a more advanced approach and use Monte Carlo simulation or related techniques (for example to derive a probability distribution on NPV or BCR) that would be welcome as it simplifies the outputs seen by the decision-makers, although it does place an additional burden on the analysts;
- Known problems with the data should be acknowledged and acted upon, e.g. UMTRI et al. (2006) excluded a proportion of drivers whose trials were invalidated (in that case 9 out of 87 drivers), and some trips by the remaining drivers. Well-known problems with the omission of unreported accidents from data have prompted Bickel et al. (2006, Table 5.1) to provide adjustment factors for different accident severities and types.

Record keeping and data storage are important. This includes qualitative/subjective data, and evidence gathered during deliberative studies (e.g. UMTRI et al (2006) ensured that focus group evidence was captured on video and by a court stenographer).

Finally, the US NHTSA observes that "the validity of any experimental test results depends on the experimental condition effects that were placed on the drivers" (NHTSA, 1996: p36). Care is needed, therefore, when extrapolating data from short-term experiments to long-term term adjustments in behaviour and demand for ICT systems — e.g. the CAS Benefits Study found that "a better estimation of the safety benefits … can be achieved as more relevant test data are gathered especially from long-term, large-fleet field operational tests" (p. C-8).

11 References

Aarts, L. and van Schagen, I. (2006). Driving speed and the risk of road crashes: a review. Accident Analysis and Prevention, 38, pp. 215-224.

Abhishek, V. (2008). European Telematics & ITS. A progress report on V2X communication in Europe.

Assing, K., Baum, H., Bühne, J.-A., Geißler, T., Grawenhoff, S., Peters, H., Schulz, W. H. and Westerkamp, U. (2006). Methodological framework and database for socio-economic evaluation of Intelligent Vehicle Safety Systems. Deliverable D3, eIMPACT: Socio-economic impact assessment of stand-alone and co-operative Intelligent Vehicle Safety Systems (IVSS) in Europe, EU 6th Framework Programme Contract no.: 027421. TNO, Delft. www.eimpact.eu

BASt, DITS, KuSS, TRL, UOC and CDV (2005). Road safety and Environmental Benefit-Cost and Cost-Effectiveness Analysis for Use in Decision Making, Deliverable D7. Funded by EC DG-TREN, under project GTC2/2000/33020. http://partnet.vtt.fi/rosebud/index.html

Batelle Memorial Institute (2003). Evaluation of the Freightliner Intelligent Vehicle Initiative Field Operational Test, Final Report to the US Department of Transportation, Project DTFH61-96-C-00077 Workorder 7718. USDoT, Washington DC. http://www.itsdocs.fhwa.dot.gov//JPODOCS/REPTS_TE//13871.html

Baum, H., Geißler, T., Grawenhoff, S. and Schulz, W. H. (2006). Cost-Benefit Analyses of Intelligent Vehicle Safety Systems – some empirical case studies. Zeitschrift für Verkehrswissenschaft. 77(3), pp. 226-254.

Baum, H., Grawenhoff, S. and Geißler, T. (2007). Cost-Benefit-Analysis of the Electronic Stability Program (ESP). Summary Report. Institute for Transport Economics, University of Cologne.
www.chooseesc.eu/download/press/University %20of %20Cologne_ESC_cost_benefit.pdf

Besseling, H. and van Boxtel, A. (2001). Intelligent Speed Adaptation — Results of the Dutch ISA Tilburg trial. Ministry of Transport, Public Works, and Water management, AVV Transport Research Center.

Bickel, P., Friedrich, R., Burgess, A., Fagiani, P., Hunt, A., De Jong, G., Laird, J., Lieb, C., Lindberg, G., Mackie, P., Navrud, S., Odgaard, T., Ricci, A., Shires, J., Tavasszy, L. (2006). Developing Harmonised European Approaches for Transport Costing and Project Assessment, HEATCO Deliverable 5. Germany. http://heatco.ier.uni-stuttgart.de/

Carsten, O, Fowkes, M., Lai, F., Chorlton, K., Jamson, S., Tate, F. and Simpkin, B. (2008). ISA-UK: Final Report. Institute for Transport Studies, University of Leeds.

Cochran W. G. (1977). Sampling techniques. 3rd edition. Wiley & Sons, New York.

COWI, ECN, Ernst & Young Europe, ECORYS (2006). Cost Benefit assessment and prioritisation of vehicle safety technologies. Final Report.

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., and Knipling, R. R. The 100-Car Naturalistic Driving Study, Phase II — Results of the 100-Car Field Experiment Virginia Tech Transportation Institute, Blacksburg, VA, NHTSA DOT HS 810 593.

Draskóczy, M., Carsten, O. and Kulmala, R. (1998). Road safety guidelines. Deliverable B5.2 of CODE project (TR1103). Atkins Wootton Jeffreys, Birmingham, UK.

Ehrlich, J. et al. (2003). LAVIA, The French ISA project: main issues and first technical results, the French ISA project. Proceedings of the 10th World Congress on Intelligent Transport Systems. Madrid, Spain, 16-19 November 2003.

Elvik, R., Christensen, P. and Amundsen, A. (2004). Speed and road accidents: an evaluation of the Power Model. TOI Research Report 740/2004, Institute of Transport Economics, Oslo.

Ervin, R. D., Sayer, J., LeBlanc, D., Bogard, S., Mefford, M., Hagan, M., Bareket, Z., Winkler, C. (2005). Automotive collision avoidance system field operational test methodology and results, Volume 1: Technical Report. UMTRI-2005-7-1; University of Michigan, Ann Arbor, Transportation Research Institute.

Everitt, B. and Dunn, F. (2000). Applied multivariate data analysis. Arnold Publication.

ESOP 99 Commission of the European Communities (1999). Commission Recommendation of 21 December 1999 on safe and efficient in-vehicle information and communication systems: A European statement of principles on human machine interface. Doc. no. C(1999) 4786, 2000/53/EC.

Fancher, P., Ervin, R., Sayer, J., Hagan, M., Bogard, S., Bareket, Z., Mefford, M. & Haugen, J. (1998). Intelligent Cruise Control Field Operational Test. Final Report UMTRI-98-17; University of Michigan, Ann Arbor, Transportation Research Institute.

FGSV (2001). Forschungsgruppe für Strassen- und Verkehrswesen. Handbuch für die Bemessung von Straßenverkehrsanlagen. FGSV, Köln. http://www.fgsv.de/

Gibbons, J. D. and Chakraborti, S. (2003). Nonparametric statistical inference. 4th edition. CRC.

Goldstein, H. (2003). Multilevel statistical models. 3rd edition, Arnold, London.

HCM (2000). Highway Capacity Manual. Transportation Research Board, Washington D.C.

Hjälmdahl, M., Várhelyi, A. (2003). Speed regulation by in-car active accelerator pedal effects on driver behaviour. Transportation Research Part F. Vol. 7, Issue 2, pp. 77-94.

ITF (2008). The International Road Traffic and Accident Database (IRTAD). www.irtad.net

Jacobson, I., Bylund, S., Jonsson, P., and Ehneboom, S. (1995), "Modeling with Use Cases: Using contracts and use cases to build plugable architectures". Journal of Object Oriented Programming, Vol. 8, No. 2, pp. 18-24.

Kato, K. et al. (1999). Plan for verification experiments of advanced cruise-assist highway system (AHS) in Japan. Proceedings of 6th World Congress on Intelligent Transport Systems (ITS), held Toronto, Canada; http://www.ahsra.or.jp/eng/d01e/index.html

Lassarre, S. and Saad, F. (2006). Présentation générale du dispositif expérimental: justification des choix. In: Carnet de route du LAVIA. Limiteur s'adaptant à la vitesse autorisée. Paris: Actes du colloque LAVIA, pp. 11-17.

Lassarre, S. and Romon, S. (2006). Utilisation du LAVIA et influence sur les vitesses pratiquées en vue de l'évaluation de l'utilité. In: Carnet de route du LAVIA. Limiteur s'adaptant à la vitesse autorisée. Paris: Actes du colloque LAVIA, pp. 53-60.

Lebart, L., Morineau, A. and Piron, M. (1997). Statistique exploratoire multidimensionnelle. Dunod, Paris.

Lecoutre, J. P. and Tassi, P. (1987). Statistique non paramétrique et robustesse. Econometrica, Paris.

Li, B. (2004). The latest development of intelligent highway system in China. Proceedings of the 11th World Congress on ITS, Nagoya, Japan.

Michon, J. A. (1985). A critical view of driver behaviour models: what do we know, what should we do? In: L. Evans and R. C. Schwing (Eds.). Human Behaviour and Traffic Safety, pp. 485-524. New York: Plenum Press.

Najm, W. G., Stearns, M. D., Howarth, H., Koopman J. and Hitz, J. (2006). Evaluation of an automotive rear-end collision avoidance system. DOT HS 810 569. DOT VNTSC-NHTSA-06-01. Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.

NHTSA Benefits Working Group (1996). Preliminary Assessment of Crash Avoidance Systems Benefits. NHTSA, Washington D.C. http://www.trb.org/news/blurb_detail.asp?id=1166

Nilsson, G. (2004). Traffic safety dimensions and the power model to describe the effect of speed on safety. Bulletin 221. Lund Institute of Technology, Lund University. http://www.lub.lu.se/luft/diss/tec_733/tec_733.pdf

Oguri, Y. (2007). Viewing traffic safety issues from victims' standpoints – developing wider social acceptance of safe speed initiatives. ICTCT Extra Workshop, Beijing.

Orban, J., Hadden, J., Stark, G. and Brown, V. (2006). Evaluation of the mack intelligent vehicle initiative Field Operational Test; Final Report. FMCSA-06-016. Batelle, Columbus, Ohio.

Page, J. (2004). A final technical report on the Belgian intelligent speed adaptation (ISA) trial. 11th World Congress on ITS in Nagoya, Japan. IBSR-Report V05/04 available in French and Dutch.

ProgTrans (2004). European Transport Report. Basel.

Regan, M. A., Lee, J. D., and Young, K. L. (2008). Driver distraction: Theory, effects and mitigation. Florida, USA: CRC Press. (In Press – available October 2008).

Regan, M., Triggs, T., Young, K., Tomasevic, N., Mitsopoulos, E., Stephan, K., and Tingvall, C. (2006). On-road evaluation of intelligent speed adaptation, following distance warning and seatbelt reminder systems: Final Results of the Australian TAC SafeCar Project. Volume 1: Report. Monash University Accident Research Centre Report 253. MUARC: Melbourne, Australia. http://www.monash.edu.au/muarc/reports/muarc253.html

Regan, M., Triggs, T., Young, K., Tomasevic, N., Mitsopoulos, E., Stephan, K., and Tingvall, C. (2006). On-road evaluation of intelligent speed adaptation, following distance warning and seatbelt reminder systems: Final Results of the Australian TAC SafeCar Project. Volume 2: Appendices. Monash University Accident Research Centre Report 253. MUARC: Melbourne, Australia. http://www.monash.edu.au/muarc/reports/muarc253.html

Saad, F. (1997). Contribution of observation and verbal report techniques to an analysis of road situations and drivers' activity. In: T. Rothengatter and Carbonell Vaya (Eds.). Traffic and Transport Psychology, Theory and Application, pp. 183-192. Pergamon.

Saad, F. (2006). Some critical issues when studying behavioural adaptations to new driver support systems. Cognition, Technology and Work, 8, pp. 175-181.

Saad, F. and Dionisio, C. (2007). Pre-evaluation of the "Mandatory Active" LAVIA: assessment of usability, utility and acceptance. In: Proceedings of the 14th World Congress and Exhibition on Intelligent Transport Systems and Services. 8-12 October 2007, Beijing, Paper 2257.

Sanghoon, B. (1998). Evaluation of ITS Field Operational Test in Kwachon Korea. The 5th ITS World Congress, Seoul, Korea, October 12-16, 1998.

Sanghoon, B. (1998): Issues and lessons learned from ITS Field Operational Test in Korea. The 1st Asia Pacific Conference on Transportation and the Environment, Singapore, May 13-15.

Särndahl, C.-E., Swensson, B. and Wretman, J. (1992). Model assisted survey sampling. Springer Verlag.

Taylor H. and Karlin S. (1994). An introduction to stochastic modeling. Academic Press.

Technische Universiteit Delft, Faculteit Civiele Techniek en Geowetenschappen, Afdeling Transport & Planning (2008). Full Traffic – WP Dataloggers & WP Verkeersimpact. TU Delft, Netherlands.

University of Michigan Transportation Research Institute (UMTRI), Visteon Corporation and AssistWare Technology Inc. (2006). Road Departure Crash Warning System Field Operational Test: Methodology and Results. Volume 1: Technical Report and Volume 2: Appendices. NHTSA, Washington D.C.
http://deepblue.lib.umich.edu/bitstream/2027.42/49242/1/99788.pdf and
http://deepblue.lib.umich.edu/bitstream/2027.42/49242/1/99789.pdf

US Department for Transportation (2005). Volvo Trucks Field Operational Test: Evaluation of Advanced Safety Systems for heavy truck tractors. Final Report. DTFH61-99-X-00102. Washington, DC: US DOT Federal Highway Administration.

US Department of Transportation, NHTSA (2006). Automotive Collision Avoidance System Field Operational Test Report: Methodology and Results. http://deepblue.lib.umich.edu/bitstream/2027.42/49539/1/99798.pdf

VDI/VDE-IT and ITE University of Cologne (2005). Exploratory Study on the potential socio-economic impact of the introduction of Intelligent Safety Systems in Road Vehicles (SEiSS). Final Report to the European Commission DG Information Society and Media. www.esafetysupport.org/en/esafety_activities/related_studies_and_reports

Volvo Trucks North America Inc (2007). Volvo Trucks Field Operational Test: Evaluation of Advanced Safety Systems for heavy truck tractors. Final Report to US Department of Transportation, Federal Highway Administration, Cooperative Agreement No. DTFH61-99-X-00102. USDoT, Washington DC. http://www.itsdocs.fhwa.dot.gov/JPODOCS/REPTS_TE/14352.htm

Wannacott, T. and Wannacott, R. (1990). Introductory statistics for business and economics. 4th edition. John Wiley & Sons.

Wasserman L. (2007). All of nonparametric statistics. Springer.

Wilmink, I., Janssen, W. Jonkers, E., Malone, K., van Noort, M., Klunder, G., Rämä, P., Sihvola, N., Kulmala, R., Schirokoff, A., Lind, G., Benz, T., Peters, H., Schonebeck, S. (2008). Socio-economic Impact Assessment of Stand-Alone and Co-operative Intelligent Vehicle Safety Systems (IVSS) in Europe. Deliverable D4. eIMPACT: Socio-economic Impact Assessment of stand-alone and co-operative Intelligent Vehicle Safety Systems (IVSS) in Europe, EU 6th Framework Programme. TNO, Delft, 2008. http://www.eimpact.eu

Annexes

Annex A Legal and ethical issues in the execution of FOTs – Worked Example

The aim of this chapter is to sensitise the reader to the legal issues that will prove to be relevant in planning and carrying out a Field Operational Test (FOT). Due to the fact that the details of future field tests cannot be foreseen, all obviously relevant legal areas are covered; necessarily giving abstract information devoid of any warranty as far as completeness and accuracy for the concrete test arrangement is concerned. Considering the legal importance of details in test arrangements, it must be pointed out that it is vital to involve legal expertise from the country in question when planning a Field Operational Test. The overview given here can, furthermore, not substitute legal advice in a particular case.

For the purpose of well-structured information, a chart is included in the sub-section A.1. In this chart a classification in terms of system design and superior legal aspects has been taken out. Each segment refers to relevant sub-sections that deliver respective information. It is meant for fast reading.

A.1 Legal definition of “Field Operational Test”

Before the legal issues related with FOTs are discussed in detail, it must be defined that for the purpose of this section on legal and ethical issues, FOT is considered to be a test arrangement that is accomplished within real-life traffic conditions. This implies that an unknown number of persons not involved in the actual testing procedure, form the surrounding traffic. Usually third parties will not even know about testing being performed. Thus this first characteristic feature excludes artificial, isolated test arrangements and has important legal implications over all legal issues in question.

A.2 Information for test persons (briefing) / contractual agreements

The legal relationship between the organisation carrying out the field test and the test person will most likely have to be agreed upon in a contract.

A further characteristic feature of field operational testing is the data acquisition. In some way the driving will be recorded, possibly even the location might be tracked or videos of the driver and/or surrounding traffic recorded. This has an influence in terms of test persons’ and third persons’ data privacy and will partially be subject to consent on their side.

A.2.1 Preliminary considerations

To give substantial legal advice on exactly which information must be provided to test persons and estimate which arrangements (insurance, exceptional licenses, etc.) are necessary, rather detailed knowledge on the experimental setup and the systems to be evaluated is required. In order to conclude a contract with test persons, very detailed knowledge on the exact test design and testing procedure is necessary (and this at an early stage of the FOT).

As certain legal consequences may turn out to be unwanted, the following should be taken into account right from the beginning in order to adjust the design of the FOT accordingly.

A.2.2 Information provided to test persons

In order to obtain a valid consent from the driver to log the data and allow for safe participation in the FOT, information on the testing procedure and setup must be comprehensive. This does not include any technical knowledge about the layout itself, but all the consequences for the user: This will make the provision of information necessary on which kind of data is being logged in the first place and who will have access (esp. in case of accidents, administrative fines, etc. this will be important (see section A.3.2.2; A.4; A.5.4; A.5.6)). Possible legal consequences in case of dangerous driving by the test person should be specified (like possible recourse in case of grossly negligent or intentional behaviour – if applicable (and considered appropriate)). Special attention must be paid towards boundaries and how to deal with malfunctions (the two must be distinguished, see A.3.2.1; A.3.2.2). The same must be considered in case of a possible overload of the driver in terms of information, warnings, etc. That the responsibility for safe use as for administrative fines remains with the driver must be pointed out explicitly (see section A.4) as must be the fact that the driver remains fully responsible for his/ her driving and is not exempt from full responsibility due to participation in the FOT (this may not fully be applicable in case of non-overrideable systems, however, this will lead to many further questions, see section A.8.). (Regan, 2006, VOL 2)

Apart from this, the effect of the systems on driving – especially in case of some kind of unusual interference – should be pointed out, too. (This might be the case e.g. when applying a visual, acoustic or haptic warning-strategy that might unsettle the driver. (Regan, 2006, VOL 2] Special information is particularly advisable in case of any interference into steering or braking, etc.). The risk here is to not sufficiently prepare the test person for safe use of the system. This may under certain conditions lead to liability of the responsible research scientist/ head of department and possibly to liability of the organisation (if possible according to national tort law). Apart from this, a researcher (even negligently) causing damage to the health of a test person may be considered criminally liable for unsafe test design, insufficient instructions and many further substantial breaches of his responsibility with obviously negative connotations for safety.

As a rule of thumb, about the same information and warnings, etc. should therefore be provided that would be necessary in a driver's manual, in case the system is meant to be used under real-life traffic conditions without further surveillance. Thereby, reasonably foreseeable misuse must be taken into account. The information should be provided in a way that the least informed test-person, who is therefore most exposed to a danger, can drive safely. The provision of information and warnings can, however, be achieved otherwise than in a written manual: Personal briefing, presentations on how to handle a system under certain conditions or the training of drivers with the possibility to experience the functioning and ask questions are legally sufficient as well. The possibility to ask questions at any time later during the FOT should be provided for e.g. by means of a telephone hotline (again in order to avoid insufficient briefing which may result in unsafe use by test persons). Furthermore, it might ex post turn out to be difficult to prove that a certain piece of information (that would have been necessary to avoid an accident) has actually been provided to the respective test person. Therefore it seems advisable to incorporate at least the most important information into the contract with a test person in form of some kind of notice or to refer to another document that has made the information available (see Regan, 2006, volume 2).

A.2.3 Information on system boundaries

A special issue in the context of briefing is system boundaries. System boundaries are those features of a system that are not a defect but still lead to wrong information or an erroneous intervention due to a lack of overall system intelligence (for further details and examples, see Deliverable D6.3).

As system boundaries always occur in certain situations, they are predictable and will not bring about liability issues as long as they have sufficiently been made clear to the test person. A test person who is able to anticipate the system behaviour in the case of all system boundaries will in this respect be considered well-informed.

A.2.4 Information on possible malfunctions

In case the FOT is taken out to evaluate a premature system, possible malfunctions will usually have to be taken into account too. Most important in case of malfunctions will again be to give test persons all the information necessary. That is first of all to provide the information, that a malfunction can occur and instruct thoroughly how to deal with the resulting situations (Regan, 2006, VOL 2). If technically feasible, recognisability of malfunctions should be made possible. In most cases it will be sufficient to provide for the possibility to switch off the system and thus ensure safety. Even this might not be necessary as long as the malfunction will not impair safe driving at all. In case of intervening systems, however, much depends on the period of time available for a reaction of the driver: If this is too short, safety will potentially be impaired by any (disturbing) intervention.

For details on possible malfunctions related to the respective systems' intervention or information concept, see Deliverable D6.3.

A.2.5 Information on data recording

As far as data privacy is concerned, details are provided in a separate section (see A.5). For the briefing of test persons it is important to point out the relevant issues for data processing as well as access rights (Regan, 2006, VOL 2).

It is legally required that the driver knows which data is being logged. It should also be pointed out, which conclusions can be drawn from the data available and this should involve all imaginable data sources and their combination (including external sources that can be resorted to). The meaning of anonymisation and pseudonymisation as well as any other measures to achieve data privacy should be described too. In case de-personalisation of data is possible and intended, it must be pointed out at which point of data-handling this is realised.

Example a: Within a FOT, data on speed as well as location is recorded. It is possible to anonymise the data for scientific use. However, when logging the data in the car, it can naturally be traced back to the driver (even if personal information is not logged). In this case, anonymisation might come into effect as soon as the data is read into a database with many similar recordings so that traceability of the test person is barred. Traceability would also be barred, in case the advice in section A.5.4 is applied. However, the risks of accessibility in spite of these measures (see section A.5.6) – e.g. until pseudonymisation has been realised – should be pointed out too (in order to avoid incomplete information).

The most important measure to comply with data privacy regulations will be to inform the test person thoroughly as far as data acquisition is concerned and (voluntarily!) gain his/ her consent. This consent must – due to the considerable impairment of data privacy combined with FOTs – be stipulated in written form. For all further details and advice, see section A.5 in this handbook (and details in the related Deliverable D6.3).

A.2.6 Agreements on cost allocation and liabilities (including insurance issues)

Another important aspect in terms of contractual agreements is the allocation of costs as well as special agreements in terms of liability. Some aspects will most certainly be regarded appropriate; some might seem disadvantageous in light of volunteer recruitment for a FOT. However, the possibilities in terms of contractual agreements are broad as long as true freedom of decision is ensured (and participation must be voluntary anyway).

Appropriate agreements within a FOT will e.g. presumably be agreements on the allocation of fuel costs that will be borne by the test person. It may furthermore be regarded adequate to agree on a certain sum per mileage for the use of the test vehicle (as long as the vehicle can be employed in everyday use). This again may be combined with other agreements – in case of long term testing – within a lease contract, etc. (Regan, 2006, VOL 1).

Of great importance in so far will be the agreements concerning the presumably valuable equipment for data acquisition (and possibly the units installed for evaluation). Here agreements on liability might be necessary as might be a special insurance in order to avoid a financial strain on the test person (and solve this conflict pragmatically). Compare section A.6.

Special agreements will be necessary on data provision by the test persons. As this will mostly be personal data, it shall in so far be referred to section A.5. However, it should be noted that apart from all the agreements necessary in terms of data privacy itself, agreements will also be necessary on how often data shall be retrieved, how this shall take place and the whereabouts of e.g. vehicle return, possibly the demounting of data acquisition components or systems (in case the vehicle remains in the property of the test person). In the latter case special attention must also be paid to possible damages brought about by installation of the FOT-equipment and how these shall be dealt with (Regan, 2006, VOL 2).

In case the vehicle does not belong to the property of the test person, special agreements might be necessary in order to assure that the car is not used for dangerous driving. This will be evident from the data retrieved and in severe cases an obligation to intervene might even be brought about as the knowledge on the side of the researcher is evident (and the participation in a FOT might even provoke dangerous driving depending on the test-persons character). Therefore an appropriate contractual obligation may be stipulated by agreeing on immediate termination of testing, in case dangerous driving is observed (Regan, 2006, VOL 2). However, it must be pointed out that this knowledge on dangerous driving usually belongs to the private sphere of the person concerned. In case this knowledge would e.g. be disclosed towards the employer, this would severely compromise the test person and must therefore be dealt with in compliance with the guidelines depicted below (see section A.5). Any disclosure to third parties must therefore – all the more – be refrained from.

Another important aspect in terms of liability of the researcher is to ensure that the test-person is fit to participate in the FOT. This will definitely not imply any detailed inquiries as far as health is concerned (and this would even be considered intimate knowledge in terms of data privacy). Yet, the researcher should not allow a test person to participate in case an unfavourable medical condition is obvious. Apart from this, it might be a good idea to enlighten the need of good health in the FOT-information provided (especially e.g. as far as eyesight is concerned) and this might also be included in the contract dealing with all the details of FOT-participation. The same is true for any substance abuse (see Regan, 2006, VOL 1).

Furthermore, information should be provided on the insurances concluded for the test vehicle (in order to point out remaining risks). Depending on the FOT model chosen, this might, of course, only be a recommendation to the test person on which insurances should be concluded (and may even be left completely to the test person in case the systems to be evaluated are mature, in no way critical and the test person owns the vehicle participating in the FOT). Special attention must be paid towards the insurance of data-logging equipment and special agreements might have to be made/ insurance issues pointed out to provide for sufficient information (see section A.6.3).

A.3 Administrative fines

In Germany, administrative fines are related to the personal responsibility of the perpetrator. If traditional driving is considered, no doubts exist on whether responsibility for any breach of traffic law remains with the vehicle driver. However, even systems that only provide information to the driver, tend to point out that traffic rules and traffic signs are prior to information provided (e.g. navigation systems). Therefore the following possibilities for system-design must be considered separately:

A.3.1 Informing Systems

As far as informing systems are concerned, two different types of information must be distinguished: First of all, information may be (more or less) legally irrelevant (e.g. a system providing information on present fuel consumption). Often, especially in case of safety-relevant ADAS, the information will, however, be legally relevant after all. Here again it must be distinguished: On the one hand there is information e.g. on legal speed limits, sign-posted dangerous bends or information provided by road traffic codes. This kind of information has a direct legal implication as it is directly linked to the provisions of road traffic and thus to the conduct legally required. On the other hand, the information that lacks this direct link may become legally important, e.g. in terms of compensation for damages. The latter is, however, much subject to the contractual agreement and information provided to the test person (for further information on this, see section A.3).

As far as those informing systems are concerned that have been circumscribed to be directly linked to the provisions of road traffic, the question might arise, whether false information provided by the system will excuse or charge (as the case may be) the driver in terms of an administrative fine.

Example A.4.1a: The driver negligently misses a sign-posted speed limit of 50 km per hour at the road side. His car is equipped with a speed alert system so he checks the speed limit displayed there. For some reason the information provided is, however, wrong, a speed limit of 70 km per hour is displayed. The driver relies on the information of his system and drives at 70 km per hour. The driver is fined for speeding.

In terms of administrative fines it does not matter how the driver has been instructed (at least, if the driver has been aware of the fact that he must generally comply with traffic rules when taking part in the FOT – this information must be provided to the test person, see Section A.3.). In example A.4.1a it can be expected from the driver to adhere to traffic signs: Only traffic codes and sign-posted traffic information are legally relevant (no in-vehicle-applications have been introduced in a legally relevant way so far). Because the driver misses the sign-posted speed limit negligently, he can be charged for speeding. All the other information (such as the display of the speed alert system) has no legal implication (it is only a factual "add-on"-information). So even though the driver in example A.4.1a only relies on the wrong information displayed, this will not excuse him legally in a way that the fine cannot be imposed on him (ALBRECHT, 2005).

Example A.4.1b: The driver is speeding and is additionally warned by a speed alert system that he is going to fast. Due to data collection in the car, the display and acoustical signal of the speed-limit warning is recorded. As a camera is also installed, it can be proved that the driver has noticed the warning provided on the display. The driver does, however, not reduce his speed and is fined.

In example A.4.1b the driver is – apart from the sign post or general traffic rule – additionally warned by the in-vehicle-application (such as a "speed alert" system) and has obviously been aware of the speed limit. Therefore his breach of traffic law might be considered intentional, which may have effect on the height/amount of the fine: In Germany e.g. it is generally assumed that speeding is a negligent act. In case intention can be proved – which would be promising given all the data recorded here – the fine will turn out to be higher (ALBRECHT, 2005). This problem is also dealt with in Section A.5 (data privacy issues) as far as data usage in terms of prosecution is concerned. In case it proves to be necessary to record this data, the test person concerned (driver) must at least be aware of the risk he/ she is running (which is again subject to the information provided by the organiser of the FOT).

A.3.2 Intervening, overrideable systems

As far as intervening, overrideable systems are concerned, most important is to point out that they must be actually overrideable in any case and at any time (otherwise see Section A.4.3). If they are overrideable, the driver is still fully responsible for every movement of his vehicle. Usually the intervention will either serve as a basis for information transmission (e.g. vibration of the steering wheel) or will simply intervene by carrying out (a part of the) driving task automatically (e.g. an Adaptive Cruise Control). As far as the transmission of information is concerned, the same will apply in terms of fines (this has been discussed in section A.4.1, see above).

In case the driving task is partially carried out automatically, the system boundaries and functioning of the system must be made completely clear (to provide for full control over the vehicle) and it must be pointed out that the responsibility – even for the aspect of the driving task carried out by the system – remains with the driver. It is therefore necessary to override the system, if this is legally required (and this must, of course, be possible!). All this is subject to the information provided to the test person (see section A.3). As full control over the vehicle will then still be immediately available, administrative fines can be imposed on the driver in case of a negligent or intentional breach of traffic law.

A.3.3 Intervening, non-overrideable systems

In case of intervening, non-overrideable systems, it should for the means of this handbook briefly be pointed out that these are generally considered non-permissible and call for exceptional licenses and a specific insurance (see section A.8 and A.6.4).

Apart from this, the driver is no longer capable of (fully) putting his will into execution as far as the control over his vehicle is concerned. In so far as the administrative fine arises from an aspect that no longer belongs to the drivers' control, the breach can no longer be considered negligent or intentional. Therefore administrative fines can no longer be imposed on the driver. [ALBRECHT, 2005]

A.3.4 Cooperative Systems

In case of cooperative systems many aspects may come to effect that have been discussed above in the sub-sections A.4.1, A.4.2 and A.4.3. The respective effect a cooperative system has within the car and the information on the cooperative system provided to the driver will then be respectively valid here. In other words, the same will then apply to cooperative systems.

A.4 Data privacy

A.4.1 Introduction/ general comments/ minimum standard within the EU

Data privacy is in Germany based on basic (constitutional) human rights (for Germany: Art. 1 para. 1 and Art. 2 para. 1 of the "Grundgesetz"= German constitution). This right is termed "informational self-determination" ("informationelle Selbstbestimmung"). The Federal Constitutional Court of Germany characterises this basic right as the authority of the individual to decide on the disclosure and use of his/ her personal data. This is substantiated with the argument that who cannot overlook which personal information is available in certain fields of his/ her social environment and therefore cannot estimate the knowledge of contact persons, can be substantially hindered to exercise his personal freedom of free decision and planning. This should well circumscribe the scope of protection data privacy acts bring about. [BfD-INFO 1, 2002]

Within Europe, the minimum standard of data privacy ("data protection") is stipulated by the EU Directive 95/46/EG. This directive was issued in 1995 to ensure data privacy of natural persons in the processing of personal data. The directive describes the minimum standard for data protection that must be guaranteed throughout the EU by national law (the directive itself is generally not directly exercisable). In Germany the directive lead to some modifications of national data protection acts such as the "Bundesdatenschutzgesetz" (BDSG). [BfD-INFO 1, 2002], [ROSSNAGEL, 2003]

The extent of protection by data privacy acts in Germany is rather dense. If therefore the data protection principles valid for Germany can be applied to the design of a FOT, it is most likely that this will be sufficient in terms of data protection for other countries of the EU too. However, it must be pointed out that the following statements can only claim definite validity for FOTs in Germany. In case of doubt, it seems advisable to contact the national data protection officer (if applicable for the respective country) for further advice (the same applies in case of any specific questions). It must also be pointed out that the standard of data protection can turn out to be lower in other countries, should, however, not drop below the minimum standard described in the EU-directive mentioned above. This minimum standard must be complied with, especially when taking out a FOT within an EU research activity. The minimum standard within the EU directive has also been referred to as far as possible.

A.4.2 Legally relevant data and general measures to ensure data privacy

Data privacy regulations are generally based on basic human rights. Therefore the scope of relevant data is restricted to personal data. Personal data are particulars on personal or factual relations of a defined or definable person. In some European Countries (Austria, Luxembourg, Denmark) even legal bodies are covered by data protection rules, however, as far as of interest for FOTs, data privacy of the natural person is in question (see Sec. 3 BDSG, Art. 3 EU Directive 95/46/EG). For further examples see Deliverable D6.3.

Anonymisation and pseudonymisation are measures to assure data privacy.

Anonymisation is the de-personalisation (a modification) of personal data. The data can then not be traced back to the natural person. However, it must be kept in mind, that complete anonymity cannot be achieved, in case the data is so particular that it will apply to only one person. Whether a data set can be considered anonymous, may be dependent on the number of particulars, the available methodological and mathematical instruments as well as the availability of additional information allowing re-personalisation. Therefore, anonymity must be considered a relative term.

In case of pseudonymisation the name or other identification criteria of a person is modified and replaced by a pseudonym (usually a multi-digit number, nick-name or combination of numbers and letters, the so called "code"). This will considerably complicate the identification of the person behind a data-set. However, in contrast to anonymisation the re-identification remains possible (and is not restricted to chance, mathematical or methodological instruments). With the help of the key that has been separated from the original data set – possibly a list linking the names to the code – re-identification can be achieved. The protection of privacy is much dependent on how well the separation of the key and data-set is ensured. If the key is destroyed, the data would be considered anonymised. [ROSSNAGEL, 2003]

A.4.3 Sub-constitutional law and general principles

For Germany the basic right of informational self determination has been further developed in sub-constitutional law. Such are the federal law on data privacy ("Bundesdatenschutzgesetz (BDSG)") as well as respective acts in every single federal state. Depending on the background of the organisation taking out the FOT (company (= private) or public authority) different measures are applied in terms of data privacy. For the means of this report, the description of legal framework will focus on the BDSG as this code is generally applicable in case of data privacy for private organisations (companies, etc.) and valid for all federal public bodies (does, however, not apply to the public bodies within the federal states which are large in number: For these the respective act in the respective federal state is applicable. The provisions tend to be very similar, though). [BfD-INFO 1, 2002]

As a rule of thumb, the provisions are, generally speaking, rather strict in case of data acquisition, processing and use by public bodies and more liberal in case of (private) companies. Speaking for Germany, this leads to the situation that only those private companies, institutions, etc. are subject to the federal law on data privacy that collect, process or use data by means of data processing equipment (automated or not). The use of personal data in any other way e.g. by private entities is not subject to the act in the first place. However, data processing equipment will be the rule for field operational testing, so data privacy restrictions will in so far be applicable.

The basic principle of data privacy provisions in Germany is that any form of data acquisition and processing is interdicted, if not subject to explicit authorization within the same act (or some regulation by special law). [BfD-INFO 1, 2002]

Consent of test persons

Based on the provision that data acquisition and processing is generally interdicted, the most important exception from this rule for FOTs is the consent of the person concerned (see Sec. 4 and 4a BDSG, Art. 7 and 10 EU Directive 95/46/EG). For any consent given in terms of data acquisition, processing or use, the person concerned must:

- make this statement in written form (if not certain circumstances make an other form necessary)
- the consequences must be clarified (intended purpose of acquisition/ processing/ use), including the consequences, if consent is not given
- the consent must even be specially highlighted, if the consent to data acquisition/ processing/ use is issued together with other statements (consent to the use, etc. of data concerning health – which might be of interest for FOTs – will call for a special (separate) consent in this respect)
- consent must always be given voluntarily!

[BfD-INFO 1, 2002]

Principle of purpose limitation

Another important principle is purpose limitation (Sec. 14, 28, 29 BDSG; Art. 6 EU Directive 95/46/EG). This means that data may only be processed for the same purpose it has been collected for (or saved, in case prior acquisition is non-applicable).

However, the act comprises a number of exceptions to this rule. As far as relevant for a FOT, acquisition, processing and use of data is not limited to the same purpose in certain cases (see Sec. 14 BDSG valid for public administration and Sec. 28 BDSG for private bodies). [BfD-INFO 1, 2002] For details on these exceptions see Deliverable D6.3.

Data acquisition (extent and limitations)

The general principle for data acquisition is that data must be collected frankly from the person concerned (and not otherwise). As far as data acquisition is taken out by a public body, this is only possible to such an extent as necessary to fulfil the legitimate tasks (Sec. 4, 13, 28, 29 BDSG; Art. 5-7 EU Directive 95/46/EG). The relevant limitations for private entities in case of a FOT are in so far

- specified in the contract on which data acquisition is based
- restricted to explicit consent in case of intimate and very private data (such as racial and ethnical background, political opinion, religious and philosophical belief, health and sex life). Especially data on the health of a test person might be important for the FOT. The acquisition will most likely prove permissible again for the reason of research: Here data acquisition must meet the legal principle of proportionality (and therefore go conform to an evaluation of the higher and therefore more valuable legal right) (see Art. 8 EU Directive 95/46/EG).

Apart from this, data acquisition is also limited by the principle of data economy (Sec. 3a BDSG). This is to say that no more personal data shall be collected and saved than is really necessary to fulfil the purpose in question (i.e. any unnecessary data compilation shall be avoided). [BfD-INFO 1, 2002]

Technical and organisational measures

An important and rather costly aspect of data privacy is the technical and organisational standard that must be applied. Generally speaking, those measures are necessary that will guarantee the compliance with data privacy (see Art. 16, 17 EU Directive 95/46/EG; Sec. 9 BDSG; Sec. 10 BDSG calls for further technical measures but concerns the automatically generated release order and should not be applicable to a FOT). What this rather general description can imply, has been stipulated in an annexe to the German data privacy act. The effort needed to ensure data privacy in case of automatic processing and usage is dependent on the character: Intimate data is most strictly protected and forms the core area that may not be impaired at all (will seldom be relevant in case of FOTs), personal, private data is strongly protected and data with a relation to other people that is generally known, is the least critical. The character of the data in question indicates the effort to be applied in order to achieve reasonable care.

General measures to ensure data privacy on the technical and organisational level are described in detail in the Deliverable 6.3.

A.4.4 Data privacy in research activities

Research is in itself a basic right of constitutional weight – as is data privacy, see above. Therefore, an appreciation of both values must be carried out for the case in question. Research is facilitated within data protection regulations such as Art. 6, 11 and 13 Directive 95/46/EG and sec. 40 BDSG.

The regulation in sec. 40 BDSG emphasises the principle of purpose limitation to the object of research and explicitly calls for an anonymisation of data at the earliest possible stage. Until anonymisation can be achieved, the characteristics of the person concerned must be saved separate from the particulars on personal or factual relations and must only be brought together in case this is required by the object of research. Any publication of personal data can only be admissible in case the person concerned gives his/ her consent (if relevant for FOTs in the first place).

However (for Germany), no right of professional discretion has been stipulated in the field of research activities (as existing e.g. concerning confidential medical communication of a medical practitioner). This has important implications within criminal law, see below section A.5.6. This may be completely different in other countries of the EU as the Directive 95/46/EG gives sufficient leeway for a deviant regulation. [ROSSNAGEL, 2003]

In order to achieve data privacy for the test persons in spite of these regulations, it has been suggested [ROSSNAGEL, 2003] to deposit data necessary for re-identification (after having pseudonymised the data) with a bearer of secrets (such as lawyers). Such a bearer of secrets can refuse to release data he/ she is entrusted with. It has further been proposed to store the personal data with the (test) person concerned [ROSSNAGEL, 2003]. This implies the risk, not to obtain the data, because the test person might finally decide on wanting not to disclose the personal data at all. However, for a FOT this is an option to be considered: As far as the data recorded is stored within the car e.g. by means of a SD-card etc. this would allow the test person to remove the personal data and take care of it by himself/ herself (until it is handed over to the organisation taking out the research activity). The personal data is thus fully placed at the disposal of the test person up to its voluntary release. The test person is, therefore, free to decide upon the further use or existence of the data (and has every right to destroy the data if he/ she pleases). This will avoid conflicts arising from data acquisition and balance the risks implied to a great extent. From a practical point of view it should technically be ensured that the storage medium is easily accessible and removable (and it should not be too expensive either, because the test-person would supposedly be required to come up for a replacement in case of the storage medium’s destruction).

A.4.5 Video recording

Video surveillance might be of great importance for FOTs: Only knowledge from surveillance of the surrounding traffic as well as the driver can determine the impact of a system and/or reveal why a sudden driver action is necessary, etc.

The conflict arises from the high quality of video data that can be achieved by today's technical possibilities in video recording. Additional processing of this data will therefore often reveal the personal identity. Privacy measures and regulations on data privacy in these cases are existent, but they are difficult to handle, see below. The problem with video data is that it is usually impossible to anonymise or pseudonymise images or even videos. This is why this data is especially 'delicate' in terms of data privacy and comprises great dangers. As such must be considered the threat that video-data might be made publicly accessible over the internet (which would be illegal, if the video contains personal data and the test person has not granted his/ her consent prior to this release). Such a video, however, would only then be of great interest, in case the driver or a third party behaves in a way that calls for voyeurism. As the harm to data privacy of the person concerned (and thus his/ her basic right of informational self-determination) can be tremendous, it must be considered reasonable to delete such a sequence entirely as soon as possible (which is as soon as respective knowledge on the existence of the sequence is available). Otherwise, measures must be taken to secure the data adequately (which would be challenging).

From a legal point of view, it must be differentiated between data acquisition outside the car (surrounding traffic) and data acquisition inside the car. This should apply to the technical requirements as well.

Generally speaking, it is advisable, according to the basic principles described above, to do without any video surveillance (in case this proves possible). And pure surveillance (without any recording) is always the less inculpatory measure. Therefore data acquisition by means of video recording should only be employed if indispensable. However, this will usually be the case with FOTs so the following statements will concentrate on necessary video recordings. [ROSSNAGEL, 2003], BfD-INFO 1, 2002]

Video recording of third parties

Video surveillance by means of optical-electronic devices (video surveillance) is only permissible according to the guidelines provided in Sec. 6b BDSG and is also considered critical within the Directive 95/46/EG (see Art. 33 Directive 95/46/EG that contains a revision clause to enhance data privacy in case of video and image surveillance). As far as Sec. 6b BDSG is concerned, it has been criticised that a differentiation of private (video-) data acquisition and video surveillance by public bodies is not made. The scope of sec. 6b BDSG therefore comprises public as well as private bodies and according to its wording it is only applicable to the "surveillance" of publicly accessible places (which would apply to roads). This, however, does not lead to the conclusion that other recordings are irrelevant in terms of data privacy (this is only the case with really private recordings such as those taken within a family). Any (other) video recordings must therefore meet the provisions stipulated in sec. 6b BDSG, in order to achieve compliance in case of video-data acquisition by non-public institutions. Data acquisition must therefore comply with the following requirements:

Data acquisition must be taken out frankly. A hidden camera will therefore not be regarded permissible as long as consent of the persons concerned has not been obtained.

A legitimate interest for video-data acquisition must exist (see sec. 6 para. 1, No. 3 BDSG). In case of a FOT the concrete legitimate interest will be the same for which video data must be recorded in the first place (and can not be done without). In case of research (a value of constitutional weight too, see above and Art. 5 para. 3 Grundgesetz = German constitution) the concrete legitimate interest is implicated by the motivation (research).

The video data must be necessary in order to achieve the purpose identified as the legitimate interest. Here it must be considered that the storage of data must also be indispensable, as any storage of video data is considered ultima ratio. This should be overcome in case of a FOT, as long as there is sufficient need of video data.

The video data must be deleted as soon as it is no longer necessary in order to achieve the legitimate purpose of research, see sec. 6b para. 5 BDSG. The same will apply in case of a protection-worthy interest of a third party which is in conflict with further data storage.

[ROSSNAGEL, 2003], [BfD-INFO 1, 2002]

Video recording of the driver

The situation for the driver differs strongly. Basically, however, the same requirements and regulations are valid in this case. The main difference is that the driver will always have to give his/ her consent to the video recording as the permanent recording within the car is strongly invasive. By no means may the recording be taken out with a hidden camera and without informing the driver beforehand. (see ROSSNAGEL, 2003]

Further care must be taken not to record any video data of other passengers, if this can be avoided by technical means (as usually a legitimate interest in so far will not exist, and, as mentioned above, the principle of data economy must be applied). In case the video recording of other passengers is inevitable, it must be ensured that the camera is well on view and thus obvious that data is being recorded. The designated test persons must in this case further be sensitised to inform any passenger of the recording. The same will, of course, apply to any further drivers of the car. It may, however, for a number of reasons be a good idea to restrict the use of the participating vehicles to the actual test person or provide for a “switch off” of all systems and data logging (for safety and data privacy reasons) in case other drivers use the vehicle. This will not only provide for consistent data (which should be necessary from a scientific point of view) but at once ensure that all the contractual agreements actually reach the driver.

A.4.6 Implications of criminal law

If the FOT goes according to plan, criminal law will not be affected. However, in case of accidents, the data collected might be used otherwise. In this case, personal data of the test person will have been recorded to an extent that is generally not available in this situation in the first place. As the breach of certain traffic rules may be relevant under criminal law aspects, the data will be of interest for means of criminal prosecution as well. This, at least in case the availability of the data is known; however, due to possibly obvious modifications within the vehicles (perhaps even a camera is on view for video-recordings) further inquiries in this respect seem probable. It must therefore be expected that the data might even be subpoenaed on application of a public prosecutor by a judge. For Germany this possibility is given, see sec. 94 seqq. StPO (Strafprozessordnung = Code of criminal procedure). In this context it must be pointed out that these legal effects will be tolerated in Germany and the recording for research-reasons will not privilege the test-person (i.e. it would not be barred to confiscate the data for the reason of criminal prosecution). And data privacy provisions (for Germany) will not bar the use of this data for the reason of criminal prosecution either.

In case the data is already in hold of the organisation doing the research, it would have to be released anyhow (see sec. 95 StPO) – in spite of the fact that this might mean a moral dilemma for the researcher involved in the FOT. These effects, however, may be largely avoided if the procedures suggested for the means of 'data privacy in case of research activities' (see section A.5.4) are taken into effect.

In this context, it must be pointed out that a suspected person always has the right to remain silent in order to avoid self-incrimination (i.e. the accused is not obliged to cooperate actively in the own conviction: "nemo tenetur se ipsum accusare"). It will therefore not be considered a criminal offense in itself, in case the accused would delete or destroy the data recorded. As far as a civil court, however, will decide on compensation for loss suffered, conclusions can be drawn from the fact that the data has been deleted/ destroyed.

A.5 Insurance

A.5.1 Introduction

This section exemplarily goes into the legal situation for Germany as far as road traffic liability and the associated insurance issues are concerned. This should allow for sufficient insight in this aspect to sensitise for possible arrangements and precautions to be taken as far as the insurance of the test-vehicles is concerned.

A.5.2 Road traffic liabilities in Germany

According to national road traffic liability law in Germany, accident victims can potentially claim for compensation from the "keeper" of a vehicle ("Fahrzeughalter", see below), the driver and the vehicle's insurance.

The "keeper" of the vehicle will usually at once be the legal owner; however, this is not invariably true. From a legal point of view, the "keeper" is generally defined as the person that makes use of the vehicle at own expense, i.e. comes up for the costs and has the capitalised use [HENTSCHEL, 2007]. The "keeper" will be liable for any damage to the legally protected interest resulting from the operational hazard. The only (basic) requirements for a claim in so far are that the vehicle was in use at the time the damage occurred (this, however, can even be assumed when parking in public space) and that the vehicle's use (i.e. its operational hazard) has lead to a damage of the legally protected interest. In case a further vehicle plays a part in the emergence of the damage to the legally protected interest, its respective contribution will be considered too. The same applies to contributory negligence of the damaged person. In so far this applies to the causation giving rise to the damage of the legally protected interest. In a further step the remoteness of further damages incurred that can be traced back to the damaging event are considered too.

Unlike the "keeper", the driver will only be liable in case of fault (e.g. any driving mistake, etc. that leads to a damage of the protected interest). Apart from this, the driver is, generally speaking, liable for the same damages to legally protected interests as the "keeper". If the driver's and "keeper's" liability is given, they will both be jointly and separately liable (together with the insurance, see below) for the damage (a term that describes that the damaged person can decide freely which debtor to claim against for the whole damage – which is then settled between the two or more debtors).

Of course, the damage the "keeper" as well as the driver are liable for, is insured via the same compulsory car insurance. By provisions of law the "keeper" is obliged to contract such an insurance in case he wishes to operate his vehicle on public roads. It is regulated that the contract covers the damage on account of the driver as well as the compensation for damages imposed on the "keeper" (and as long as the contractual obligations are adhered to, no recourse will be taken). In Germany, a direct claim of the aggrieved party against the insurance is admissible according to provisions of law. [ALBRECHT, 2005]

A.5.3 Insurances for road traffic in Germany

In so far as material damages are concerned, it is – according to the situation in Germany — important to distinguish between many different types of insurances. First of all, the compulsory road traffic insurance will cover the damage to the property or health of a third party (automobile third party insurance – as stated above, this insurance is compulsory in Germany and therefore widespread and generally referred to as the "car insurance"). It will, however, neither cover the physical damage to the "own" vehicle nor the damage to the health of the driver and other occupants. As far as the physical damage to the car is concerned, a special insurance can be obtained to cover this (comprehensive insurance/ comprehensive coverage insurance including collision). As far as the health of occupants or other passengers is concerned, a special motor passenger personal accident insurance type exists that will come up for damages to passengers. However, it must be kept in mind that the insurance sum for this insurance is usually restricted (and will generally not be sufficient to adequately compensate for serious injuries, special medical care requirements, etc.). Furthermore, the driver might be excluded in this insurance – here a special insurance may prove necessary (driver personal accident insurance). This, however, is again usually limited to certain insurance sums that may not prove to be sufficient for full coverage. Therefore it may appear to be reasonable in some cases – according to the field test design chosen – to obtain some kind of special insurance tailored to the special needs of the specific field trial (e.g. clinical trials insurance).

Most important in all cases will be to disclose the fact towards the insurance that the vehicle is participating in a FOT (which in general should simply be accepted by the insurance). Insurance rates might rise, however, depending on the systems integrated in the vehicle (subject to the FOT, likely in case of premature systems which might involve additional risks). Disclosing this information and possibly incorporating a respective clause in the contract will be a reasonable method to avoid legal uncertainties as far as insurance coverage in case of an accident is concerned.

Automobile Third Party Insurance

As stated above, this insurance is compulsory by law. The minimum insurance sum for this insurance type is e.g. in Germany fixed for motor vehicles at 2.5 Million Euro in case of damages to health (in case of fatal injuries or more than three persons injured: 7.5 Million

Euro) and in case of damage to property even limited to 500 000 Euro (see annexe 1 to Sec. 4 of the obligatory insurance law = "Pflichtversicherungsgesetz"). These, however, do not necessarily cover the whole damage in case of serious accidents and even the maximum compensation sums according to the German road traffic act (which is below the minimum insurance sum) can be exceeded in case the claim is based on the law of torts (in this case a maximum compensation sum no longer exists).

It is therefore reasonable to raise the test persons' awareness to these (general) limitations (see section A.3.3.) and, if considered necessary, the test vehicle should be insured to better conditions.

Comprehensive insurance/ comprehensive coverage insurance including collision

It must further be decided on the necessity of a comprehensive coverage insurance. This insurance will replace the material damage to a vehicle even in case of self-inflicted accidents (depending on the contract). The insurance will usually exclude intentional damages and may exclude damages resulting from gross negligence. This insurance is not compulsory. If it is renounced, it must beforehand be decided and agreed upon who will come up for these material damages (this insurance would cover) in order to provide for legal certainty. Possibly this has influence on the information that should be provided to the test person, see section A.3.

Motor passenger personal accident insurance

In case of damage to the passengers, compensation can be obtained within the automobile third party insurance of the injuring party. In case of hit and run accidents, absence of insurance coverage of the third party, etc. compensation can be claimed from the personal accident insurance. It will cover all those damages to health involved from the time the passenger gets in the car until he gets out again [HIMMELREICH/HALM, 2006].

This personal motor passenger accident insurance will generally cover costs for medical treatment as well as the costs involved in the event of motor passengers' death.

This insurance shall not be enlarged on within this report. However, selected aspects shall be pointed out: This insurance will not necessarily cover the damage to the driver (which may, however, be the case). And insurance sums vary strongly. They may not be sufficient to cover severe health injuries or the costs involved in case of disability. Special attention must therefore be paid towards the maximum sums. Further information on this type of insurance should be gathered in the planning of the FOT (if applicable and considered necessary within the test design chosen). [HIMMELREICH/HALM, 2006]

Driver Supplementary Insurance

The Driver Supplementary insurance (= "Fahrerzusatzversicherung") is a fairly new insurance type and is largely unknown. The conditions of insurance may differ strongly. Motor passengers are in Germany (since 1st Aug. 2002) covered by the Automobile Third Party Insurance. This is even the case, if it comes to a self-inflicted accident caused solely by the driver of the vehicle they are occupying. In this case, the passengers can claim for compensation against the vehicle's Third Party Insurance. Therefore today only very particular cases depicted above will leave passengers without insurance coverage.

The driver, however, is the only car occupant who may not be able to claim for damages (or only have a partial claim against the third party's insurance). This Driver Supplementary Insurance will cover these damages as far as compensation cannot be obtained otherwise. This insurance type is generally considered reasonable. [HIMMELREICH/HALM, 2006]

Clinical Trials Insurance

As stated just above in section A.6.3.4, the risk of damage to the health of the driver (who is at once the test person) is severe and must be considered beforehand. Of course, the testing in open traffic will also involve many further risks to third parties which must nonetheless be considered. Therefore the justifiable risk will be limited strongly in the first place.

Shall a greater risk nonetheless be taken (and this be considered otherwise permissible), a clinical trials insurance may be necessary to cover the risks involved. As far as (medical) clinical trials are concerned, this insurance type is common. For the purpose of road traffic such an insurance would have to be tailored according to the specific needs of field operational testing.

Test Equipment Insurance

The data-logging equipment and possibly prototype systems may have to be insured too by means of property insurance. This should be kept in mind when planning a Field Operational Test. This electronic equipment will usually not be covered by the comprehensive insurance/ comprehensive coverage insurance including collision (as maximum insurance sums for common electronic equipment in vehicles will assumedly be exceeded by far).

A.5.4 Insurance issues in case of non-overrideable systems

A basic principle of European road traffic is driver's full control. Therefore non-overrideable systems that influence the driving task, comprise unsolved and complex legal questions that cannot be fully assessed at present (see "Communication from the Government of the Federal Republic of Germany to the European Commission of 27 June 2007" page 6, stipulating conclusions of the eSafety Conference in Berlin on 5th/6th June 2007].

As far as insurance issues are concerned, it must therefore be taken into consideration that any type of insurance known today assumes full driver's control. In case the Field Operational Test is otherwise regarded permissible (see sections A.7 and A.8), special arrangements for specific insurance coverage must be taken into consideration.

A.5.5 Insurance issues in case of cooperative systems

Cooperative systems may comprise very specific insurance issues in case the influence on the driving task is strong: If the cooperative aspect therefore involves any kind of vehicle control, it must further be taken into consideration that all present regulations on road traffic are based on the assumption of a vehicles (and driver's) autonomy. In case control of a vehicle is therefore dependent on other vehicles (the same for road side beacons) it must be assessed whether common liabilities (and thus insurance contracts) will sufficiently cover all possible damages in between the linked ("cooperative") vehicles and towards surrounding traffic. If this is not longer the case, it might turn out to be a legally challenging task to tailor the insurance contract to fit the actual "cooperative" situation.

A.6 Vehicle licensing requirements

A.6.1 Licensing requirements for motor vehicles in general

As long as the Field Operational Test is focussed on the evaluation of applications already approved of as optional or standard fitting of the test vehicles, no vehicle licensing requirements will be in question.

The licensing of a vehicle for Europe is generally taken out by means of type approval according to technical rules and regulations in international law. For Europe the so called ECE-Regulations are binding (and their fulfilment will be considered sufficient for road admission throughout Europe). However, gaining a type approval certificate for a new vehicle type is challenging and costly and can hardly be considered an appropriate approach for field testing. Apart from this, it must be taken into consideration that in practice an approved vehicle will serve as a basis for further system integration in a FOT.

In this case the approved vehicle is modified for the purpose of field testing. These modifications may – much depending on the character of the modifications – lead to the cancellation of the vehicle’s operating licence. Whether this is the case strongly depends on national licensing requirements. These may still be in existence (as is the StVZO = “Straßenverkehrszulassungsordnung” in Germany). According to the provisions therein, the operating license will expire in case of certain modifications (see Sec. 19 para. 2 StVZO).

This does not apply, in case the vehicle parts integrated have a general approval of their own (which is further specified) or have already been approved of as they are – fitted to the vehicle (and have thus been included in the operational licence of the respective vehicle). In order to make the legal effects of modifications manageable, the German Federal Ministry of Transport has established a catalogue of possible modifications and their impact on the vehicle’s operating licence (this catalogue is not legally binding). [HENTSCHEL, 2007] The catalogue will provide a good overview in terms of challenges to be overcome for field testing according to the modifications envisaged. Generally speaking, minor changes such as the safe integration of a display, a separate power supply or data logging equipment will not lead to the cancellation of the vehicle’s operating licence. The modifications must, however, be made transparent.

A.6.2 Special regulation for vehicle manufacturers

In this context an important regulation shall be pointed out that will partly exempt vehicle manufacturers (in hold of the type approval certificate for the respective vehicle) from special licensing requirements (see sec. 19 para. 6 StVZO). In Germany, a vehicle that is used for testing by the manufacturer and registered as such, will not be deprived its operating licence, if further parts are integrated for the purpose of testing. [HENTSCHEL, 2007] This regulation will, however, not permit the modification of vehicles privately owned and registered.

A.6.3 Licensing requirements of “premature” systems/ applications in general

For the purpose of this report a “premature system” shall be considered a system that has so far not been approved of within vehicle type approval and is not separately approved as car accessory either. In order to evaluate such a “premature system” in a Field Operational Test, a special approval might be required to maintain the vehicles operating licence.

For Germany the law within the federal state is decisive as far as the responsibility of the local public authority is concerned (see Sec. 68 StVZO). The responsible public authority will then decide on the necessity of a report by an officially recognised expert certifying consistency with legal provisions. This will usually not be necessary for manufacturers, see above, section A.7.2.

A.7 Special licences (exceptional licences within road traffic law)

A.7.1 Introduction

Further exceptional licences should normally not be necessary for a Field Operational Test – apart from those discussed above in case of modifications on the vehicle’s side. This finding will also apply to the drivers’ driving licences: Driving licences correspond to certain vehicle types and their use will therefore cover any driver assistance as well as any driver information system that does not put full driver’s control into question. And whether data-logging equipment is implemented in a vehicle or not will – if at all – influence the operational licence of the vehicle and does not call for any further special licence.

However, further attention must be raised towards those systems that may intervene beyond full driver’s control. In this case, exceptional licences may be necessary after all.

A.7.2 Full control of the human driver

The technology available in the past, as well as the Vienna Convention on Road Traffic (1968) have likewise lead to the assumption of the driver’s responsibility to ensure full control over his vehicle. Art. 8 para. 5 and Art. 13 para. 1 of the Vienna Convention explicitly stipulate this full control of the driver over his vehicle “under all circumstances” (Art. 13 para. 1 Vienna Convention on Road Traffic).

The Vienna Convention on Road Traffic formulates a minimum set of requirements in purpose of free (and safe) flow of cross-border transport between the signatory states. The document has had strong influence on the development of national Road Traffic codes and the all-underlying idea of full control of a human driver has thus found its way into many legal provisions concerning road traffic in Germany as well as other countries throughout the EU (and worldwide). In fact, the number of legal provisions based on this idea of full driver’s control went without saying and can even be traced back to national road traffic liabilities (which will again influence insurance issues of such systems, see above, section A.6.4).

These findings are common for the EU at large and must be taken into consideration, in case a system shall be evaluated in a Field Operational Test that overrules full control of the driver. In this case, special legal advice as to the consequences the specific system might bring about will be necessary, as will be the application for exceptional licences. These restrictions will, however, not affect systems that do not put the full control of the driver into question. As such must be considered systems that optimise driver initiated functions (e.g. ABS), advisory systems (e.g. speed alert) and fully overrideable ADAS (e.g. adaptive cruise control). Permissible must further be considered those non-overrideable ADAS that have the same effect as traditional technical limits in vehicle performance (e.g. speed limiters) or intervene in situations that cannot be handled by the driver in time and ensure that the intervention keeps in line with the drivers' intentions and will (e.g. ESP and automatic emergency braking). [ALBRECHT, 2005]

Interventions into driving, however, that counteract to the intentions and will of a driver still able to perform the driving task would bring about legal consequences that cannot be predicted at present. (see "Communication from the Government of the Federal Republic of Germany to the European Commission of 27 June 2007" page 6, stipulating conclusions of the eSafety Conference in Berlin on 5th/6th June 2007]

Therefore a need for an exceptional license will arise whenever a non-overrideable system that does not ensure full control of the driver shall be subject to a Field Operational Test.

A.8 Ethical rules

Ethics can be considered as a sub-discipline of philosophy and moral principles guiding behaviour, i.e. they help to distinguish, if a certain conduct is right or wrong. Ethical rules apply and must be obeyed in all kinds of research activities on living organisms and, of course, in particular with human beings. The currently most important ethical rules relevant for research but also professional work on human beings have recently been reviewed as subject of the NoE HUMANIST TF 2 activity on ethical laws and guidelines that apply to behavioural experimental studies [HANZLIKOVA, 2004]. However, in the context of the present report and the planning and preparation of FOTs it does not seem to be necessary to review and discuss all principles which apply when e.g. performing medical or genetic research. Here it seems to be sufficient to refer to the key principles for the evaluation of research which are according to [HANZLIKOVA, 2004, p.5]:

*"- **Respect for the personality** and his or her autonomy, dignity and self-determination*

***Beneficience:** a commitment to maximise potential benefit and minimise possible risks"*

Regarding the planning and performance of FOTs as research projects in the 7th Framework Programme the European Commission makes a clear point when stating: *"All research activities carried out under the seventh Framework Programme must be carried out in compliance with fundamental ethical principles"* (Decision no. 1982/2006/EC; see http://ec.europa.eu/research/science-society). For this reason research proposals shall be evaluated by an independent panel of experts if ethical aspects have been properly addressed. For practical purposes of writing a proposal the EC provides a checklist with critical questions which is designed to help proposers to identify possible relevant ethical issues (see http://ec.europa.eu/research/science-society). With regard to FOTs the following two questions on "Privacy" seem to be most relevant:

Does the proposal involve processing of genetic or personal data?

Does the proposal involve tracking the location or observation of people?

In case of "yes" proposers are required to describe at least the procedures for obtaining informed consent from the persons and the procedures for protecting confidentiality.

Moreover, the process of anonymisation or encoding of the data shall be described and it has to be indicated, if the data are used for commercial purposes. These aspects will naturally correspond to legal data privacy provisions, see section A.5.

A.9 Conclusions

From an overall perspective on legal and ethical issues for FOTs, a number of aspects must be taken into account in terms of planning and accomplishing of such testing. This is mainly due to the significant difference between normal driving and FOTs which lies in the evaluation of possibly immature systems (as far as legally permissible), the great extent of data logged and the unique and possibly unprecedented situation test drivers will be confronted with in open traffic.

To sum up, prohibitive difficulties neither from a legal nor from an ethical point of view are in so far to be expected. As long as the advice provided in this report is considered, potential risks – as far as presently foreseeable – can either be settled, avoided or safely handled. It must, however, further be taken into account that advice provided herein is void of knowledge on the concrete system design and thus specific dangers that might be implied in case of particular systems cannot be covered. It can be expected that such concrete difficulties – apart from those indicated as delicate in the report – can be overcome, this, however, will call for further support on legal and ethical issues within the concrete FOT.

Annex B FOT Implementation Plan (FOTIP)

(To be read in conjunction with Chapter 2)

FOT Teams and People

1. Research Institute contracted to run FOT
2. Project Manager
3. Research Team
4. Technical Support Team
5. Administrative Support Team
6. Project Steering Committee
7. Project Management Team
8. Accounting/Auditing Advisor
9. Legal and Ethical Advisors
10. Sub-Contractors
11. Public Relations and Communications advisor
12. Project Sponsor(s)

Legend for FOT teams and people identified in column 3 of the FOTIP

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity *(See above for Legend to teams and people)*	Critical Considerations (the "dos" and "don'ts") *(Italics denote most critical issues)*	General Advice *(Italics denote most critical issues)*
1. **Convene FOT teams and people**	• Appoint FOT project manager • Appoint research team • Appoint technical support team • Appoint administrative support team • Appoint team leaders in each of the research, technical and administrative teams • Appoint project steering committee • Appoint project management team (for –day-to-day management) • Appoint accounting/auditing advisor • Appoint a legal and ethics advisor. • Appoint sub-contractors • Appoint a public relations/communications advisor • Sign off on agreed research and support structure.	• 1 • 2 • 2, 6 • 2 • 2 • 2, 6, 7, 11, 12 • 2 • 2, 7 • 2, 7 • 2, 7 • 2, 7 • 2, 7, 5, 8, 12	While the project manager must have knowledge of all activities, ensure that critical knowledge is not vested in just one person. Personnel, including the project manager, may leave the project. *Ensure that there is a "standby" for all key research and management roles within the FOT.* Appoint early someone to deal with human participants/ethics committee issues. Include in the research team someone who is a "gizmo" expert – who has up to date knowledge about current ICT/ITS developments and capabilities. Civil engineering and geographical information system (GIS) expertise is also critical. Ensure the project management team meets regularly (about once a month) to resolve research issues, monitor timelines and budgets, and resolve administrative, technical and other issues. *Choose contractors that can guarantee in writing that, if a staff member leaves or is ill, there is sufficient expertise and capacity to maintain project continuity.* Maintain good relations with other partners involved in the FOT. *Ensure that the FOT evaluation process will be, and be recognised as, independent.* It is not necessary to appoint all teams/people at the same time – appointments should coincide with project needs.	Although this Activity precedes Activity 2, the choice of teams and people will be determined to some extent by the aims and objectives of the FOT. Appoint a project manager with excellent research, project management and communication skills. (Note. In some FOTs, the FOT project manager is responsible for both the administrative and scientific management of the FOT. In other FOTs, a senior researcher may be responsible for the scientific, but not the administrative, management of the FOT. This requirement will depend on the scale of the FOT.) The research team should be multi-disciplinary and would typically include psychologists, civil, mechanical, electrical and electronics engineers, statisticians, human factors experts, traffic safety experts,

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
			Identify a final internal arbiter, acceptable to all parties, who can resolve scientific, administrative, legal and other disputes. Decide early in the project the frequency and timing of project Steering Committee meetings	and socio-economic modelling experts. The technical support team would normally include computer software engineers, communications engineers, mechanical, traffic, civil and electronic engineers, and GIS experts. The project Steering Committee sets the strategic direction of the project and keeps it aligned with the project aims and objectives. Normally it would include the FOT project manager, selected members of the research and project management teams (e.g. the team leaders), along with key stakeholders and the sponsor(s). Members should have authority to commit their organizations to the aims, objectives and implementation of the FOT. For smaller FOT projects, the stakeholder committee may not be necessary. The project management team is led by the FOT project manager and includes selected members of the

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
				research (e. g, the team leaders), technical and administrative teams. A legal advisor should support the FOT over the full duration of the project (a lawyer's office providing advice whenever needed is sufficient). Legal knowledge must be available on the legal situation in the country, or countries, in which the FOT is conducted.
2. **Define aims, objectives, research questions and hypotheses**	• Define aims and objectives of FOT, in conjunction with relevant stakeholders • Identify systems and functions to be tested • Identify use cases/ situations in which systems and functions are to be tested • Define research questions and prioritise them • Formulate hypotheses to be tested, deriving from research questions • Determine and resolve constraints which may prevent the aims and objectives from being met • Define final aims and objectives of the FOT, and seek agreement from relevant stakeholders.	• 2, 3, 4, 6 • 2, 3, 4, 6 • 2, 3, 4, 6 • 2, 3, 4, 6 • 2, 3, 4, 6 • 2, 3, 4, 6, 7 • 2, 3, 4, 6	Be prepared for the potential for FOT aims and objectives to change when new administrations come in. Be prepared for the potential for conflict in objectives by different stakeholders. e.g. a car manufacturer wants a deep understanding of product use and driver behaviour and acceptance versus a public authority more interested in determining the impact of system use on traffic and on the transport system.	See Chapters 4, 5 and 6 of the FESTA Handbook for further advice on defining the aims, objectives, research questions and hypotheses for a FOT. Constraints which may prevent the aims and objectives from being met might include cost, lack of supporting infrastructure, time, willingness and commitment of key stakeholders to cooperate in providing supporting infrastructure, their likely support in promoting the aims and objectives of the FOT, the availability of appropriate data, etc

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	• Sign off on aims and objectives of FOT	• 2, 3, 4, 6, 7, 11, 12		Commonly cited aims are: - evaluate system(s) effectiveness in changing behaviour and performance - evaluate driver acceptance of system(s), including willingness to purchase -evaluate system technical operation - stimulate societal demand for new technologies - evaluate safety impacts - evaluate environmental impacts - evaluate impacts on traffic (e.g. congestion, mobility) - evaluate socio-economic cost-benefits - evaluate commercial impacts (e.g. productivity, return on investment, direct cost savings, incremental revenues by getting more customers, customer loyalty, etc.) *Defining the research questions and prioritizing them at an early stage will ensure they stay at the focus of the FOT and help protect from subsequent "mission*

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
				creep"
3. Develop FOT project management plan	• Define project activities, tasks and sub-tasks • Decide who is accountable for completion of activities, tasks and sub-tasks • Determine timelines for completion of activities, tasks and sub-tasks • Determine budget for project activities, tasks and timelines • Develop a project GANTT chart to guide project management • Implement procedures for monitoring project activities, timelines, budgets and resources (e.g. project management team meetings) • Undertake a risk assessment for the FOT and plan contingencies as required. • Determine sign off procedures (meetings and documents) to ensure that there is sign off on all critical decisions and stages in the FOT by all relevant parties • Agree on project issues which are confidential and implement mechanisms for safeguarding their confidentiality. • Develop a manual for conducting the FOT that documents critical	• 2, 7 • 2, 7 • 2, 7 • 2, 7 • 2, 7 • 2, 7 • 2, 7 and risk management consultant • 2, 7 • 2, 7, 12 • 2, 3, 4	*Include in the total budget some "contingency" that can be used to pay for unforeseen activities and tasks (especially meetings) that cannot be anticipated.* 5 -10 percent of the total project cost is recommended. Different elements of the project may require different proportions of this contingency. It should be held and allocated by the project manager, not sub-activity leaders or partners. Identify and document in the GANTT chart the dependencies that exist between different activities, tasks and sub-tasks. Anticipate the need and budget for specialist consultants with skills and expertise that does not exist within the project team (e.g. training experts, software developers, lawyers etc) Anticipate changes to 3rd party vehicle fleets (e.g. vehicle upgrades and changes in operating routes) during the course of the FOT. Be aware that technical efforts are most likely to incur risk in terms of time and budget (especially the hardening up/refinement of systems, where these are developed within the FOT) *Don't under-estimate the time required and the cost of designing, running, analysing and de-commissioning the FOT. It will be greater than you think.* Assume that some further modifications to, and fine tuning of, the project management plan will be required. It is impossible to foresee everything that is required in running a FOT. Develop procedural manuals for those conducting the	*Documentation of all project meetings is critical to record critical decisions, document the lessons learnt and justify possible blowouts in budgets and timelines.* A budgeting structure that accommodates the uncertainties associated with running FOTs is desirable – for example, a series of prospective budgets for each critical stage of the FOT. *Be aware that in some jurisdictions project papers from publicly funded projects are public documents and copies can be requested by members of the public.*

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the “dos” and “don’ts”)	General Advice
	procedural knowledge. • Sign off on project management plan.	• 2, 7, 12	FOT to ensure that, if project staff leave, all procedural knowledge does not leave with them. These should be developed for each activity.	
4. **Implement procedures and protocols for communicating with stakeholders**	• Commission communications advisor to design communications plan • Develop and implement communications plan • Appoint media spokespeople • Sign off on agreed communication protocols.	• 2, 7 • 2, 7, 11 • 2, 7, 6 • 2, 7, 11, 12	*The media can be an important ally in supporting and promoting the value and outcomes of FOTs. Assume, however, that you will be mis-represented at times by the media. Try and limit media attention until data collection is complete, and at least after participant recruitment is complete.* *Agree in the contract with the sponsor who is responsible for press releases and dissemination of information and results.* *FOTs attract a lot of unsolicited media attention. Provide adequate time and budget for unsolicited communication with stakeholders, especially with the media.* Ensure that the project steering committee has input to the communications plan. *Ensure that there is appropriate control of communication with the media, through the appointed media spokesperson. For EU projects, involving multiple partners, it may be necessary to appoint more than one media spokesperson.* Everyone involved in the project must know who the media spokespeople are. *The media spokesperson should consult with the project management group before speaking to the media, especially on sensitive issues.* Provide media training for appointed spokespeople. *Build political support for the FOT early in the project,*	*Open communication with key stakeholders is important at an early stage of the FOT to ensure that the aims and objectives of the FOT are clear, that stakeholders are committed to the project, and that the aims and objectives of the FOT are not misquoted, misrepresented or misunderstood.* There should be an agreed minimum level of transparency and result sharing in the FOT — avoid “confidential FOTs”. It may be beneficial to engage a professional press office to handle external communications, particularly with the media. FOT drivers and FOT researchers are usually of most interest to the media. Decide in advance with stakeholders a minimum time for approval for statements

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
			and maintain it during and after the FOT. *Be aware that there may be some key stakeholders who believe that FOTs are an impediment to system rollout. These people, in particular, must be made aware of the rationale for FOTs.* Plan to have some results available at early stages of the project. If desirable, they should be released to an informed audience (e.g. at a conference) but not to the media as they could contaminate subsequent data collection. Plan for annual public meetings, and a project website, to disseminate information and findings. *Don't undermine the scientific integrity of the research program by mis-timing communications with the media and other stakeholders.* *Have a response prepared in case of serious incidents – such as a crash involving a test vehicle. Anticipate media contact between the media and participant drivers.* Be aware that fleet/truck drivers may be more inclined to disclose opinions to the media if asked.	released to the media. Be prepared for the possibility that politicians may at times want to veto communications between the FOT project team, the media and other stakeholders. Building support outside the project for the FOT aims can help provide protection against strong partners/sponsors who may wish to change its direction in a way that could compromise it. *Early negative media attention may have a significant impact on participant recruitment and/or colour participant expectations of system performance. Try to prevent any media awareness until after the recruitment phase is complete.*
5. **Design the Study**	• Become familiar with the methods, measures and procedures of previous FOTs: ○ Read the FESTA handbook ○ Attend FOTNET seminars and similar events and networking activities	• 2, 3, 4	*Ensure that necessary historic data (e.g. data on vehicle speeds on certain roads) is available for baseline comparisons or Cost Benefit Analysis.* Allow sufficient time between vehicle allocations for system maintenance and verification, servicing and repairs to be undertaken. Accept that it is impossible to design a perfect FOT.	See Chapter 6 of the FESTA Handbook for detailed advice on designing the research study. See the FESTA Handbook reference list for published reports on previous FOTs.

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	○ Talk to experts who have conducted FOTs previously. ○ Review the relevant literature • Identify the performance indicators necessary to test the hypotheses derived in Activity 1 • Select measures (objective and subjective) that allow performance indicators to be derived to test the hypotheses • Identify the sensors and sensor requirements for obtaining the required measures • Design the experimental methods, tools and procedures for testing the hypotheses • Define methods, tools, requirements and procedures for acquiring, storing, transferring, de-coding, reducing/transcribing, filtering, backing up and verifying the data. • Define methods, tools and procedures for analyzing the data. • Determine optimal sample size (conduct power analyses) to ensure sufficient statistical power.	 • 2, 3, 4 • 2, 3, 4 • 2, 3, 4 • 2, 3, 4 • 2, 3, 4 • 2, 3, 4 • 2, 3	Many practical issues – including time and money — will constrain the final experimental design. Remember that an FOT is not an experiment – control is limited, and counterbalancing may not be possible. *Design into the FOT a contingency plan, in case there is an unexpected requirement to reduce or increase the scope of the study (e. g, to save money or time).* Employ a multidisciplinary team in developing hypotheses. This should include researchers and people with expert knowledge about the systems to be tested. *Design the study in a way that allows for direct comparisons to be made between objective data (logged by the vehicle) and subjective data (collected through questionnaires, focus groups etc).* Keep to an acceptable minimum the number and size of questionnaires that must be completed by participants at different points of the study, to maximise the likelihood of them being completed. A sub 2-hour completion duration is a useful target, as longer sessions may tend to remind participants that they are part of a scientific study. *Don't be tempted to reduce the sample size in order to save money – conducting a study with too few participants leads to a lack of statistical power to detect effects, and may ultimately be a waste of time and money.* Make sure that everyone understands the FOT study design, so that they appreciate the timing issues and the consequences of wanting to make changes to it e.g. if wanting to reduce the scope of the study. Delays in one area of the program cannot necessarily be made up by making sacrifices to other areas.	Where it is not possible, for ethical, practical or safety reasons, to investigate an issue in a FOT, consider safe alternative means for doing the research (e.g. simulators, test tracks). *The level of driver familiarity with the test vehicle may influence driver performance during the early stages of the FOT.* Ethical incentives that can be given to discourage driver attrition from the study should be agreed on early in the project. The models for estimating safety and other benefits may need to be updated in response to recent literature when making the estimation. *For the business sector, the commercial impact of the technologies deployed (e.g. in terms of productivity, return on investment, cost savings, incremental revenues by getting more customers, customer loyalty, etc) will be important to evaluate.*

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the “dos” and “don’ts”)	General Advice
	• Select models for estimating the potential safety, environmental and other benefits of the technologies tested. • Sign off on study design, methods and tools, questionnaires and associated procedures.	• 2, 3 • 2, 7, 6, 9, 12	*Don’t assume that FOT drivers are the only ones who will drive the FOT vehicles.* Don’t be pressured into changing the design of the study if, in doing so, it compromises the scientific integrity of the study. Ensure that all terms and phrases making up the research questions and hypotheses are clearly defined and unambiguous. This will facilitate interpretation of the FOT outcomes and comparisons with previous and future FOTS. *When performing the sample size calculations, allow for participant attrition, e.g. if using fleet drivers, some may leave the company during the FOT period.* Where hypotheses are not supported, consider conducting a process evaluation. This can help determine whether the system did not work, or whether any implementation issues may have impacted on the results.	
6. **Identify and resolve FOT legal and ethical issues**	• Seek specialist advice to identify relevant legal and ethical issues • Resolve all legal and ethical issues that can be identified in advance • Create contracts and confidentiality agreements with all relevant parties (e.g. car leasing organisations, suppliers, consultants, fleet managers, researchers etc) for all relevant issues (e.g. data collection, provision and usage, theft, insurance, privacy, duty of care, property, disposal of vehicles after	• 2, 8, 9 • 2, 7, 8, 9 • 2, 7, 8, 9	*There must be mutual agreement on the relative risks to all parties before contracts are signed*. *Double check that the final design and conduct of the FOT accords with ethical and legal requirements in all jurisdictions in which the FOT will physically occur.* Ensure that all intellectual property issues are identified and resolved “up front”. Ensure permission to drive (and necessary insurance cover) restrictions are understood by all parties, particularly participants. Identify the conditions under which a participant will be expelled from the study, and ensure these are made known to participants before the FOT commences.	See Deliverable D6.3 Annex A and Chapter 3 of this FESTA Handbook for detailed advice on legal and ethical issues.

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	the study, etc) • Seek ethics approval to conduct study (where required) from relevant ethics committee • Seek expert advice regarding liability issues and to ensure insurance provision is adequate for all foreseeable eventualities • Ensure that vehicle's licensing requirements are adhered to in spite of the modifications (implementation of data logging equipment and possibly systems to be evaluated, etc.) • Obtain informed consent of participants before they are allowed to participate in the FOT • Sign off on all aspects of the FOT design and procedures pertaining to legal and ethical matters.	• 2, 3, 4, 9 • 2, 8, 9 • 2, 7,8, 9 • 2, 3, 9 • 2, 7, 8, 9, 12	Ensure that all participating drivers are fully licensed to drive the test vehicles. *Don't forget about the need to adhere to contractual obligations and confidentiality agreements. FOTs often extend over long periods, making it easy to lose sight of obligations and agreements*. Clarify participant responsibilities and the study's obligations to the participants. Participant responsibilities should include routine vehicle maintenance activities e.g. checking fluid levels. *Ensure all relevant health and safety requirements of participants and the study team are met.* Clarify data use with participants in order to allow for anonomised data to be passed to 3rd parties. (NB with GPS and video data it may be very difficult to guarantee anonymity). All project staff must understand who has access to project data, especially video data. *All study team members must understand the agreed response should a major event, such as an accident, occur. Media comment should only be made by the appointed spokesperson.* Don't underestimate the complexity and the time commitment involved in identifying and resolving the legal and ethical issues associated with the conduct of a FOT. *Ensure that all methods, tools, procedures and materials used in the study that require legal and ethics approval are approved by the ethics committee at appropriate points in the study.*	

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
7. **Select and obtain FOT Vehicles**	• Specify functional requirements, performance specifications and user requirements for the test vehicles needed for the study. • Specify functional requirements and performance specifications for the *integration* into vehicles of all technologies needed for the FOT (FOT technologies, support technologies and data collection technologies), if these are not already in the vehicles. • Select test vehicles (makes and models) that meet above requirements. • Purchase, lease, hire or borrow (where the driver owns the vehicle) the test vehicles. • Sign off on selection and obtaining of test vehicles.	• 2, 3, 4 • 2, 3, 4 • 2, 4 • 2, 8 • 2, 4, 7, 12	The choice of vehicles may well impinge on the selection of participants which, in itself, will impact on the research questions. Choice of vehicles must be undertaken at an early stage in the project's planning. Be aware of the large costs associated with leasing vehicles that are used in FOTs. *Consider obtaining one or two extra vehicles. These can be used as spare vehicles in case of vehicle/system failure and as "showcasing" vehicles. The latter can be driven at appropriate times by politicians and other high ranking officials in positions of authority to promote and deploy the systems on a wider scale.* Be aware that vehicle choice may affect participant response if the test vehicle is significantly better/worse than the vehicle they are used to driving. Choose a conservative model. *Do consider vehicle maintenance requirements and the dealer network that is available in the FOT area. If the FOT will take place in a limited area, consider advising the local dealer(s) of the study. This may be important if a participant takes a test car to a dealer to fix a problem.*	See also Chapter 7 of this Handbook for information relevant to this Activity. The test vehicle will vary, depending on the nature of the FOT. In some FOTs, the test vehicles will already contain mature OEM systems. In others, the systems will need to be developed (fully or partly) and integrated into the vehicles. In some FOTs, the systems will be integrated into drivers' own vehicles; in others, they will be integrated into company fleet vehicles. The vehicles must be capable of hosting the technologies to be evaluated (OEM, aftermarket and nomadic) and the data logging and support systems. In some FOTs, where pre-production systems are being tested, one or two additional vehicles have been obtained, which serve as pilot platforms for ironing out bugs prior to deployment of the test vehicles.

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
8. **Select and obtain systems and functions to be evaluated during FOT (if they are not already implemented in the vehicles)**	• Develop selection criteria for choosing systems and functions (OEM, aftermarket and nomadic) to be tested (if the technologies to be tested have not already been selected by the sponsor; see General Advice column). • Use above selection criteria to select and obtain systems to be tested • If commercial systems do not exist, that meet the above criteria, develop functional requirements and performance specifications for systems that do, (including for HMI and security issues). • Develop functional requirements and performance specifications for the infrastructure needed to support the deployment of the technologies to be tested (e.g. digital maps, roadside beacons). • Source infrastructure that meets the above functional requirements and specifications. • Where infrastructure is not commercially available, develop supporting infrastructure that meets the above functional requirements and performance specifications. • If appropriate, issue Expressions of Interest/Requests for Tenders for provision of systems and supporting	• 2, 3, 4, 6 • 2, 3, 4 • 2, 3, 4 and (if appropriate) consultant • 2, 3, 4, 6, and (if appropriate) consultant • 2, 4 and (if appropriate) consultant • 2, 3, 4 and (if appropriate) consultant • 2, 7, 8, 9	*Ensure that criteria for the selection of candidate technologies (where this is required) to be evaluated are developed in consultation with relevant stakeholders, to ensure the systems to be tested meet the needs of all relevant stakeholders and are suitable for in-car use (this includes good interface design).* Selection of technologies must be undertaken with consideration of the data-logging system. If not, problems of interfacing may result. Beware of hidden costs of hardware and software development if these technologies are not originally designed for research purposes. *Do not underestimate the amount of time it will take to obtain services if public service organisations are called on to provide infrastructure to support the FOT, especially for cooperative systems eg digital maps. It may take months or even years.*	See also Chapter 7 of this Handbook for information relevant to this Activity. Criteria for selection of candidate technologies in the FOT (if they have not been pre-selected by the sponsor) could include: likely safety or environmental benefit, likely benefit in increasing commercial productivity and efficiency, availability, compatibility with host vehicles, technical performance, cost, reliability, maintainability, likely acceptability to drivers, usability, compliance with relevant human factors/ergonomic guidelines, compliance with local legal requirements, compliance with relevant standards, crashworthiness etc. If prototype systems are tested, then estimates of durability, reliability, maintenance costs etc of production systems will be difficult, and full Cost Benefit Analyses may not be possible.

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	infrastructure. • If appropriate select preferred tenderers, negotiate contracts and award contracts. • Decide what will be done with the test vehicles, and the equipment in them, once the FOT has been completed. • Sign off on selection and obtaining of technologies to be evaluated during the FOT	• 2, 7, 8 • 2, 3, 4, 5, 6, 9, 12 • 2, 4, 7, 12		
9. **Select and obtain data collection and transfer systems for FOT vehicles**	• Specify data to be logged (measures and sampling rate) • Specify functional requirements and performance specifications for systems for collecting and transferring the data to be logged. • Source, purchase and/or develop systems for logging and transferring the data that meet the above functional requirements and performance specifications. • Sign off on selection and obtaining of data collection and transfer system for test vehicles.	• 2, 3, 4 • 2, 3, 4 • 2, 4 and (if appropriate) sub-contractors. • 2, 7, 9, 12	Implement re-calibration procedures that will ensure accuracy of measurements/sensors over time and help prevent data drift issues. Plan for software upgrade and revision during the FOT and try to ensure that all software systems are updated together. Ideally, this should be possible remotely. *In-vehicle data logging systems need to be unobtrusive, safe and secure — but they also need to be accessible to enable routine repairs.* *Provide a local location for vehicle support and a vehicle tracking capability.* Minimise driver involvement in data download from test vehicles. *Ensure boot-up time for test systems and data logging systems is sufficiently fast to prevent data loss at the beginning of each trip.* *Ensure that a common time stamp is used for all recorded data sources.*	See also Chapters 7 and 8 of this Handbook for information relevant to this Activity. The technologies fitted to the test vehicles may also include supplementary technologies (such as sensor technologies; e.g. forward looking radars, GPS) that are needed to, for example, measure inter-vehicle following distances in order to determine whether speeds are free or constrained (e.g. see Regan et al, 2006, Volume 1). See Deliverable D6.3, Annexe A and Chapter 3 on legal issues of data privacy to be aware of, possible dangers,

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
			Verify the definition of signals provided by 3rd parties (e.g. CAN message definitions by car manufacturers) Do not allow data collection to proceed automatically without active confirmation of data capture and validity. This may include the generation of warning messages (SMS?) when out of tolerance data is recorded. Recognise that some data is much more important than others and should be given a relatively higher priority. *Do keep a stock of spares for critical items and anticipate that some components (e.g. PC cards) may become unobtainable during the study.* Consider the opportunities for ad-hoc and post-hoc interrogation of raw data files to answer additional questions. This may not be possible if data collection is triggered.	and legal provisions.
10. **Select and obtain support systems for FOT vehicles**	• Define the support systems needed (see General Advice Column) • Develop functional requirements and performance specifications for systems needed to support the study • Where appropriate, develop functional requirements and performance specifications for the HMI, to ensure that the HMI for support systems is safe and user-friendly • Source, purchase and/or develop support systems that meet above functional requirements and performance specifications	• 2, 3, 4 and (if appropriate) consultant • 2, 3, 4 and (if appropriate) consultant • 2, 3, 4 and (if appropriate) consultant • 2, 4 and (if appropriate) sub-contractors.	Ensure that the data logging system is capable of logging whether the test vehicle is driving forward and in reverse. If possible, support systems should be capable of remote operation to allow, for example, remote system re-boot. In the case of very large naturalistic studies it may not be practicable to provide operator support systems. In these cases attempt to automate as much as possible. Anticipate data analysis requirements before specifying data to be logged (e.g. rates and resolution). Ensure that missing data are clearly indicated – e.g. if the data collection system malfunctions, missing data should NOT be indicated with a zero, where zero is a valid measure (e.g. speed). If in doubt about the final list of measures to be logged, log more parameters if performance of the data logging	See also Chapter 7 of this Handbook for information relevant to this Activity. Support systems have multiple purposes: e.g. to display information to drivers; to automatically turn systems on and off where multiple systems are being tested and exposure to each is kept constant across drivers; for manually disabling systems in the event of malfunctions (i.e. "panic buttons"); for preventing use of systems by non-participants; for diagnosing system status and

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	• Sign off on selection and obtaining of support systems for test vehicles.	• 2, 7, 12	system or storage capacity are not affected.	faults; etc.
11. **Equip FOT vehicles with all technologies**	• Prepare a system installation/integration specification. • Equip test vehicles with the FOT technologies to be evaluated (if not already in vehicles) • Equip test vehicles with data collection and transfer systems • Equip vehicles with FOT support systems (e.g. panic button, for turning systems on and off etc) • Sign off on system integration activities, ensuring that all systems have been installed in accordance with the system installation/integration specification.	• 2, 4 and (if appropriate) sub-contractors. • 2, 4 and (if appropriate) sub-contractors. • 2, 4 and (if appropriate) sub-contractors. • 2, 4 and (if appropriate) sub-contractors. • 2, 7, 12	Ensure that the in-car computer driving all systems (FOT, data collection and support) has sufficient computing power to avoid processing delays. *Ensure that all systems (FOT, data collection and support) operate identically across test vehicles.* *Allow all new vehicles a "burn-in" period (around 1000km) so that vehicle faults, that could disrupt the FOT, can be detected.* Be aware that 'identical' vehicles, fresh off the production line, may perform differently due to variation in components and manufacturing variability. Check for differences that may be critical for the FOT. Try and make all adaptations to test vehicles (e.g. fitment of novel display systems) invisible to reduce the likelihood of theft or behaviour modification by other drivers. *Create protocols that standardise the procedure for installing all in-vehicle equipment.*	See also Chapter 7 of this Handbook for information relevant to this Activity.
12. **Design and implement driver feedback and reporting systems**	• Design, develop and implement systems and procedures to allow drivers to report technical problems in a timely manner. • Design, develop and implement systems and procedures to allow drivers to provide	• 2, 3, 4 • 2, 3, 4	Possibly implement 'driver diaries' to allow confirmation of driver identity and trip details (if this process cannot be automated using a smart card, i-button or other technology). This may, however, encourage the drivers to behave less naturally. Implement a timetable for the timely collection of qualitative data so that participants don't have to rely on their memories.	See also Chapters 6 and 7 of this Handbook for information relevant to this Activity.

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	feedback to researchers, in real time or retrospectively (e.g. usability problems, opinions of systems, confirmation that systems are operating as required etc) • Design, develop and implement systems and procedures that allow researchers to monitor participant progress (e.g. to ensure they are adhering to study requirements). • Sign off on implementation of driver feedback and reporting systems and procedures	• 2, 3, 4 • 2, 3, 4, 7, 12	*Anticipate that drivers may not complete diaries accurately or consistently and may fail to attend for de-briefing interviews. Appoint driver liaison staff as a single point of contact.* *Ensure that the project team can respond to emergencies and incidents on a 24/7 basis.* Ask participants to announce when they are going on holiday or not driving for an extended period. Keep a record of all reported problems, and document these in relevant reports. Ensure that all feedback and reporting procedures are documented in a manual for quick reference by the research and technical support team as required. Consider whether you need to design, develop and implement a system to allow for the collection of fuel consumption information. Where fuel consumption is calculated manually, anticipate that drivers will not always use fuel cards, return fuel dockets or fill in the fuel logbook.	
13. **Select, obtain and implement standard relational database for storing FOT data**	• Design, develop and implement a database for storing data logged from the test vehicles • Design, develop and implement a database for storing the subjective data collected from participants (e.g. from questionnaires, from focus groups, from feedback lines etc) • Develop data navigation and visualization tools	• 2, 3, 4 • 2, 3, 4 • 2, 3, 4	Before an FOT is launched, the database architecture should be reviewed by a system evaluator to ensure that all requirements are fulfilled. *Ensure copies are made of raw data, reduced raw data and all processed data files and store these securely, separate from the primary data store.* Use an industry standard relational database to store the data. *Ensure that unauthorised access to the database is not possible. Preferably, do not give the database host an IP number.* Careful database design can reduce the need for post-	See Chapter 8 of the FESTA handbook for more detailed advice relating to this activity. Basic legal advice on this issue is also provided in Chapter 3, and Deliverable D6.3 and the Annexe A.

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	• Sign off on database for storing FOT data.	• 2, 3, 4, 7, 9, 12	collection manipulation if the database is designed to feed directly into a statistical package for data cleaning and analysis. *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 meta data should be kept, and how.*	
14. **Test all technologies against functional requirements and performance specifications**	• Develop "acceptance testing" protocols (see comment column). • Test the technologies for acceptance, using the acceptance testing protocol. • Develop a usability test plan for the purpose of assessing the systems for usability. • Conduct usability testing, using the usability testing plan, to ensure systems are user-friendly and that the systems meet all usability assessment criteria. • Obtain or develop a valid and reliable ergonomic checklist.	• 2, 3, 4 • 2, 3, 4 • 2, 3 with consultant (if appropriate) • 2, 3, with consultant (if appropriate) • 2, 3	*Do not sign off on the outputs of any of the previous activities until all technologies have been tested and, where appropriate, refined.* Be sure that all systems are designed so they do not drain the battery when the engine is not running. *Be sure that retrofitted systems are properly secured and meet all relevant crashworthiness requirements.* If sub-contractors are appointed to install or maintain test equipment, implement a quality assurance programme. *Be aware that system clocks can drift significantly if left to run independently – although, GPS time can be used to correct system clock error.* *Implement procedures to ensure that alignment and calibration of sensors is maintained and tested in all potential weather conditions.* Various guidelines, standards and checklists exist for assessing the ergonomic quality of the human-machine	This activity is *not* about pilot testing — it is about testing the performance, security and reliability of systems – to ensure that all technologies to be deployed perform in accordance with the functional requirements and performance specifications developed for them in previous activities. An Acceptance testing Protocol is a test protocol for testing that all systems to be used in the study (FOT systems, data collection systems and support systems) meet the functional requirements and

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	• Assess systems, using the ergonomic checklist, to ensure that they meet all relevant criteria. • Assess vehicles against relevant certification procedures to ensure that vehicles are safe, roadworthy and comply with all relevant National, State and Territory laws, treaties and other protocols. • Ensure that all vehicle modifications that affect primary safety are signed off by a competent engineer or appropriate testing authority. • Rectify all technical, usability, ergonomic and certification issues where deficiencies are noted. • Sign off on completion of all systems tests.	• 2, 3 • 2, 4 with consultant (if appropriate) • 2, 4 with consultant (if appropriate) • 2, 3, 4 with consultant (if appropriate) • 2, 3, 4, 7, 12	interface for ICT systems (see Chapter 25 of Regan, Lee and Young, 2008, for a summary). Also see: AIDE (EU-Project) Deliverable 4.3.1: "Report on the review of available guidelines and standards" – publicly available over the internet. *Be aware that some system components may become corrupted over time with continuous use (e.g. flash memory cards).* Create an installation manual for all vehicle modification procedures. Consider the need to obtain waivers/special licences from regulatory authorities for equipment that is non-compliant (e.g. radars that operate outside legal bandwidths). Standard testing of vehicle modifications by a competent authority may be necessary with respect to safety features (e.g. proper deployment of airbags following modification to vehicle interiors). *Be aware that some systems (e.g. displays) that are not OEM-installed may fail in automotive environments.* *Where appropriate, test for radio frequency (RF) interference effects (e.g. from overhead tram wires), which may adversely affect system operation. Also ensure that normal vehicle systems (e.g. FM radio and remote locking) are not affected by installed equipment.* Ensure that the in-car computer powering the data collection system and support systems is powerful enough to ensure that the data sampling rate is consistent and at the rate specified. Don't assume that OEM systems that are already installed in test vehicles have been ergonomically assessed against appropriate standards and guidelines.	performance specifications developed for them by the FOT project team, under all foreseeable operating conditions. The term "usability" can mean different things to different people. The test plan should use a standard definition of usability (e.g. ISO 9241). Be aware that the frequency used by some radar-based systems may interfere with the operation of other systems used by Police, emergency services or other operators (or vice versa) when used in other countries or jurisdictions. This must be investigated where the FOT is conducted across State and international boundaries.

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
			Ergonomic assessment of systems prior to system deployment can be useful in identifying ergonomic problems that may explain or confound treatment effects. Provide a written statement for the participants to keep (in the vehicle) which confirms their participation in the FOT and the nature of vehicle modifications – in case they are challenged by Police or other authorities. *Resolving any technical, usability, ergonomic, and certification issues may require several iterations. Do not underestimate the time required for this process.*	
15. **Develop FOT recruitment strategy and materials**	• Develop recruitment strategy, including driver entry and exit requirements and procedures. • Develop recruitment materials and procedures • Sign off on recruitment strategy, materials and procedures.	• 2, 3, 9 • 2, 3, 11 • 2, 3, 7	Where possible, ensure drivers are representative of the relevant driving population to ensure results can be generalised. Assume that there will be an attrition rate of about 10 to 15 % when using company drivers, who come and go, and retire. Be aware that, when company drivers change jobs within their companies, this may have a dramatic effect on their annual mileage rates. If fleet drivers are recruited via a fleet owner or manager it is also necessary to get buy-in from individual drivers. *Assume that it is much harder to recruit women than men when using company drivers.* With respect to safety, select drivers who do not pose a risk to themselves, others or the project, but without biasing the participant sample. *It is harder to recruit company drivers than lay people.* Do not underestimate the complexities involved in recruiting company drivers (see Regan et al., 2006, Vols 1 and 2).	See Chapters 5 and 6 of this Handbook for further advice relevant to this Activity. *The Ethical requirements for recruitment of drivers may be difficult to adhere to when recruiting company drivers. In any case: ensure voluntary participation.* Ideal companies to approach to recruit fleet vehicle drivers have the following characteristics: many vehicles; drivers have high mileage rates; drivers drive primarily in the geographical areas of interest in the FOT; and management has a commitment to the aims and objectives of the FOT. It may not be possible in some

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
			Be aware that some commercial operations may have driver turn-over rates approaching 100 % per annum.	countries (e.g. Germany), to obtain personal information about drivers that can be used to screen them for inclusion in the study (e.g. has a drunk driving record). It may not be possible in some countries to obtain directly from car dealers the names of drivers of particular makes and models of vehicles. In some countries (e.g. France), potential participants must be screened by a registered doctor. *The recruitment materials and procedures will need to have been incorporated and approved as part of the FOT ethics and legal approval processes.*
16. **Develop FOT driver training and briefing materials**	• Conduct training needs analysis (TNA) to identify training requirements (if appropriate) • Design and develop driver briefing and training materials, based on outputs of the TNA. • Design and develop briefing materials for participating car/truck fleet managers (if appropriate)	• 2, 3 with consultant (if appropriate) • 2, 3 with consultant (if appropriate) • 2, 3, 9	*Ensure that training programs and briefing materials are designed in a way that does not confound experimental treatment effects.* Ensure all drivers understand existing in-vehicle systems as well as test systems, especially if use of them is required as part of a baseline comparison. Don't underestimate the time required for the development of briefing and training materials — it is a time consuming activity. Anticipate that some car manufacturers will not wish the participating drivers to receive any training about	See Chapters 3 and 6 of the FESTA handbook for further advice relevant to this Activity See Regan et al, 2006 (Volume 2) for examples of training and briefing materials used in a previous FOT. *Refresher training may be required if FOT systems are not activated for several weeks or months into the FOT.*

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	• Design and develop FOT system(s) user manual (if appropriate) • Design and document the procedures for the delivery of the briefing and training to the FOT participants • Sign off on driver training and driver (and company) briefing materials and delivery processes.	• 2, 3, 9 • 2, 3 • 2, 3, 7	implemented systems. In such cases, "structured familiarisation" may be more acceptable. Provide drivers with a mini-operating manual to keep in the vehicle and prepare written materials (brochures, DVDs & CDs) that can be taken away after briefing sessions as memory joggers for important information.	The training and briefing materials and procedures will need to have been incorporated and approved as part of the FOT ethics and legal approval processes.
17. **Pilot test FOT equipment, methods and procedures**	• Develop protocol for pilot testing FOT equipment, methods, procedures and materials (including training, briefing materials and data collection, downloading and analysis procedures) • Recruit, brief and train pilot participants • Deploy a small sample of FOT vehicles under a representative range of driving conditions that will be experienced in the FOT, as per the pilot testing protocol. • Fine tune FOT vehicles and technologies, systems, procedures and protocols, as required, on the basis of the pilot data yielded.	• 2, 3, 4, 9 • 2, 3 • 2, 3, 4 • 2, 3, 4, 7	*Do not truncate your pilot test plan, and do not underestimate the time required for comprehensive pilot testing. The importance of pilot testing cannot be overstated.* Undertake a 'full dress rehearsal' of the FOT on a scale that is smaller than the FOT but big enough to properly test all systems, procedures, and equipment. Use pilot testing also as a means of estimating the amount of time required to complete activities, as this will enable more accurate budgeting during the remainder of the project. Pre-test all data analysis procedures to ensure appropriate data is collected – particularly data related to event recording triggers. Ensure that the routes used in pilot studies maximise the likelihood of critical situations of relevance to the FOT. Add independent monitoring systems to pilot vehicles to ensure the validity of data derived from sensors. *In the pilot phase listen to the drivers as well as the owners of the vehicle fleet – their ideas are likely to be*	See also Chapters 6, 7, 8 and 9 for further advice relevant to this Activity. For data collection systems, ensure that data is being recorded, determine the accuracy of data recorded, test downloading procedures and equipment, test reader software and analyse samples of pilot data.

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	• Sign off on pilot testing.	• 2, 3, 4, 7	*different.*	
18. Run the FOT	• Ensure that all sign offs have occurred for previous activities. • Manage the FOT: • monitor project activities, timelines, budgets and resources • prepare regular progress and financial reports for sponsor • convene and attend regular meetings with research and support teams • maintain communication with sponsor and key stakeholders • Recruit participants • Organise training session times/materials • Brief and train participants • Brief fleet managers (if appropriate) • Deploy FOT vehicles • Regularly monitor participant progress, including kilometres travelled • Administer questionnaires and implement other data collection methods at pre-	• 2, 7 • 2, 3, 4, 5, 7 • 2, 3 • 2, 3 • 2, 3 • 2, 3 • 2, 3, 4 • 2, 3 • 2, 3	Anticipate, and plan for, driver 'dropout' throughout the FOT — over-sample. It is rarely possible to replace drivers who drop out after more than a few days without affecting the timing plan. *Develop protocols for responding to drivers with technical and other problems (e.g. provide drivers with a dedicated cell phone to report problems; ensure at least two people have pagers to receive problem calls; etc) Timely responses will keep drivers happy.* Anticipate problems that may increase the drop out rate (e.g. higher fuel consumption in the FOT vehicle than in the drivers' own vehicle) and take steps to prevent or mitigate these problems. Monitor closely system usage for drivers who you suspect may be tempted to 'demonstrate' novel systems to friends and neighbours. Adhere to quality control mechanisms to ensure that data is being properly recorded and downloaded. Adhere to calibration procedures to ensure accuracy of measurements/sensors over time and help prevent data drift issues. Find a suitable location for training drivers where you can also assess transfer of training to the test vehicles in a safe environment *If the number of kilometres driven by drivers is being controlled for, conduct regular calibration checks of cumulative distance travelled.* Assume that it will take you 50 % longer than you think to recruit participants if recruiting company drivers. Check logged data as soon as you receive it to verify	More detailed advice relevant to this Activity can be found in other chapters of this Handbook: Chapter 6 for participant recruitment; Chapters 3 and 6 for organising training sessions; Chapters 6 and 7 for implementing data collection methods at pre-determined intervals and for collecting and storing subjective data; Chapter 9 for preliminary data analysis; Chapters 3, 7 and 8 for vehicle maintenance and compliance with laws; and Chapter 3 for reporting of dangerous driving, where appropriate. *Ongoing communication with key stakeholders is important during the FOT to ensure that the aims and objectives of the FOT are clear, that stakeholders stay committed to the project, and that the aims and objectives of the FOT are not misquoted, misrepresented or misunderstood.*

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the “dos” and “don’ts”)	General Advice
	determined intervals • Collect, enter into database (unless automated) and store subjective data • Record, download and store objective (i.e., logged) data • Collect special data (e.g. fuel dockets) needed to analyse surrogate performance indicators • Monitor for, collect and document data on technical problems and user feedback • Commence preliminary evaluation of data, to identify instances of dangerous driving and any other findings of interest/relevance to FOT outcomes • Repair and re-deploy vehicles (as required) • Routinely ensure vehicles and vehicle systems are properly maintained and legal in other ways (e.g. registered, licensed, tyres properly inflated) • Report dangerous driving behaviours (if legally required) • Conduct exit interviews with drivers and fleet managers • Remove systems and	 • 2, 3 • 2, 3, 4 • 2, 3 • 2, 3, 4 • 2, 3, 4 • 2, 4 • 2, 3, 4 • 2, 3, 4 • 2, 3 • 2, 4	accuracy and completeness of data and verify kilometres travelled. *Monitor and record critical factors that could have an impact on the measured outcomes/dependent variables (e.g. changes in Police enforcement strategies, unseasonal weather conditions). If these are not controlled for in the experimental design, or accounted for in the analyses, they could confound the measured effects of the systems being tested.* Where company fleet vehicles are involved in the study, advise fleet managers not to “demonstrate” their vehicles, as this may compromise the aims of the study. *Give sponsors early warning of potential problems that could compromise the integrity of the study, or increase the budget.* Encourage participants to report technical problems as soon as possible. *Don’t assume that all systems in the test vehicles are functioning as required. Develop systems to check, at appropriate times, that they are operating properly.* Don’t assume that drivers will do what you ask them to do (e.g. to fill out questionnaires; maintain vehicles). They need regular reminding and follow-up. Where data downloading is manual, don’t forget to replace flash memory cards, or other storage devices, with new (empty) ones on a regular basis. Do not always assume that drivers will clock up their kilometres evenly over the trial. Contact them on a regular basis to check cumulative distance logged. If legally required, don’t forget to report to the appropriate authorities (e.g. company fleet managers)	

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	equipment from private vehicles (if used) • Sign off on completion of this activity of the FOT.	• 2, 3, 4, 5, 6, 7, 8, 10	recorded instances of dangerous driving by test drivers. *Don't assume that drivers will drive the vehicles without trailers, bike racks and other accessories. These may affect the operation of some FOT systems (e.g. reverse collision warning devices).* Minimise interference to commercial operations during FOTS, especially trucking operations. Problems that compromise commercial productivity may result in companies withdrawing vehicles from the FOT. *Make sure fleet managers are, and remain, motivated. Their support is critical.* Be careful about the feedback given to drivers. They may be concerned about the possibilities of 'unintended consequences' e.g. their managers learning how and when they take rest breaks etc. Participants are more likely to comply with what is asked of them if they engage with the project. Ongoing communication and even small incentives can enhance perceived engagement and improve compliance. However, the level of engagement must not compromise the outcomes of the study. Remember that long-term involvement in a research study can be onerous for a participant. At all times treat them as participants in the study process, not simply subjects of a study. *Allow sufficient time for any data entry which has to be done manually (e.g. responses from pencil and paper questionnaires, focus groups). As far as is possible, manual data entry should be carried out routinely during the course of the data collection phase and not all left to the end.* A system for basic inventory management is	

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
			recommended for FOTs with more than a few vehicles in use. For such a system to be efficient, sensors, data acquisition system units, vehicles and all other equipment need to be included, as well as relevant supporting procedures developed.	
19. **Analyse FOT data**	• Develop a data analysis plan • Analyse objective (i.e., logged and recorded data) • Analyse subjective data (i.e., data obtained from interviews, questionnaires, focus groups, hotlines, etc) • Draw conclusions with respect to the hypotheses generated for the FOT • Sign off on completion of all required analyses	• 2, 3, 4 • 2, 3 • 2, 3 • 2, 3 • 2, 3, 7, 12	Plan for the fact that there will be constant demand for study findings, such as general trends in the data, early in the project, even though the data may not be statistically reliable enough to report with any confidence. In a well-powered study, null findings (i.e., where no effect is found) are potentially as interesting as when the hypotheses are supported. *Anticipate the requirement to have to perform supplementary analyses for the funding organisation, which may be expensive and not originally budgeted for. This will require negotiation with the sponsor if these analyses are expected to be carried out within the original budget.* Anticipate that, unless distance travelled is controlled for in the FOT, the distance travelled by different drivers will vary significantly. Take this into account in the analysis to ensure results are not skewed. Don't forget to run "reality checks" on the data, to be sure that the data are "clean". This is essential. *If data is reduced/aggregated, always keep a copy of un-aggregated data.* Ensure that all data analysts have used the test vehicles and understand the circumstances in which data was/is collected.	See Chapters 6, 8 and 9 of this Handbook for detailed advice on data analysis tools and methods. There may be a requirement to conduct ongoing analysis, such as ongoing identification of dangerous drivers, determining whether adaptation to systems is occurring early enough to warrant a shorter FOT duration (e.g. to save money and time), and to identify early trends in the data. These checks should be built into the analysis plan at the start of the project. Some FOTs have developed novel ways of turning ADAS technologies on and off to control precisely the amount of exposure to the technologies that are being evaluated (see Regan et al, 2006, Vols 1 and 2).

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
			All team members who handle participant data should receive appropriate training in relation to data privacy issues. Work out how to best filter logged data and deal with missing data.	Sponsors need to be calibrated about the relative costs of running FOTs. For example, the cost of running simulation models at the end of the FOT to estimate safety and other benefits of ICT technologies is a fraction of the cost of preparing and deploying the FOT vehicles.
20. **Write minutes and reports**	• Write minutes of regular project management team meetings • Write regular minutes of Project Steering Committee meetings • Write quarterly progress reports for the sponsor(s) • Write the draft FOT report • Send the draft FOT report to relevant stakeholders and peers for peer-review • Convene 1 or 2 meetings to discuss feedback with sponsor/peers • Incorporate feedback and write final report. • Deliver final report to sponsor(s)	• 2 • 2 • 2 • 2, 3, 4 • 2 • 2 • 2, 3, 4 • 2	*Use regular progress reports to document problems, solutions and lessons learnt.* Allow sufficient time for sponsor review of draft and final reports, but not so long that the review process drags out unduly. Six to 8 weeks is recommended. Consider peer review of major outputs; this will improve their quality, but delay their release. *Document all lessons learnt in the final FOT report.* Ensure that the final report contains practical recommendations for wider scale deployment of those systems found to be effective, and for fine-tuning of those with potential to be more effective. *Develop, in consultation with the Project Steering Committee, a suggested plan for implementing the recommendations deriving from the FOT. Document the suggested implementation plan in the FOT final report.*	*The FOT lifecycle is long. Hence, it is advisable to write separate reports on each critical stage of the FOT, particularly the lessons learned, to ensure that nothing important that should be documented is forgotten.* Formal meeting minutes are a critical resource for the project in confirming departures from the project plan.

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the “dos” and “don’ts”)	General Advice
	• Sign off on completion of all required reports	• 2, 3, 4, 7, 12		
21. **Disseminate the FOT findings**	• Send regular project reports to the sponsor • Disseminate preliminary and final findings at seminars, conferences and special events • Prepare reports on preliminary findings for the sponsor • Send sponsor draft and final FOT reports • Provide other stakeholders with access to FOT final report (s) and, if allowed, raw or filtered data from the FOT • Showcase the vehicles at relevant events during the FOT (e.g. Smart Demos, motor shows) to promote awareness and wider deployment of systems. • Sign off on completion of all dissemination activities	• 2 • 2, 3, 4 • 2, 3, 4 • 2 • 2, 3, 4 • 2, 4, 6, 7, 12 • 2, 3, 7, 12	Disseminate the findings in accordance with the previously agreed communications plan, and other contractual obligations (e.g. as specified in EC-funded projects). *Agree on what can and cannot be disseminated and said at different points in the study.* Seek necessary permissions prior to divulging FOT findings to any third party. FOT reports are large and expensive to print. Allocate sufficient budget at the beginning of the project for printing, if required. FOT reports are large and hard to read. It is desirable to produce conference papers along the way that document the outputs of the study at different phases. Prepare a concise 1 or 2 page synopsis of the study outcomes that can be read and easily digested by politicians, chief executives and relevant others in positions of authority. Agree in advance who is empowered to release and comment on results.	Where private industry is a participant in the FOT, it may be necessary to seek permission from the manufacturer before divulging certain information deriving from the FOT. This must be established. Maintain at least one vehicle for demonstrations; preferably at a location that is convenient to politicians, officials and the press. A demonstration and briefing to an influential politician is likely to be far more effective than sending them a report.
22. **Decommission the FOT**	• Conduct de-briefing interviews with participants to elicit feedback on the FOT that can be used to improve future FOTs. • Dispose of test vehicles which are no longer needed (if	• 2, 3 • 2, 3, 4, 5, 7, 8, 12	Ensure that participants return relevant items at the end of the study (e.g. flash memory cards, i-buttons) and perform other required activities to decommission the FOT vehicles (e.g. disconnect power to support systems). *Keep one vehicle until all data analyses are complete.* Consider providing public access to FOT databases, where ethically allowed, that enables others to use the	*Consider keeping one or two vehicles as showcasing vehicles after the study, to allow stakeholders in positions of authority to experience the look and feel of the vehicles.*

Activities	Tasks and Sub-Tasks	Person/ Team/ Organisation Responsible for Activity	Critical Considerations (the "dos" and "don'ts")	General Advice
	vehicles are not privately owned). • Retrieve installed data logging equipment (if vehicles are privately owned) • Sign off on completion of all FOT activities	• 2, 4, 5, 8 • 2, 3, 4, 5, 7, 8, 9	data for other research purposes after the FOT has been de-commissioned (but remember to fully explore and address anonymity issues). The data collected and stored after the FOT is de-commissioned should be regarded as "living data". *Don't lose momentum at the end of the FOT. Lobby stakeholders to ensure that there is commitment to implementing the recommendations of the FOT.*	It may be necessary to consider legal issues when decommissioning the FOT as far as the de-installation of data logging equipment is concerned (in a contract with participants). See Deliverable 6.3 and Annexe A on legal issues.