



This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.



**CC creative commons**  
COMMONS DEED

**Attribution-NonCommercial-NoDerivs 2.5**

**You are free:**

- to copy, distribute, display, and perform the work

**Under the following conditions:**

**BY:** **Attribution.** You must attribute the work in the manner specified by the author or licensor.

**Noncommercial.** You may not use this work for commercial purposes.

**No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

**Your fair use and other rights are in no way affected by the above.**

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:  
<http://creativecommons.org/licenses/by-nc-nd/2.5/>



Large Scale Collaborative Project  
7<sup>th</sup> Framework Programme  
INFSO-ICT 224067

## D1.7 Evaluation Plan

Deliverable n.	D1.7	Evaluation Plan	
Sub Project	SP2	FOT Framework	
Work package	IP-level report		
Task n.			
Author(s)	Franzén, S. Mononen, P. Paigle, K. Morris, A.P., Montanari, R.	File name	TeleFOT_D1.7_Evaluation Plan _v.4.0.doc
Status	Final		
Distribution	Public (PU)		
Issue date	2011- 09 - 15	Creation date	2011-03-01
Project start and duration	1 <sup>st</sup> of June, 2008 – 48 months		



Project co-funded by the European Commission

DG-Information Society and Media

in the 7th Framework Programme



## TABLE OF CONTENT

LIST OF FIGURES .....	3
LIST OF ABBREVIATIONS .....	3
REVISION CHART AND HISTORY LOG.....	4
EXECUTIVE SUMMARY.....	5
1 INTRODUCTION.....	6
2 FIELD OPERATIONAL TESTS.....	8
2.1 REQUIREMENTS FOR FIELD OPERATIONAL TESTS.....	8
2.2 THE FESTA HANDBOOK.....	10
3 THE TELEFOT EVALUATION STRATEGY .....	12
3.1 BLOCK: THE CONTEXT.....	12
3.1.1 <i>TeleFOT boundaries</i> .....	14
3.1.2 <i>Functions Specification</i> .....	15
3.2 BLOCK: RESEARCH QUESTIONS AND HYPOTHESES.....	16
3.2.1 <i>The Top-down approach</i> .....	17
3.2.2 <i>The Bottom-up approach</i> .....	19
3.2.3 <i>Common research questions and hypotheses</i> .....	20
3.2.4 <i>Some recommendations</i> .....	21
3.3 BLOCK: PERFORMANCE INDICATORS AND MEASURES.....	21
3.4 BLOCK: STUDY DESIGN .....	22
3.4.1 <i>Within Subject Design</i> .....	22
3.4.2 <i>Between Subject Design</i> .....	23
3.4.3 <i>Implementation</i> .....	23
3.4.4 <i>Participants</i> .....	24
3.4.5 <i>Study environment</i> .....	26
3.5 BLOCK: DATA ACQUISITION .....	27
3.5.1 <i>Data collection</i> .....	28
3.5.2 <i>Subjective data</i> .....	28
3.5.3 <i>Objective data</i> .....	29
4 PILOT TESTS .....	31
5 DISCUSSION.....	33
6 CONCLUDING REMARKS .....	37
7 REFERENCES.....	38
APPENDIX I .....	39
APPENDIX II .....	44

## LIST OF FIGURES

Figure 1. The steps that typically have to be considered when conducting an FOT. The large arrows indicate the time line. ....	10
Figure 2 L-FOTs vs. D-FOTs .....	13
Figure 3 TeleFOT application area in the field of Intelligent Vehicle safety Systems. ..	15
Figure 4 Integration of top-down and bottom-up approaches for hypotheses generation for all impact areas (safety is referred to in the figure). ....	17
Figure 5 Description of procedure, from impact area and research questions to measures.....	21

## LIST OF ABBREVIATIONS

ABBREVIATION	DESCRIPTION
CAA	Cockpit Activity Assessment Module
CAN	Controller Area Network
D	Deliverable
DAS	Data Acquisition System
DoW	Description of Work
D-FOT	Detailed FOT
GDS	Green Driving Support
GPS	Global Positioning System
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communication
FOT	Field Operational Test
L-FOT	Large Scale FOT
NSS	Navigation Support (static)
NSD	Navigation Support (dynamic)
PI	Performance Indicator
SA	Speed Alert
SLI	Speed Limit Information
SP	Sub Project
TI	Traffic Information
IR	Internal Report
WP	Work Package

## REVISION CHART AND HISTORY LOG

REV	DATE	AUTHOR	REASON
1.0	2011-03-01	Franzén, S.	Synopsis
2.0	2011-06-01	Franzén, S.	Modified synopsis (based on feedback from partners)
3.0	2011-07-10	Franzén, S.	First draft
4.0	2011-09-15	Franzén, S. Mononen, P. Paigle, K. Morris, A.P., Montanari, R.	Final report (with peer review input from the SP Leaders)

## EXECUTIVE SUMMARY

The overall objectives of TeleFOT are to assess the impacts of aftermarket and nomadic devices used in vehicles for driver support and to raise the awareness of the functions and potential that these devices offer. This report starts from the scientific and technological objectives that will make these overall objectives more concrete. They are

1. *Build, mobilise and integrate European test communities for long term testing and assessment of driver support functions through aftermarket and nomadic devices*
2. *Create a methodological framework for executing the tests and analysing the data*
3. *Study aftermarket and nomadic devices in different technical contexts*
4. *Study different levels of impacts on drivers and society*
5. *Focus on functions and services for safe, efficient and economical travel*
6. *Investigate the contents of functions provided for cooperative driver support*
7. *Develop effective procedures of enhancing awareness and take-up of driver support ICT systems among the public*
8. *Focus also on aspects in the use of aftermarket and nomadic devices that may decrease safety*

TeleFOT is supposed adopt the approach of Field Operational Test (FOT). When going into the actual work in TeleFOT, as laid out in the DoW, it was seen as a useful step to make use of the structure of the deliverable *D2.2.1 Testing and Evaluation strategy I*, based on the FESTA FOT Chain (from the FESTA Handbook). This was done in order to identify what findings in the TeleFOT project (so far) has a unique and from the FESTA Handbook deviating approach.

The intention has been to highlight these deviations (or improvements) in order to widen the potential use of the FOT methodology in the future. This IP-level deliverable is therefore focused on what constitutes the unique features of TeleFOT that could have an impact also on a more general level. This is especially important as new FOTs are planned in the area "cooperative driving". Most of the TeleFOT deliverables until today have been consulted including the first series of deliverables addressing the Data Analysis Plans for all the impact areas to be covered by TeleFOT; they are Efficiency, Environment, Mobility, Safety and User Uptake.

It is concluded that the first of the overall objectives has been met to quite a high degree, even if there still are some important steps that must be finished. The second overall objective is not yet addressed in a systematic way. However, there are WPs in the TeleFOT DoW that are supposed to cover these aspects in the last part of the project.

## 1 INTRODUCTION

The general objectives of the TeleFOT project as presented in the TeleFOT Description of Work (DoW) are firstly, to assess the impacts of aftermarket and nomadic devices used in vehicles for driver support and secondly, to raise the awareness of the functions and potential that these devices offer.

To operationalize these general objectives they have been further elaborated into scientific and technological objectives, which are the following:

- 9. Build, mobilise and integrate European test communities for long term testing and assessment of driver support functions through aftermarket and nomadic devices*
- 10. Create a methodological framework for executing the tests and analysing the data*
- 11. Study aftermarket and nomadic devices in different technical contexts*
- 12. Study different levels of impacts on drivers and society*
- 13. Focus on functions and services for safe, efficient and economical travel*
- 14. Investigate the contents of functions provided for cooperative driver support*
- 15. Develop effective procedures of enhancing awareness and take-up of driver support ICT systems among the public*
- 16. Focus also on aspects in the use of aftermarket and nomadic devices that may decrease safety*

It was assumed that the work in TeleFOT should adopt the approach of a Field Operational Test (FOT). In the European context the EU-funded project FESTA has compiled and developed international experience from other FOTs, expert knowledge from the evaluation area, etc. into the FESTA Handbook. It provides the guidelines for setting up, performing and analysing an FOT and one important element (that in principle summarises the guidelines) is the FOT Chain (or the FESTA V).

In this report, starting from the original objectives, the FESTA FOT Chain forms the structure of report of progress that has been encountered during the first 30 months of work in TeleFOT. This IP-level deliverable is to some extent focused on what constitutes the unique features of TeleFOT that could have an impact also on a more general level. This is especially important as new FOTs are planned in the area "cooperative driving".

This deliverable starts with a short presentation of what constitutes an FOT, based on historical information but also on the outcome of the FESTA project, i.e. the FESTA Handbook.

The FESTA FOT Chain (the FESTA V) is then used for exploring the evaluation strategy that has been applied for TeleFOT as presented in deliverable *D2.2.1 Testing and Evaluation strategy I*. The unique and novel elements of this strategy are highlighted. The importance of fully understanding the context of the FOTs to be performed in TeleFOT (assessing the impacts of functions provided by nomadic devices) is emphasised. The functions addressed are introduced (cp. *D2.5.1 Functions Specification*). More information about the context is found in the deliverables *D3.3.2 Test communities overview* and *D3.4 .1 FOT Plans*, where the different test communities and the dedicated FOT plans are presented.

An integrated top-down and bottom-up approach (based on relevant impact areas) for generating research questions and related hypotheses was applied. The top-down approach is based on the theoretical background to the impact areas addressed, i.e. Efficiency, Environment, Mobility, Safety and User Uptake.

When it comes to the study design it is important to notice the introduction of the concept of Large Scale FOTs (L-FOTs) and Detailed FOTs (D-FOTs) and their links to naturalistic studies and a more experimental approach. A within-subject design is preferred compared to a between-subject design and an explanation is given. Data acquisition is crucial and apart from logging devices also questionnaires, travel diaries, interviews and focus group discussions are introduced. Reference is made to the deliverable *D2.3.1 Data Specification and Quality* for more details on the efforts related to data collection and to deliverable *D2.2.1 Testing and Evaluation Strategy I* for more information about subjective "sensors" in the form of questionnaires, etc. Measures to make this data collection procedure as automatic as possible have been taken.



## 2 FIELD OPERATIONAL TESTS

In this section the concept of FOT is introduced and reference is made to the FESTA project and its main result, which is the guidelines for setting up, performing and analysing an FOT, i.e. The FESTA Handbook.

### 2.1 Requirements for Field Operational Tests

The European version of Field Operational Test (FOT) was introduced by the EC in the beginning of 2007, when two workshops and an Expert meeting were organised for interested stakeholders to clarify and discuss the FOT idea.

The following direct quotations from the Strategic Research Agenda about FOTs will be of help to understand the TeleFOT scope:

*"... there is still a great need to investigate the behaviour of the user in the real environment when being equipped with the new ICT systems for safety and efficiency as compared to the user's behaviour without the ICT systems. The short and long term effect of the use of such systems is also of great importance to assess as justifications of the systems."*

*"Field Operational Tests (FOTs) have during the later years developed as a powerful tool to gain insight into how new functions and systems suit the user when operated in the real context under sufficiently long time to reach the "daily operational behaviour level"."*

*"FOTs for ICT based traffic safety and efficiency functions and systems are viewed as very important for understanding the users' ability to use these and to make cost/benefit assessments as well as to evaluate the impact on safety, traffic efficiency, environmental aspects and the behaviour of drivers and other road users. FOTs are also important in the technical development since they can provide valuable feedback regarding the performance and improvement of the technical system."*

*"In addition, FOTs will provide the necessary assurance for both the roll-out: technical and commercial feasibility. It is only by associating a business model to a mature technical system that benefits can be delivered to users."*

*".. FOTs can be the basis for broadening the awareness among the public for these ICT systems."*

The EC also set six objectives for FOTs at the Experts meeting held in Brussels 5 June 2007 [Field Operational Tests 7 FP]. They are the following:

1. Validate the effectiveness of ICT based systems for safer, cleaner and more efficient transport in real environment,
2. Analyse driver behaviour and user acceptance,
3. Assess the impact of intelligent safety and efficiency functions using real data,
4. Improve awareness on the potential of intelligent transport systems and create socio economic acceptance,
5. Obtain technical data for system design and product development,
6. Ensure the transferability of the FOT results at national, European and International level.

Furthermore, EC stressed the nature of FOT and proposed that

*"[a] FOT [has the] focus on technically mature ICT systems, including technical, user acceptance, efficiency and deployment aspects. "*

To conclude, based on the above quotations and objectives, the following issues and research questions need to be addressed in FOTs:

- Focus on ICT systems assisting drivers,
- Focus on mature systems and technology,
- Investigate different things such as
  - usability and willingness-to-have aspects (*"how new functions and systems suit the user"*)
  - impacts on behaviour,
  - impacts on safety,
  - impacts on efficiency,
  - impacts on environment (economic / "green" driving),
  - socio-economic impacts,
- Investigate technical and commercial feasibility / business models,
- Study also long-term impacts,
- Use robust scientific methods to study the impacts of ICT systems,

- Raise awareness among the public of the ICT systems in assisting drivers.

## 2.2 The FESTA Handbook

The main purpose of the FESTA 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 FESTA 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 the steps that ideally need to be carried out during an FOT are presented. They are explained in detail in the different chapters of the FESTA Handbook. The FOT Implementation Plan takes up all the steps and integrates them into one big table which can be used as a reference or checklist when actually carrying out an FOT.

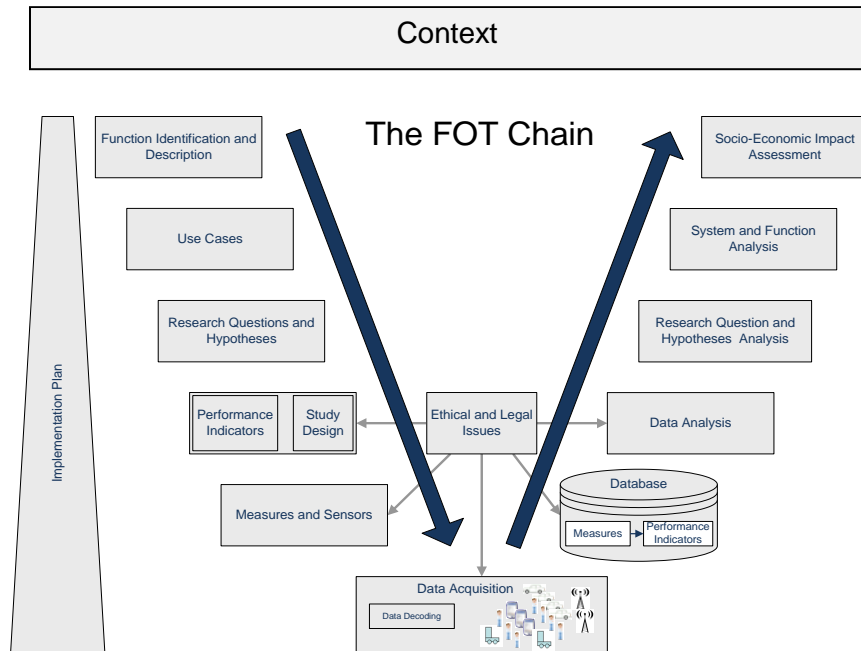


Figure 1. The steps that typically have to be considered when conducting an FOT. The large arrows indicate the time line.

In order to make the picture more complete (compared to the original version) a horizontal bar is added on top of the diagram that in principle summarises the context in which the FOT is supposed to take place. For instance, the choice of a function to be tested implies that there is either a problem that is to be addressed and that the chosen function is defined to solve the problem or that a policy objective is stated and that the function tested can be used to reach the objective. An FOT can always be related to a wider view on the exercise than is defined by just a description of the function to be tested.

This can be summarised as 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, dealing with the more general aspects of an FOT and with aggregation of the results. The further down the FOT Chain V-Shape that 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.

## 3 THE TELEFOT EVALUATION STRATEGY

The FESTA Handbook starts with a general introduction where the approach is visualised by the FOT Chain (Figure 1). In the figure the steps that need to be carried out during an FOT are presented. In relation to this deliverable only the left hand side (and part of the bottom side) of the FOT Chain is covered. The bottom to right hand side, being the topics of data acquisition, data analysis and impact assessment work, is the focus of the last half of the TeleFOT project and is not addressed in this report.

This report is held on a generic (IP) level and tries to disclose what has been learned over the first two and a half project years. When necessary reference is made to other TeleFOT deliverables, where suitable details can be found and further consulted.

### 3.1 BLOCK: The Context

TeleFOT is an example of an FOT, or rather several FOTs. In order to define the key characteristics of the TeleFOT project and its different FOTs, it is necessary to first describe the following key concepts: field test, naturalistic study, experiment, and field operational test. One important deliverable is *D2.2.1 Testing and Evaluation Strategy I*. A summary of the TeleFOT strategy is found in Appendix I.

- By *field tests* is generally meant to test something, e.g. a product, under actual operating conditions or in actual situations reflecting intended use.
- By *naturalistic study* is most often referred to a study where researchers observe and record some behaviour or phenomenon, often over a longer period of time, in its natural setting while interfering as little as possible with the subjects/participants or the phenomena.
- *Experiment* is defined as a test or trial carried out for the purpose of discovering something unknown or of testing a principle, supposition, etc.
- By *controlled experiment* is meant an experiment that isolates the effect of one variable on a system by holding constant all variables but the one under observation.
- A *Field Operational Test* is generally described as a test run under normal operating conditions in the environment typically encountered by the subjects and the equipment being tested. Normally a FOT involves a larger number of users using the systems and services in their daily life in actual use conditions.

The TeleFOT project consists of Large Scale FOTs (L-FOTs) and Detailed FOTs (D-FOTs). L-FOTs are *naturalistic studies* in the sense that they are studies in which will be investigated normal, everyday, use of a set of nomadic and aftermarket devices and different functions. The studies concern conditions in which the participants receive, use and react to functions and services provided to them and data is to be collected over a longer period of time from a larger number of participants. The studies are also *experiments* in the sense that tests are undertaken in order to find out the answers to questions and hypotheses posed. Nevertheless, they are not controlled experiments even though as rigid a test procedure is to be executed.

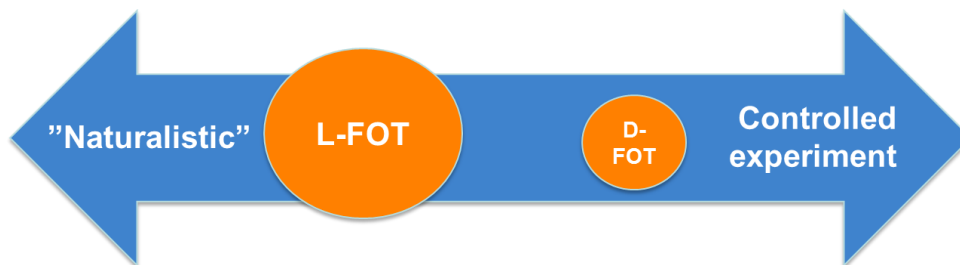


Figure 2 L-FOTs vs. D-FOTs

Also the D-FOTs will be carried out as experiments in the sense that the tests are undertaken in order to find out the answers to questions and hypotheses posed. However, even though not all D-FOTs may not be carried out as completely controlled experiments, the D-FOTs will be run with more control than the L-FOTs, e.g. the participants will be asked to drive certain routes, as well as under certain conditions. Furthermore, less vehicles and less participants will be involved but the vehicles will be equipped with additional equipment so that additional, as well as more detailed, data will be collected (e.g. on acceleration patterns, speed, petrol consumption, etc.) than will be the case in L-FOTs.

However, it is important to stress that the L-FOTs constitute the core of the TeleFOT project. The main purpose of running L-FOTs and D-FOTs across different test sites, and in parallel, is to benefit from the particular strengths of the respective approaches, i.e. the higher ecological validity of the naturalistic driving test results, providing evidence of behaviours and behavioural changes over time, and the higher reliability of the experiment, offering possibilities to identify more causal types of explanations to e.g. the

drivers' behaviour than will L-FOTs. Thus, the D-FOTs will complement the L-FOTs, providing information for the analysis and interpretation of the results.

### 3.1.1 TeleFOT boundaries

Field Operational Tests, (FOTs), have been carried out in Europe, in Japan, in Australia, and in the U.S. Even though the general objective of these types of tests can be described as testing new transport technologies in general, a common focus of previous projects has been in-car use of driver assistance systems, in particular different safety systems.

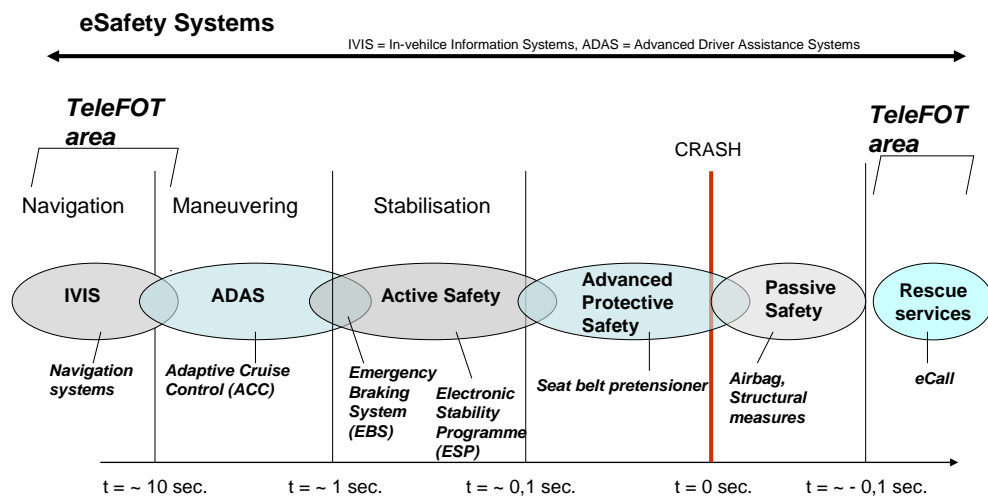
TeleFOT is to assess the impacts of functions provided by different aftermarket and nomadic devices, including future interactive traffic services. As stated in the Description of Work (DoW) the project is to test "... driver support functions...", but also that "...mobile services may have impacts on the whole travel chain – not only on driving". In contrast to several of the former FOTs, which more often involved a more narrow focus and a system boundary defined by driver-vehicle-road, the TeleFOT project system boundary is the larger context of traffic, transport and travelling, and in addition the user in different roles, i.e. as driver, passenger, as well as traveller.

The project will further address the use of functions and services in relation to pre-trip (e.g. the use of functions for planning a trip), during trip (e.g. the use of functions to re-plan a trip or to help way-finding), as well as post-trip (the use of a function, e.g. green-driving support, to assess the outcome of a trip). This means that data will be collected beyond that of in-vehicle use and that it is necessary to collect information also by means of e.g. travel diaries and questionnaires in addition to the data gathered by different logging devices.

Furthermore, earlier FOTs have often had a focus on Safety related to driving. Also in TeleFOT one of the main impact areas concerns driver behaviour and safety but other areas are also to be equally considered. These include: User Uptake (i.e. issues related to acceptance and adoption of new technology and new behaviours), Efficiency Environment ("green driving") and Mobility, as well as socio-economic impacts on the transport system as a whole. In TeleFOT, all L-FOTs are to address all impact areas whereas D-FOTs will have a more narrow focus. For instance is Mobility considered difficult to address in a D-FOT.

### 3.1.2 Functions Specification

The TeleFOT application areas cover all information coming from beyond the driver's visual field, i.e. typical cooperation driving application areas. In addition to that, the project focus is on navigation type of tasks and providing personalised speed management and "Green driving" support, where the benefits can be seen in the manoeuvring level, and in the appropriate choice of speed such as speed limit information or speed alert.



**Figure 3 TeleFOT application area in the field of Intelligent Vehicle safety Systems.**

The project concept relates closely to cooperative driving (I2V – V2I) , since aftermarket and nomadic devices used in vehicles receive the information contents from background systems through traffic information databases, GPS, telecom operators and service providers. For this reason, the project provides an excellent opportunity to test functions and services foreseen in future cooperative driving systems. The project also has close links to existing projects using traffic information for better safety and ecological driving.

Six main functions were addressed; Traffic Information (TI), Speed Limit Information (SLI), Speed Alert (SA), Navigation Support (static) (NSS), Navigation Support (dynamic) (NSD) and Green Driving Support (GDS). As combinations of functions will be tested as well, such combinations have to be identified and specified as one entity.

The eCall function will also be tested but the arrangements cannot follow the FOT set-up procedures as proposed by FESTA and is laid out in the TeleFOT Deliverable D2.2.1,



Testing and Evaluation Strategy I. A specific dedicated approach will be applied in due course.

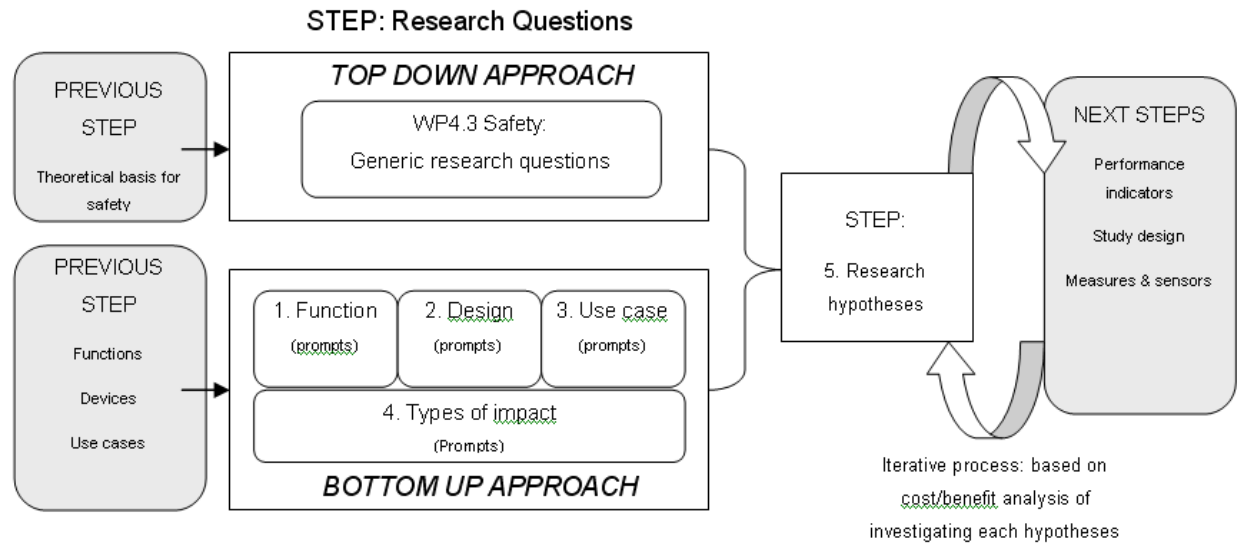
The process for the specification of functions started with a description of the functions to be tested and was presented as a generic *Functions specification* for the six main functions addressed. Based on reports from the test site managers, the specifics of every function being implemented in a local FOT has now been clarified. It is obvious that the implementations of each FOT are unique (if they are not using exactly the same device and function provider) with minor but (possibly) important differences in the features made available to the driver (user).

Finally, the combination of functions now being implemented in the different FOTs are addressed and a tentative classification of critical and not so critical combinations is presented as an input to the impact evaluation work to be performed under the umbrella of SP4

### **3.2 BLOCK: Research questions and hypotheses**

Within the context of TeleFOT, it is important that the number of hypotheses to be tested is kept to the minimum whilst ensuring that all impact areas are covered. The studies must be designed to enable the hypotheses to be tested.

According to the FESTA Handbook the formulation of use cases provides the main basis for generating research questions and hypotheses to be tested in a FOT. This approach has however, proven insufficient for the TeleFOT project with its many functions and combinations of functions and given that not all functions were determined.



**Figure 4** Integration of top-down and bottom-up approaches for hypotheses generation for all impact areas (safety is referred to in the figure).

Therefore, a generic modified process for generating hypotheses for different FOTs has been developed in TeleFOT (Figure 4). An integrated Top-down and Bottom-up approach will be used to firstly generate the research questions and subsequently formulate the hypotheses. This integrated method ensures that the hypotheses have a foundation in both the theoretical aspects of the impact area under consideration (Top-Down) and in the system functionality (Bottom-Up considering both intended and unintended effects). This in turn helps ensure that all the potential impacts are identified. This process is also one important step for ensuring that the number of hypotheses that will be addressed across the whole of TeleFOT is kept to a manageable level. These approaches are explained in more detail in the following sections.

### 3.2.1 The Top-down approach

The basic principle for generating hypotheses using a top-down approach lies in a theoretical understanding of the factors that influence the different impact areas. It should be noted that there is likely to be overlaps of these factors among the impact areas under consideration and hence the same research question and resulting hypotheses will be applicable across more than one impact area. The approach will result in *generic research questions* that are independent of the any system functionality (and can thus be used also in other FOTs and for other, additional devices and functions than those initially tested in TeleFOT).

The procedure for generating hypotheses in a top-down approach is as follows;

- The impact area should be considered in its entire context and primary measures affecting that area identified.
- Secondary factors of these measures are then identified that can be used to explain the variations in the primary measures.
- Finally the variables affecting the secondary measures are identified.
- The variables identified form the basis of the generic research questions *‘‘Is there a change in the variable?’’* and the hypothesis based upon an anticipated effect of the variable *‘The variable will increase/decrease.’*

This procedure should be undertaken for each of the five impact areas.

Using the Safety Impact as an example;

- The primary measures affecting safety would be the ‘Number of events (accidents, near misses) that occur’ and the ‘Severity of the event’.
- Secondary factors affecting the first of these measures would, for example be ‘Exposure of the vehicle on the road’, ‘The driving style of the driver’, ‘The distraction of the driver from the driving task’ and ‘Any interaction with the fitted device’.
- Considering the factor ‘Exposure’, this can be measured with the following variables: ‘Length of journey’, ‘Number of trips undertaken’ and ‘Road type used’.
- These variables lead to the following research questions
  - *Does the system affect the length (miles) of journeys?*
  - *Does the system affect the duration (hours) of journeys?*
  - *Does the system affect the number of journeys undertaken?*
  - *Does the system affect the road type used?*

This leads to the generic hypotheses that can be tested in a statistical manner. The direction each hypothesis should take (e.g. increase or decrease) is based upon the anticipated effect once the top-down approach is integrated with the bottom-up (system defined) approach.

- *Journey lengths will increase/decrease when the system is used compared to when it is not used.*

- *Journey duration will increase/decrease when the system is used compared to when it is not used.*
- *The number of journeys will increase/decrease when the system is used compared to when it is not used.*
- *The use of rural roads/motorways/major roads will increase/decrease when the system is used compared to when it is not used.*

For TeleFOT, generic research questions and hypotheses for the impact areas Efficiency, Environment (“green” driving), Mobility, Safety, and User Uptake have been generated.

### 3.2.2 The Bottom-up approach

The bottom-up approach can only be undertaken once the functions, devices and use cases have been defined for the particular test site. For each impact assessment, the detailed research questions were developed based on a consideration of four sets of factors.

1. **Function:** A functional description of the system (i.e. what it does) and the effect that the function may have on a user in the context of each Assessment domain.
2. **Design:** The implementation of the system (i.e. how it is designed), and the impact the design attributes have on the user–system interaction in the context of each Assessment domain
3. **Use Case:** The use cases (i.e. the context of use factors) and their relationship with consequences of use within the real world
4. **Types of Impact:** The types of impact that are being considered.

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. This process leads to the development of hypotheses concerning specific scenarios. These scenarios are derived from the combination of Use Cases and Situations (TeleFOT Internal Report IR2.2.1).

A huge number of research questions and associated hypotheses from the top-down and the bottom-up approaches were developed. A key task has been to integrate both sets of hypotheses in the context of each FOT. The bottom-up approach formed the basis of the hypotheses list for a FOT and that the top-down approach was used in order to check that nothing significant for a particular impact area has been omitted.

After the integration has taken place, the list of hypotheses was likely to be large. In order to derive a final, manageable set of research questions and hypotheses that can be applied throughout the various test sites, a cost-benefit approach was proposed and used. An assessment was made regarding the likely "costs" of collecting the data. Costs can be represented in terms of effort required to derive a performance indicator expressed predominantly in terms of resources. This should be offset against the likely "benefit" that proving/disproving the hypotheses will have. This is measured by way of the likely contribution towards providing a significant answer the research question and thus the level of contribution to the impact assessment. To some degree, this depends upon the stakeholder needs and requirements, and therefore a prioritisation of their 'needs' was considered.

### 3.2.3 Common research questions and hypotheses

A common set of research questions and hypotheses are to be addressed in TeleFOT FOTs. A tentative list of common research questions and hypotheses has been generated based upon the top-down and bottom-up approaches, and by taking into account their feasibility in D-FOTs and L-FOTs, and they are found in Appendix I. Consideration of the 'feasibility' of addressing specific research questions and hypotheses within TeleFOT was a particularly important step in reducing the number of hypotheses from the starting point. For example, 'Fatigue' is known to be an important research area and research hypotheses were developed to address this in the context of After-market and Nomadic Devices. However, when consideration was given to the feasibility of measuring Fatigue in an FOT, it was clear that it could not be satisfactorily handled within neither D-FOT nor L-FOT and so the research hypotheses relating to this issue were discarded.

However, from the tentative list can be concluded that one specific research question may result in several different hypotheses depending upon the function(s) to be tested, thus the top-down approach must be complemented by the bottom-up approach as described above. It can also be concluded that one and the same hypothesis can, given the anticipated impact of the function to be tested, have a positive or a negative direction. Finally, it can be assumed that not all aspects would have been covered in such a systematic way, across all functions, without the top-down approach.

### 3.2.4 Some recommendations

A structured approach should be applied linking a **top-down** approach at the global system level with a **bottom-up approach** which looks more at system states and what can arise from them. FESTA considers it mandatory to combine the two approaches.

A multidisciplinary team should *jointly* develop the hypotheses. A workshop at which participants can brainstorm and debate is recommended to achieve this. Participants in the process should include design engineers, traffic engineers and behavioural scientists, ideally including both behavioural psychologists and human factors experts. The process should iterate between the top-down and bottom-up approaches. It is not particularly important which is performed first, but it is important to cross-check one approach by using the other.

### 3.3 BLOCK: Performance indicators and measures

After defining the hypotheses, appropriate performance indicators (PIs) must be chosen to answer the hypotheses (Figure 5). Performance indicators are by definition quantitative or qualitative measurements, expressed as a percentage, index, rate or other value, which is monitored at regular or irregular intervals during the test and which can be compared with one or more criteria (cf. FESTA Handbook 2008).

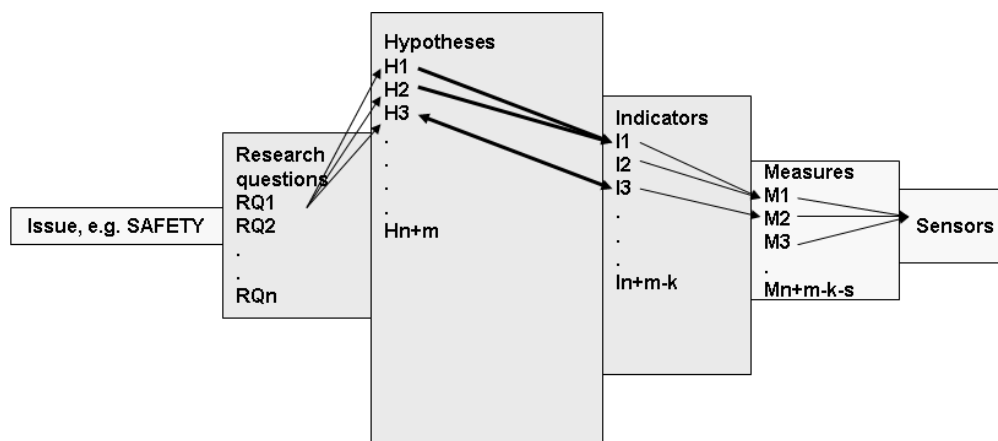


Figure 5 Description of procedure, from impact area and research questions to measures.

Appropriate performance indicators will be chosen for each hypothesis and combination of hypotheses. These performance indicators are described in the internal report IR2.2.2 Research Questions and Hypotheses. In determining the indicators the list provided in

the FESTA handbook will be used as a first option. The list serves a check list also for the coverage of the hypotheses.

The number of measurements and indicators will be kept to a minimum. However, the same performance indicator may well serve to clarify a number of hypotheses and the impacts in a number of impact areas. It is important to go through systematically all possible research questions, hypotheses and indicators to identify the most valid indicators. The final set of indicators will be chosen based on possibilities to complete the measurements, the project budget and other limitations.

### **3.4 BLOCK: Study design**

Each FOT should be designed so as to enable the chosen hypotheses to be tested in a statistically rigorous manner. Since many of the hypotheses aim to show the effect the introduction of a 'treatment' (e.g. the use of a Navigation System) it is essential that *base-line* or *control data* is collected along with the experimental data so that comparisons can be made between user response when, for example, the Navigation System is in use compared to when it is not.

There are two general study designs that can be used to generate data in both control and experimental conditions; Within Subject Design and Between Subject Design.

#### 3.4.1 Within Subject Design

In a Within Subject Design, each participant undertakes a period of time in both the control condition and the experimental condition. For example in a FOT on driving without and with a navigation system, every participant drives for some time without (control condition) and for some time with the navigation system (experimental condition).

This type of design has two main advantages:

- (i) fewer participants are needed than in the case of a Between Subject Design, and
- (ii) it is more likely to find a significant effect, given the effects are real since there is no variance in the characteristics of the subjects in the experimental and control groups.

However, there are also disadvantages. There is a risk for so called carry-over effects, which means that if a participant experiences one condition, this may affect his/her

driving in the other condition. For example, in a D-FOT – where the control and experimental data are both collected for a pre-determined route – the first measurements taken may influence the second since people gain familiarity with a specific road. Furthermore, there should be a sufficient amount of time (and therefore exposure) allocated to the control condition and the experimental condition in order to experience similar circumstances in both conditions (e.g. weather, lighting, traffic density etc.) but also allowing for adequate experimental data to be collected.

#### 3.4.2 Between Subject Design

In a Between Subject Design each participant takes part in either the control condition (control group) or the experimental condition (experimental group). The groups should then be equal in size. Thus, for the same amount of data, twice as many participants are required in a between subject design than for a within subject design. The number increases further if more than one treatment is under consideration.

The advantage of a between subject design is that carry over effects are not a problem as individuals are measured only once in every condition; one measurement is completely independent of the other. Furthermore, the duration of the FOT is shorter for a between subject design than a within subject design since both the control group and the experimental group can run in parallel.

#### 3.4.3 Implementation

In Large Scale FOTs, the drivers are in principle using their own vehicles in their daily travel. These vehicles are equipped with testing (aftermarket/nomadic) and recording devices measuring speed, position and some vehicle dynamics. These constitute the core of testing. The proportion and number of subjects using different devices are detailed in the test and evaluation strategy and reported in deliverables *D3.3.2 Test communities overview* and *D3.4 .1 FOT Plans*.

The Detailed FOTs complements large-scale FOTs in that they are more detailed in terms of behavioural parameters measured. It is intended that subjects use dedicated test vehicles during tests. The first idea was to recruit the subjects from among those involved in the large-scale tests. It is necessary to carry out detailed tests across Europe, since cultural differences explaining driver behaviour and reactions have been indicated to be significant by earlier European behavioural studies and accident statistics.



To conclude: two approaches were defined for the road tests: (i) Large-scale FOTs subjects use their own cars and devices. Secondly, (ii) samples of subjects will be chosen for more precise testing in Detailed FOTs and here experimental, well-equipped test cars are used.

Concerning the differences in experimental design between the large-scale and detailed testing, it can be stated that large-scale testing does not allow strict observance of experimental design principles. It concerns especially conditions in which the drivers receive, use and react to functions and services provided to them. The subjects are using the services in the daily travel and these conditions cannot be manipulated by making them use the same routes at the same time of day.

This stricter adherence to the principles of experimental design is reserved to detailed testing where subjects are instructed to follow given tests routes and react according to given instructions. In this way, detailed testing enables deeper, causal type of explanation of the behaviour than large-scale testing. Detailed tests are complementing large-scale tests.

#### 3.4.4 Participants

A random sampling procedure is often considered the most desirable method to select the participants for a study. In the case of the TeleFOT FOTs, this is not feasible and other, non-randomised methods must be used. Nevertheless, a random procedure could be used in order to assign participants to the corresponding experimental condition (control group or experimental group) specified by the particular experimental design in each FOT.

A fundamental criterion for choosing participants is that the participants reflect the intended user population of the tested system/function. A characterisation of the intended user population is therefore a first step in the choice of participants. A second criterion is that the participants allow for the chosen hypotheses to be tested. A stratified sampling method may be applied. Depending upon the research questions and the functions to be tested, there could be a need to select certain groups of participants taking into account some particular characteristics in such a way that these variables will act as covariates in the analysis in order to study differences between groups. Thus, the

variables to be considered should define exhaustive and mutually exclusive groups, i.e. a participant should be assigned to a group and to one group only.

#### *Type of participants*

As previously mentioned, the participants in the FOTs should be representative of the intended user population of the systems and functions tested. The intended user population could be described on the basis of different characteristics, e.g.:

- *demographical variables* such as age, gender, socio-economic status, etc.;
- *experience*, such as driving experience (total kilometres driven, kilometres during last year, years of driving license, usual types of roads driven, etc.), experience of technology (embedded systems, nomadic and after-market devices, etc.), and experience of functions (navigation support, speed alert, etc.);
- *personality and attitudes*, considering aspects such as sensation seeking, locus of control and/or attitudes towards road safety issues (new technology, public transport, etc.).

However, on the one hand and depending upon the chosen research questions, there may be a need to select a particular group of participants for inclusion in the FOTs, ensuring that this group is in some way representative of those drivers who will ultimately interact with the system. On the other hand, considering too many factors in the choice of participants will increase the requirement for the sample size.

Based on the use case descriptions, the basic criteria for the selection of participants are:

- (as a general approach) the participants should be between 25 and 65 years old. Nevertheless, each FOT should select the age range of the participants considering the age allocation of the driving population within each country;
- the participants should have more than 3 years driving experience; and
- the participants should drive more than 10 000 km/year.

#### *Number of participants*

FOTs are described as studies in which a large number of individuals participate. In TeleFOT a difference has been made between D-FOTs and L-FOTs with implications for the number of participants to be involved. A test involving e.g. 10-12 participants will not be regarded as a L-FOT whereas a test involving, e.g. 100 participants may. On the other hand, 10-12 participants may suffice in a D-FOT.

The required number of participants in an FOT will always depend upon a number of factors, e.g. the number of functions and/or systems to be tested, the hypotheses formulated, the choice of a between or a within subjects design, etc. If the number of participants is small, it is difficult to statistically prove any effects of the function/system that are actually there whereas a large number of participants increases the chance of finding an effect. On the other hand, a large sample implies a higher investment in terms of equipment, resources, etc.

In order to ensure that the chosen sample size is representative for the behaviour of a group of users and that it is possible to statistically prove any effects that are there, a power analysis is needed to calculate the desirable sample size. A statistical power analysis exploits the relationships between the four variables involved in statistical inference: sample size, significance criterion, population effect size, and statistical power.

#### 3.4.5 Study environment

The study environment is a critical element in all FOTs, since it will determine the data that is collected and the ability to fulfil the objectives and test the hypotheses. Such environmental conditions include:

- geographical location, e.g. types of roads, traffic patterns, infrastructure and communication issues, transportation options;
- weather conditions, e.g. sun, fog, rain, snow;
- time of day; and
- seasonal effects, including e.g. the seasonal opening/closure of roads

The fundamental idea behind the TeleFOT project is that the devices and functions to be evaluated should be used in everyday life and during as natural use conditions as possible. Therefore, designing an L-FOT that *requests* the participants to drive certain roads or drive during certain weather conditions contradicts the overall idea of FOTs, i.e. real-life usage of functions and services. Research questions and hypotheses that require very specific environmental conditions in order to be tested, e.g. very particular weather conditions or very particular types of roads, should be consequently avoided in the L-FOTs whereas they may be applicable in the D-FOTs.

However, if the research questions and hypotheses chosen require specific environmental conditions in order to be tested also in L-FOTs, it is recommended this should be ensured

by other means. The design of the study may still increase the likelihood of the desired environmental conditions to occur. One such study design issue is the choice of participants. One could choose participants according to driving patterns, i.e. choose participants who often drive long distances on motorways as opposed to participants who drive predominantly in the city centre or choose participants according to geographical location, i.e. including participants who live in city centres as opposed to participants who live in rural areas. Also the choice of time of year is a study design issue, i.e. choosing to run the L-FOT during the winter months as opposed to the summer months (only) will increase the probability of slippery and icy roads.

Another feasible approach is to use post-trial categorisation of key variables to enable exploration of data, e.g. an analysis of the effect of a particular function during short as opposed to longer trips or during sunny as opposed to rainy weather conditions. This requires, however, that relevant issues are identified beforehand, and that data is captured to enable matching with hypotheses and scenarios of interest.

### **3.5 BLOCK: Data acquisition**

In TeleFOT, a first common set of data is to be collected across test sites and FOTs in order to ensure a cross-site analysis. With this approach the European dimension of the project can be addressed and the results of the separate FOTs can be compared across the different test sites. In addition, a second set of common data will be collected across those L-FOTs and D-FOTs that evaluate the same type of functions and/or devices while a third set of data consists of data that is specific for each site. This will be collected per site only.

However, overall the data collected should be kept as limited as possible, be restricted to a minimum. At the same time the data have to be sufficient for the evaluation of the defined hypotheses, both common and specific to the functions and test sites.

In general, data collection is to be carried out at three points of times: pre- test, during test, and post-test. Pre-test data will be collected in advance to the tests. This data contains information about the background of the test participants, their respective driving experience and attitudes towards the tested systems as well as experience of different nomadic devices and different functions/ services.

The main part of the data will be collected during the test, i.e. during control/baseline conditions and during experimental conditions. The data to be collected includes objective data, as well as subjective data by means of several different types of methods and tools. In order to ensure that the research questions are answered and the chosen hypotheses can be tested, subjective data as well as objective data have to be collected in both L-FOTs and D-FOTs.

### 3.5.1 Data collection

In both the D-FOTs and L-FOTs, data will be collected by means of different data collection methods. Overall, a methodological triangulation approach will be applied in TeleFOT. By triangulation is meant the application and combination of several research methodologies in the study of the same phenomenon. Methodological triangulation involves using more than one method and may consist of within method or between method strategies.

As already stated, data to be collected involve subjective as well as objective data. Subjective data can be retrieved from, e.g. participants' descriptions of events and/or assessments of and preferences for features. Objective data can be retrieved from, e.g. recording of participants' driving behaviour and the logging of speed, acceleration, etc. as described in the former sections, but also from observations.

### 3.5.2 Subjective data

Data collection methods that could be used for the collection of subjective data include questionnaires, individual interviews, and focus group interviews.

#### *Questionnaires*

A questionnaire is a research instrument consisting of a series of questions for the purpose of gathering information from respondents. Most often the questions, their arrangement, as well as their answers are determined in advanced. The particular benefits of questionnaires are that they are an efficient way to gather data from a large number of respondents, and that the same study can be repeated several times over across a longer time period.

### *Interviews*

An interview is a conversation between two or more people (the interviewer and the interviewee) where questions are asked by the interviewer to obtain information from the interviewee. A personal interview may be carried out in a very structured manner (involving predefined questions, predefined arrangement of questions, etc.) or in a semi-structured (e.g. carried out with the support of an interview guide) or even an unstructured manner.

### *Focus group interviews*

Focus group interviews are an alternative to individual interviews. A focus group interview is a particular kind of group interview supported by a moderator. The particular strengths of focus group interviews (compared to individual interviews) are that the focus group participants can reflect not only upon the questions posed but also react to and provoke one another. Through unification across the participants, common denominators become visible and through polarisation, possible segmentation and hence differences become apparent.

### 3.5.3 Objective data

Data collection methods to be used for the collection of objective data include measurements by logging but also activity-based diaries (or travel diaries) and observations. Again, a difference has to be made between the L-FOTs and the D-FOTs due to the limited access to data in the vehicles involved in the L-FOTs in comparison to the instrumented vehicles of the D-FOTs.

### *Logging*

For the collection of objective data obtained by different sensors different types of devices could be used. When collecting information on e.g. travelled distance, speed, etc., data can be collected by means of internal vehicle bus data (e.g. CAN), by a separate logging device inside the specially equipped vehicles. This kind of data logging will primarily be accomplished in the D-FOTs with less than 20 vehicles.

In those FOTs, where there will be no access to a vehicle bus (as the regular systems of the driver or the participating company are used), other loggers (e.g. Metasystems or Broadbit) have to be used and no access to the vehicle bus is available. These loggers are connected to sensors, nomadic devices or aftermarket devices, e.g. via Bluetooth.

Inside the logger a GPS receiver, for the estimation of GPS Position and Time, is contained and can be equipped with other additional sensors, like accelerometers. All loggers have to be matched to the predefined Performance Indicators, so that the needed data is available for the following evaluation subsequent to the tests.

## 4 PILOT TESTS

Conducting a pilot study is necessary to prepare for the deployment of the different FOTs, whether D-FOTs or L-FOTs, and to support the design of the relevant tools for the evaluation process (See FESTA Handbook). It represents a relevant step for the mobilisation and dialogue between the various teams involved in the FOTs, hereby promoting a common framework and consensus for the evaluation process.

The TeleFOT project is characterised by a mix of, so called, Detailed and Large Scale FOTs. The L-FOT are to be run as naturalistic driving tests (to the greatest extent possible), with a large number of vehicles and participants taking part in the test over a longer period of time, while the D-FOTs are to be completed as 'controlled experiment'. The differences have implications for study design, for the data that can and will be collected, and by what methods.

Consistent across D-FOTs and L-FOTs, however, a study design involving baseline/control conditions compared with experimental conditions will be applied, both subjective and objective data will be collected in order to address a set of common research questions and hypotheses, and data will be collected pre-test, during, and post-test. Legal and ethical matters, including privacy and security, have to be addressed in the same rigorous manner way across all test sites.

The pilot tests in TeleFOT performed have addressed the following levels of analysis:

- The first level has to check the technical function of the data collection systems in actual driving situations. They should enable to identify potential problems of sensor calibration and thus to establish the periodicity of maintenance procedures during the FOT, and to validate the data collection procedure (from data acquisition, data transmission to data storage). The technical teams involved in the FOT were in charge of these field tests and the first pilot tests were executed with the full set of objective data.



- The second level consisted of testing the feasibility of the overall evaluation process – from the selection of participants through to data collection. It was, thus, a final rehearsal before the deployment of the actual FOT. It enabled 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.

## 5 DISCUSSION

This report is an effort to disclose in what way and how far the project TeleFOT has come to fulfil the scientific and technological objectives that were stated from the very start. The structure of the Work Packages did not directly coincide with the structure of the objectives and therefore this discussion on how the results of the TeleFOT project meet these objectives is included. It should be noted that this report is related to the left-hand and part of the bottom side of the FESTA FOT Chain. In principle this report states the situation when being about half way through the project. The scientific and technological objectives are repeated in the following text, one by one, with a following section where the outcome of TeleFOT related to the specific objective is discussed.

*Build, mobilise and integrate European test communities for long term testing and assessment of driver support functions through aftermarket and nomadic devices*

This objective has been fulfilled to an acceptable level. There are three test communities; a Northern (Finland, Sweden), a central (Germany, France, UK), and a Southern (Spain, Italy, Greece) where 18 separate FOTs are performed. In each community (or country) an organisational framework has been established for setting up the test, selecting participants, organising data collection, etc. The assessment procedures are still not settled to the full extent, the assessment work not yet being the main focus of the TeleFOT activities. However, the pilot tests performed have all shown that (with some minor modifications) the FOTs now are operational also in relation to data analysis and the impact assessment work. More details are found in the SP3 Deliverables D3.3.2 and D3.4.1.

*Create a methodological framework for executing the tests and analysing the data*

This objective has also been met at an acceptable level, and a strategy for testing and evaluation has been presented in detail in the deliverable D2.2.1 Testing and Evaluation Strategy I. The starting point has been the FESTA Handbook and it has been followed quite well as far as has been possible, i.e. in relation to the FESTA FOT Chain only the left-hand (and part of the bottom) side have now been made operational. The right-hand (and part of the bottom) side of the FOT Chain has in a limited way been tested in the

pilot tests just performed, and will find their full implementation during the remainder of the project. However, it is important to highlight some aspects that have been developed further in order to fit the specifics of the TeleFOT project.

In contrast to several of the former FOTs, with a most often more narrow focus and a system boundary defined by driver–vehicle–road, the TeleFOT project system boundary is the larger context of traffic, transport and travelling, and in addition the user in different roles, i.e. as traveller, passenger, as well as driver. The project will further address the use of functions and services in relation to pre-trip (e.g. the use of functions for planning a trip), during trip (e.g. the use of functions to re-plan a trip or to help way-finding), as well as post-trip (the use of a function, e.g. green-driving support, to assess the outcome of a trip). This means that data will be collected beyond that of in-vehicle use and the necessity to collect information also by means of e.g. travel diaries and questionnaires in addition to the data gathered by different logging devices.

As part of the work on the strategy for the TeleFOT project, a generic and modified (compared to the FESTA Handbook) process for generating hypotheses for different FOTs has been developed. For the TeleFOT project an integrated Top-down and Bottom-up approach has been used to firstly generate the research questions and subsequently formulate the hypotheses. This integrated method ensures that the hypotheses have a foundation in both the theoretical aspects of the impact area under consideration (Top-Down) and in the system functionality (Bottom-Up considering both intended and unintended effects). This in turn helps ensure that all the potential impacts are identified.

<i>Study aftermarket and nomadic devices in different technical contexts</i>
--

Every test site has its own characteristics and in that sense the realisation of the functions and their accessibility differs quite substantially. However, still no specific analysis has been made about what these differences are and in what way they might influence the results on the European level. There is a Work Package (starting in project month 36) dealing with impacts on the transport systems as a whole (WP4.9) where such problems could be addressed.

*Study different levels of impacts on drivers and society*

Earlier FOTs have often had a focus on Safety related to driving. In TeleFOT one of the main impact areas concerns driver behaviour and Safety but other areas are also to be considered. These include: User Uptake (i.e. issues related to acceptance and adoption of new technology and new behaviours), Efficiency, "Environment/Green driving" and Mobility, as well as socio-economic impacts on the transport system as a whole. In TeleFOT, all L-FOTs are to address all impact areas whereas D-FOTs will have a more narrow focus. For instance is Mobility considered difficult to address in a D-FOT.

Each impact area has been addressed and (as part of SP4) *Data analysis plans* have been launched, one per impact area. This topic is not further dealt with in this report.

*Focus on functions and services for safe, efficient and economical travel*

Six main functions are addressed in TeleFOT; Traffic Information (TI), Speed Limit Information (SLI), Speed Alert (SA), Navigation Support (static) (NSS), Navigation Support (dynamic) (NSD), and Green Driving Support (GDS). As combinations of functions will be tested as well, such combinations have to be identified and specified as one entity (cp. Deliverable D2.5.1). The eCall function will also be tested but the arrangements cannot follow the FOT set-up procedures as proposed by FESTA and is laid out in the TeleFOT Deliverable *D2.2.1 Testing and Evaluation Strategy I*. A specific dedicated approach will be applied in due course.

There is no function that directly addresses Safety as its main goal. However, the safe use of the functions is of course of interest and some of the functions tested have an indirect safety implication, e.g. SLI and SA. The use of TI will of course in many cases be related to the Efficiency of travel as well as to Mobility. Furthermore, the Environmental aspects are certainly directly related to GDS, but close links to SLI and SA also have an indirect impact on the environment.

As a conclusion, the functions tested in TeleFOT cover the areas mentioned in the objective, but also other aspects like Mobility, Environment and User Up-take are

covered. The economical side of the coin will be covered in forth-coming work on operational business models (Task 2.2.3).

*Investigate the contents of functions provided for cooperative driver support*

This topic has not yet been addressed. However, making use of nomadic devices as the platform for functions related to traffic, transport and travel can be considered as a cooperative system in itself. The functions make use of infrastructure installations (e.g. mobile networks, GNSS-based systems, etc.) that can be seen as a cheap and simple approach towards cooperative driving. What remain are the dedicated tests of the introduction of vehicle-related functions in the palette of functions made available via a mobile platform.

*Develop effective procedures of enhancing awareness and take-up of driver support ICT systems among the public*

This topic will be developed further in the last part of the project. There is a close link to the User uptake assessment work in WP4.7. Also the future work on Operational business models and impacts of the transport system as a whole will be used as inputs.

*Focus also on aspects in the use of aftermarket and nomadic devices that may decrease safety*

This specific aspect has been covered by the early studies made at the ADAC laboratories in Germany. It also coincides with studies made in relation to the ESoP on HMI, where the issue of how the nomadic devices are mounted in the vehicle are addressed. It will probably be followed up by the Usability Benchmarking studies to be performed at the end of the project.

## 6 CONCLUDING REMARKS

The discussion in the preceding section has revealed that many of the scientific and technological objectives have been met.

As these objectives are supposed to reflect the overall objectives it is recognised that the first overall objective

*"to assess the impacts of aftermarket and nomadic devices used in vehicles for driver support"*

has been covered to quite a high degree. There is of course still work to be performed, especially related to the data analyses in general as well as to identify in what way the European dimension in the work will be guaranteed.

When it comes to the second overall objective

*"to raise the awareness of the functions and potential that these devices offer"*

there is quite a substantial need to clarify and determine what would be the topics to be addressed during the remainder of the project.

This IP-level deliverable has been written with the ambition to, on a general level, highlight the findings in TeleFOT that are unique and that provide new insights into the FOT community. Such work is for the moment in progress on the European level as a revision of the FESTA Handbook is on its way. This work is led by the project FOT-NeT 2 and the findings made evident in this report are fed into that work for inclusion if the revised version of the FESTA Handbook.

## 7 REFERENCES

Reports from the FESTA project available at  
<http://www.its.leeds.ac.uk/festa/downloads.php>

FESTA Handbook v2.pdf

FESTA D2.5 FOT Implementation Plan.pdf

FESTA D2.4 Data analysis and Modeling.pdf

FESTA D5 Nomadic Devices Final.pdf

FESTA D2.1 PI Matrix.xls

FESTA D4 Cooperative Systems Final.pdf

FESTA D3 Vehicle systems Final.pdf

Reports from the TeleFOT project available at  
<http://www.telefot.eu>

TeleFOT D2.2.1 Testing and Evaluation Strategy I

TeleFOT IR2.2.1 Use Cases

TeleFOT IR2.2.2 Research Questions and Hypotheses

TeleFOT IR 2.2.3 Experimental Design

TeleFOT D2.3.1 Data Specification and Quality

TeleFOT D2.5.1 Functions Specification

TeleFOT D4.3.1 Data Analysis Plan – Safety

TeleFOT D4.4.1 Data Analysis Plan – Mobility

TeleFOT D4.5.1 Data Analysis Plan – Efficiency

TeleFOT D4.6.1 Data Analysis Plan – Environment

TeleFOT D4.7.1 Data Analysis Plan – User uptake

## **APPENDIX I**

**A list of common research questions and hypotheses, generated on basis of a top-down and bottom-up approach.**



\* Apply to almost all hypotheses and is therefore not repeated. (GDS=Green Driving Support, NAV=Navigation Support, SA=Speed Alert/Speed Limit information, TI=Traffic Information)

Research question	Hypothesis	Functions			
		GDS	NAV	SA	TI
Is the travel time(s) from origin to destination affected ?	Travel times will increase due to lower speeds (with access to function compared to without).*			x	
	Travel times will decrease because the system provides information on fastest routes.		x		
	Travel time will decrease for regular trips because the driver chooses smoother routes or timing based on the information provided.				x
	Time spent travelling will increase because speed is decreased due to traffic related warnings.				x
Are any delays avoided ?	There will be less delays because the driver is informed about incidents causing delays.				x
	There will be less delays because the system supports the driver to find a more efficient route.		x		
Are any traffic jams avoided ?	There will be less exposure to traffic jams because the driver chooses an alternative route to avoid the incident(s).				x
	There will be less exposure to traffic jams because the system reroutes the driver to other routes in case of such an incident en route.		x		

Is speed affected ? Is average speed affected?	There will be a decrease/increase in speed because the driver is warned about incidents ahead. (The highest speeds are cut).				x
	The variation in speed will increase because drivers react differently to the information.			x	
	There will be a decrease in speed because the driver is warned about exceeding the speed limit.				
	Speed limit violations will decrease because the system reminds/warns the driver.			x	
	The variance in speed will decrease because the driver obeys speeds limit better.			x	
Is speed homogeneity affected?	Speed homogeneity will increase.	(x)			x
Are speeds of single vehicles more harmonized?	Drivers will choose more constant speed in order to avoid extra fuel consumption (or to minimize consumption)	x			
Is there a change in target speed?	There will be a decrease in target speed to be upheld because of access to information on ...				x

Is acceleration affected ?	There will be a decrease in incidents of sudden acceleration/deceleration.	x			(x)
Is very sudden/ heavy acceleration affected ?	There will be a decrease in harsh/sudden braking because the driver anticipates the situation due to warnings.				x
Is gear choice /RPM affected?	Higher gears will be used to a higher extent which leads to lower rpm. (and hence ... )	x			
Is the total amount of fuel used affected?	There will be an increase in total (average) fuel consumption if roads with higher speed limits are chosen (motorways).		(x)	a	
Is average fuel consumption affected?	There will be a decrease in total (average) fuel consumption if roads with lower speed limits are chosen (rural roads, highways).		(x)	x	
Is the emission of CO <sub>2</sub> affected?	There will be an increase/decrease in CO <sub>2</sub> emissions (depending on whether the consequence is less time spent/trip or more journeys are undertaken).		x		
Is the number of journeys undertaken affected?	The number of journeys undertaken will decrease as the driver becomes more aware of environmental impacts.  The number of journeys will increase/decrease because it will be easier to travel and find places also in unfamiliar environments.	x			
Is the duration (in time) of journeys affected (for regular trips)?	There will be longer travel times as a consequence of lower (average) speeds.  There will be a decrease in time spent travelling because the driver will choose a smoother route (or time) based on the information provided.  There will be a decrease in time spent travelling because less time is spent trying to find the route or place.			x	(x)
Is the distances travelled affected (for regular trips)?	There is likely to be an increase in distances travelled because the driver is warned about incidents ahead (and chooses an alternative, longer route)  There will be a decrease in distances travelled because of a reduction in detours.			x	x
Is there a change in time allocated to travel (for regular trips)?	More time will be allocated to travelling.	x		x	
Is there a change in travelling in adverse conditions (dark, fog, slippery road, etc.)?	There will be an increase in journeys undertaken in dark conditions because the driver is guided by the system.		x		(x)
Is there a change in route choice (for regular trips)?	There will be a change in route choice for regular trips because of information on incidents ahead.  Drivers will increase their use of roads with higher speed limits (motorways).  There will be an increase in 'rat running' because in some cases the system reroutes the driver to minor roads.			x	X  x

Is there a change in travel mode (for regular trips)?	<p>There will be an increase in the use of private cars as the increased access of information leads to increased travel comfort.</p> <p>There will be a decrease in the use of private cars as the driver becomes aware of the many disturbances in the traffic.</p> <p>There will be a decrease in the use of private cars as the drivers get more aware of the (negative) environmental impacts.</p>	x			X	x	
Is there a change in time (of day) for travelling (regular trips)?	Travellers will start earlier or postpone their regular journeys because of access to information on incidents.						x
Is 'eyes off road time' affected (frequency, duration)?	'Eyes off road' (time and frequency) will increase because the driver allocates visual capacity to read the information. (Refers to solutions where the information is provided in a textual format).	x	x	x			x
Is mental workload affected?	There will be a decrease in mental workload because the driver has access to information on ...		x	x			
Is there a change in users' level of stress?	<p>There will be a decrease in user's level of stress because of better access to information on ...</p> <p>There will be an increase in stress level due to annoying, repetitive messages</p>		x	x			
Is there a change in users' uncertainty?	There will be a decrease in the users' experiencing uncertainty as the user is informed about ...		x	x			
Is there a change in users' feeling of travel comfort?	<p>There is an increase in users' estimation of travel comfort as the user is better informed about ...</p> <p>There will be an decrease in travel comfort due to annoying, repetitive messages</p>		x	x			(x)
Is there a change in users' feeling of subjective safety?	<p>There is an increase in safety because the driver experiences driving more carefully (according t speed limits)</p> <p>There is an increase in safety because the driver experiences driving without disturbances (accidents, incidents, queues)</p>		x				(x)
Have the (perceived) mobility options increased?	The (perceived) mobility options will increase in the following ways: .....		x				(x)
Is there a change in perceived journey quality?	<p>There will be an increase in perceived journey quality because of a perceived feeling of control</p> <p>There will be a decrease in perceived journey quality because of a decrease in personal integrity</p>	x	x	x			x
Is there a change in users' knowledge?	The users' knowledge of the function will increase over time (and increase the potential for user uptake)	x	x	x			x
Is there a change in users' problem perception?	Users' perception of problems associated with environmental issues/negative environmental impacts will increase.	x					

	Users' perception of problems associated with mobility issues will increase.  Users' perception of problems associated with safety (speeding) will increase.			x	x
--	--	--	--	---	---

Is the device/function considered useful?	The higher the perception of usefulness, the more substantial the user uptake.	x	x	x	x
Is the device/function considered satisfying?	The higher the perception of satisfaction, the more substantial the user uptake.	x	x	x	x
Does the user trust the device/function?	The higher the perception of trust, the more substantial the user uptake.	x	x	x	x
Is the user willing to invest in the device/function?	The higher the perception of usefulness, satisfaction, trust etc., the more likely the user will be to invest in the device/function.	x	x	x	x

## **APPENDIX II**

### **TeleFOT strategy checklist**

**Context, specifics**

- *The individual FOTs within the TeleFOT project are to adopt as common a test and evaluation approach as possible. The common approach is described in the deliverable, D2.212 Test and Evaluation Strategy I.*
- *The L-FOTs will aim to assess the impact of the different functions based on a naturalistic study approach whereas the D-FOTs have scope for a more controlled experimental design.*
- *L-FOTs and D-FOTs will run across different test sites and in parallel. The purpose is to benefit from the particular strengths of the respective approaches, i.e. the increased ecological validity of the naturalistic driving test results, providing evidence of behaviours and behavioural changes, and the increased reliability of the experiment, offering possibilities to identify more causal types of explanations to e.g. the drivers' behaviour than will L-FOTs. Thus, the D-FOTs is to complement the L-FOTs and support the analysis and interpretation of the L-FOT results.*
- *The TeleFOT project system boundary is the larger context of traffic, transport and travelling, and in addition the user in different roles, i.e. as traveller, passenger, as well as driver. The assessments in TeleFOT will extend beyond the in-vehicle use to include decisions made and actions taken pre, during and post trip. This means that e.g. travel diaries and questionnaires must be used in addition to logged data of in-vehicle use of different functions and of other relevant measures.*
- *Each L-FOT should address all main impact areas, i.e. efficiency, environment, mobility, safety, and user uptake, whereas D-FOTs may have a more narrow focus (on environment, safety, and user uptake).*

**Research questions and hypotheses**

- *An integrated top-down/bottom-up approach will be used for generating the relevant research questions and hypotheses.*
  - *The top-down approach is driven by the issues of relevance to the impact areas irrespective of the system functionality.*
  - *The bottom-up approach is driven by the functionality, the system design, the use cases and the impact area.*

- *The bottom-up approach will form the basis of the hypotheses list for each FOT. The top-down approach will be used in order to check that nothing significant for a particular impact area has been omitted.*
- *The number of (common) hypotheses will be kept to a minimum. However, the same hypotheses may well serve to clarify the impact in a number of impact areas.*
- *The hypotheses will be worded so as to state the measure (e.g. journey length), the direction (e.g. longer/shorter) and the comparison conditions (e.g. system in use/system not in use).*
- *Hypotheses (and hence performance indicators to be collected) will be prioritised using a cost/benefit approach. The cost is measured by the material costs and the burden for collecting the required data. The benefits are measured by the likelihood that the hypothesis will make a significant contribution to assessing the impact. However, if the different FOTs are to contribute to the European dimension, a set of hypotheses must be the same across different FOTs and applications. Therefore a set of common hypotheses will be defined.*
- *A last step in the determination of the complete list of hypotheses is consideration of the 'feasibility' of collecting data to address each hypothesis. If it is decided that a research hypothesis, although important, cannot be easily addressed within any of the FOTs, this hypothesis will be discarded from the final list.*
- *A set of common hypotheses will therefore eventuate which will be addressed across all FOTs. (A tentative list has been generated.) Thus, depending upon the function(s) to be tested, each FOT will collect data to support as many of the chosen hypotheses as possible in order to provide a European perspective to the analysis results. However, the individual FOT may choose to add additional hypotheses to be tested in the specific FOT.*
- *Several of TeleFOT FOTs will test combinations of functions. After formulating hypotheses for single functions, the hypotheses for the combined functions and possible interactions will be addressed.*

### **Study design**

- *The FOTs, whether D-FOTs or L-FOTs, must provide minimum a control or a base-line condition.*
- *A Within Subject Design should be used in preference to a Between Subject Design.*

- *If a Between Subject Design is used then the control and experimental groups must be carefully matched in order increase the power of the analysis as much as possible.*
- *For the Within Subject Design there should be a sufficient and equal amount of time (and therefore exposure) allocated to the control condition and the experimental condition in order for the participants to experience similar driving circumstances in both conditions (e.g. weather, lighting, traffic density etc) but also allowing for adequate experimental data to be collected.*
- *Given that D-FOTs have the characteristics of more controlled experiments, the D-FOT may last for a short period of time. An L-FOT should, on the other hand, last for longer period of time (minimum six, or two times six months depending upon the study design).*
- *Some FOTs will involve integrated functions. If the integrated functions can be isolated (i.e. switched on and off independently of each other) then more complicated study designs can be implemented that will test the impact of each individual function. However, if the functions cannot be isolated then they must be considered as a single treatment and any conclusions from the FOT can be made only about the combined effect of the functions being applied simultaneously.*

### **Participants**

- *The participants should mirror the intended user population while at the same time allowing for the hypotheses to be tested. Describing the characteristics of the user population is an important first step in the recruitment and choice of participants.*
- *A non-randomised sampling procedure will be used in the different FOTs. However, a random procedure could be used in order to assign each subject to the corresponding experimental condition (control group or experimental group) specified by the particular experimental design of each FOT.*
- *The required number of participants in the FOT will depend upon a number of factors. A rule of thumb is that a test consisting of 10-12 participants cannot be denominated as a L-FOT while a test involving 100 participants may. For D-FOTs within TeleFOT a smaller number may suffice.*
- *In order to ensure that the chosen sample size is representative for the behaviour of a group of users and that it is possible to statistically prove effects that are there, a power analysis is needed to calculate the desirable sample size for L-FOTs. A*



*statistical power analysis exploits the relationships between: sample size (N), significance criterion ( $\alpha$ ), population effect size (ES), and statistical power.*

- By ensuring that different FOTs that are to test the same type of function (or combinations of function) run the trials according to the same design principles, the basis for statistical tests will radically improve.*
- A plan for participant 'drop-outs' should be developed throughout the FOT (in particular the L-FOTs), either by selecting a larger sample than is needed already from the beginning or by creating a "replacement pool", i.e. reserve an extra number of participants to be recalled at short notice if needed.*

### **Study environment**

- The L-FOTs should not request the participants to e.g. drive certain roads or during certain weather conditions. This contradicts the overall purpose of an FOT, i.e. in real-life usage of functions and services.*
- Formulation of research questions and hypotheses that require a particular study environment in terms of, e.g. road types or weather conditions should therefore be avoided in the L-FOTs. Specific conditions may instead be enforced in the D-FOTs.*
- However, if the research questions and hypotheses chosen require specific environmental conditions in order to be tested, it is important to design the study so that the likelihood of the specific circumstance increases, e.g. through the choice of participants or geographical location or time of year.*
- Also, a post-trial categorisation of key variables will enable exploration of data in order to address hypotheses referring to particular environmental conditions. This requires, however, that relevant issues are identified before the test, and that the data captured enable a matching with hypotheses and scenarios of interest.*

### **Data collection**

- A minimum set of data to be collected (subjective and objective) will be determined for all FOTs in order to address the list of common hypotheses.*
- Data collected will include both objective and subjective data in order to address the common hypotheses.*
- In L-FOTs subjective data will be collected primarily by means of web-based questionnaires (but also individual interviews and/or focus group interviews are recommended complementary methods for collection of more in-depth information).*

- *In L-FOTs a limited amount of objective data will be collected due to the restricted access to the vehicles to be engaged in the L-FOTs. It will include primarily the data that can be collected by means of GPS (and if possible additional logging devices), directly or derived. In addition will information on travel patterns etc. be collected by means of a travel diary (to be filled in for one week at a time).*
- *In D-FOTs subjective data will be collected by means of (web-based) questionnaires (but also individual interviews and/or focus group interviews are recommended complementary methods for collection of more in-depth information). The number and content of the questionnaires will, however, be more limited than in the L-FOTs.*
- *In D-FOTs, the objective data will be richer as the vehicles engaged in the test will be specially equipped. This allows access to e.g. CAN data. In D-FOTs will also be carried out usability tests according to standard testing procedures.*
- *Data is to be collected pre-test, during test, and post-test.*
  - *In pre-test phase data about the participants will be collected including e.g. gender, age, driving experience, experience of devices and function, as well as information on e.g. the expectations of the participants.*
  - *During the test will in D-FOTs primarily be collected objective, logged, data. In L-FOTs will be collected objective, logged data, but also data on travel patterns, and user uptake (at the beginning of the test, after a certain period of use, and towards the end of the test).*
  - *In the post-test phase a subjective evaluation of the functions tested should be completed, in both D-FOTs and L-FOTs.*
- *Data will be transferred to a common database for hypotheses testing and further analysis by means of relevant statistical tools, both parametric and non-parametric.*

### **Preparing the test**

- *Pilot studies will be completed before the start of the actual D-FOTs and L-FOTs. These preliminary field tests will (i) check the technical function of the data collection systems in real driving situations, as well as (ii) check the feasibility of the evaluation process including different data collection methods and tools (diaries, questionnaires, etc.).*
- *The participants must be informed about the overall purpose of the project and its organisation, and of the specific FOT, as well as about possible risks, the costs covered and not covered, whom to contact in case of breakdown, etc. To formalise*

*the arrangement between the involved organisations and participants (and their employers), a letter of agreement (or similar) should be signed.*

### **Legal and ethical issues**

- *All FOTs must take into account legal and ethical aspects.*
- *All FOTs must make certain that the laws of the country in which the FOT takes place are followed. This includes to make certain that the functions tested are not illegal (e.g. is speed alert illegal in Germany).*
- *All necessary permissions for running the FOT must have been acquired before the tests start. Each FOT is therefore advised to create a specific committee involving technicians, researchers, as well as legal and ethical experts before planning and running the FOT. This committee should ensure that all necessary permissions for running the FOT have been acquired.*
- *A Letter of Agreement (or corresponding) should be signed, stating the conditions for the participant's involvement in the FOT.*
- *All FOT test sites must make certain that the privacy of the participants is secured. Data retrieved during the FOT must not allow for external actors to identify the persons involved in the test. Any personal information should be maintained in a separate and protected database.*
- *All test sites must make certain that the security of the participants in the FOT is apprehended. The organisers of the FOT are recommended to follow a structured risk management procedure. This includes, e.g. making certain that the introduction of new devices into the vehicles do not result in modifications that influence the driver's security.*