European
Large-Scale
Field
Operational Tests on In-Vehicle Systems
7th Framework
programme
INFORMATION
AND COMMUNICATION TECHNOLOGIES
ICT-2-6.2 ICT for
Cooperative Systems
Large-scale integrating project
Grant agreement no.: 223945

Deliverable D11.3

Final Report

Version number	Version 1.1
Dissemination level	PU
Lead contractor	Ford
Due date	30.06.2012
Date of preparation	12.12.2012

Authors

Christoph Kessler, Ford

Aria Etemad, Ford

Giancarlo Alessandretti, ALC

Karsten Heinig, VTEC

Selpi, CHAL

Rino Brouwer, TNO

Andras Cserpinszky, ERT

Walter Hagleitner, ADAS

Mohamed Benmimoun, IKA

This deliverable has been compiled by the above authors, but it is a summary of individual contributions from many other authors as indicated in the relevant cited references. All results are scientific findings which are only valid inside the statistical assumptions and other limits of application. All findings have to be considered with the associated range of significance. The affiliation of the authors with any organization involved in this project does not indicate that those organizations endorse all the findings contained within this report.

Project Coordinator

Aria Etemad

Ford Research & Advanced Engineering Europe

Phone: +49 241 9421 246

Fax: +49 241 9421 301

Email: aetemad1@ford.com

Ford Forschungszentrum Aachen GmbH

Suesterfeldstr. 200

D-52072 Aachen

Germany

Revision and history chart

Version	Date	Reason
0.1	2012-01-17	Structure by Ford
0.2	2012-02-08	Layout check by EICT
0.3	2012-04-22	Parts by ALC
0.4	2012-06-11	Compilation by Ford
0.5	2012-06-12	Included Methodology.Results, Ford
0.6	2012-06-12	Review 1
0.7	2012-06-15	Include review comments&changes Ford
0.8	2012-06-15	Include review in 0.7 (crash in 2nd review doc) FORD
1.0	2012-06-18	Final version by EICT.
1.1	2012-12-12	Include reviewer's comment and revised D6.x D5.x Ford and final check by EICT

Table of contents

Table of Figures

Table of Tables

Executive Summary

The European project euroFOT developed the first large scale Field Operational Test, with a focus on Intelligent Vehicles equipped with Advanced Driver Assistance Systems (ADAS) and used by ordinary drivers in real traffic. Its motivation was to evaluate different on-board functions with regard to traffic safety, efficiency and the environment. Also usability and acceptance were exhaustively evaluated. Participants either owned their test vehicles, leased them during the experiment or took part as professional drivers employed by freight companies. Data acquisition techniques ranged from questionnaires to continuous recording of vehicle signals, and also, in some cases, additional instrumentation with video and extra sensors. The following functions have been considered for passenger cars and trucks:

- longitudinal control functions: Forward collision warning (FCW), adaptive cruise control (ACC), speed regulation system (SRS)

- lateral control functions: Blind spot information (BLIS), lane departure warning (LDW), impairment warning (IW)

- advanced applications: Curve speed warning (CSW), fuel efficiency advisor (FEA), , safe human machine interface (SafeHMI)

The project started in May 2008 and ended with a Final Event at the Autoworld Museum in Brussels, Belgium in June 2012. Several hundred Terabyte of data have been collected from around 1200 drivers driving for more than 35 million km.

This deliverable summarizes the three major phases of the project: **specification/piloting, execution and data analysis.**

Field tests are a well-known method for manufacturers to look into the way their products are used by the consumer. For the first time euroFOT has brought together major European vehicle manufacturers and research institutes in order to collect data from different ADAS equipped vehicles in different countries but all with the same task: ordinary driving on real roads. Participants drove vehicles which did not look very different from standard vehicles and could be driven without special instructions. It was therefore necessary to assemble complex computer and sensor hardware, flying wires, instrument brackets or even maintenance intensive software into a nice and clean package, requiring low maintenance and worthy of the newly acquired customer vehicle.

During the first two years, euroFOT was dedicated to defining the framework for the analysis of research questions (specifying functions, experimental procedures, hypotheses, measures, indicators and questionnaires) and for testing the chain of data collection and processing (piloting). The data acquisition systems (DAS) for CAN-data, video and extra sensor recording were selected, modified or even developed from components and programmed for the different vehicle lines involved. The acquisition of drivers had started early and many drivers did agree to participate, despite low take-rate of systems and the weak economic situation. In the third year almost 1000 vehicles were on the road and collecting data. Vehicle Management Centers (VMC) had been installed to oversee the installation and maintenance of instrumentation, to provide hot line support and to administer questionnaires. In the fourth year the comprehensive task of "Data analysis" applied the methodology that had been previously prepared and fine-tuned.

The results achieved are now available and summarised in a number of public deliverables. They can be used by research organizations, public bodies and other stakeholders in Europe and elsewhere to support the wider deployment of ADAS.

The analysis first focused on system performance and user aspects, especially in dangerous situations which could potentially lead to accidents (which have been defined as ‘incidents’). This was followed by impact studies on traffic safety, efficiency and environment. Finally, the project considered a Cost Benefit Analysis (CBA).

Results

The following points summarise the main conclusions of the analysis. Overall, the final results point to positive effects on safety, and positive effects regarding fuel consumption together with high levels of driver-acceptance.

1. For both, passenger cars and trucks, the time-headway increased significantly when drivers were following a lead vehicle while using **ACC+FCW**. In addition, the relative frequency of harsh braking events and incidents decreased. Regarding changes in driver behaviour, drivers of passenger cars using ACC+FCW were three times more likely to engage in visual secondary tasks during normal driving (e.g. reading maps, looking at passengers or objects in the car). However, this difference was not found during incidents. The results imply that drivers seem to be capable of managing secondary tasks such that they focus on the road ahead when the traffic situation requires doing so. In addition, ACC+FCW does not seem to affect the amount of drowsy driving. For trucks, no particular side effects on driver behaviour were observed.

   Projecting the safety indicators of widely deployed ACC+FCW to EU-27, it was concluded that ACC+FCW in passenger cars have a positive effect on the overall number of crashes. In trucks, this conclusion could only be made for motorways. Hence, assuming that the safety indicators are good indicators, for how the accident scenario would change if all vehicles were equipped, ACC+FCW cars could potentially affect up to 2.2-5.7% of the injury accidents on motorways, while ACC+FCW trucks could potentially affect up to 0.2-0.6% of these accidents. Further estimations based on the relevant rear-end target crash population can be made for EU-27, e.g. regarding involved injured individuals. Note that the presented percentages are based on an extensive set of assumptions, which are described in 5.1.2. They are therefore to be used with caution, and need to be put into the perspective of all the assumptions made within the analysis framework.

   Based on the positive influences of ACC+FCW on safety there were also positive (indirect) effects on traffic efficiency. Due to the potential reduction of accidents the annual incidental delay calculated in lost vehicle hours could be lowered by up to three million hours on an EU-27 level. The environmental impact was measured in terms of fuel consumption which showed a reduction of approximately 3% for passenger cars and 2% for trucks. This effect does not consider benefits from changes in traffic efficiency.

   Overall, ACC+FCW seems to be a highly appreciated and well used function that increases driver comfort as well as potentially having a positive effect on safety. Questionnaire data indicated that the driver expectations were fulfilled. The positive experiences of the drivers can also be seen in the increased use during the treatment phase (31% in travel time and 53% in the activation frequency).

2. While driving with **LDW** (**+IW** for passenger cars), drivers showed a slightly improved lateral control and a small increase in turn indicator usage. There was also a trend (although not statistically significant) toward less involvement in lateral incidents. Regarding driver behaviour, drivers were more likely to engage in secondary tasks. However, this difference was not found during incidents. This result mirrors the

outcome for ACC+FCW and seems to indicate that drivers are capable of adjusting secondary task engagement to the traffic demand.

Overall, drivers indicated in the questionnaires that LDW is a useful function. Some of them perceived the association between warning and actual crash risk as weaker than for the ACC+FCW. Hence, many warnings were perceived as unnecessary. An effective warning strategy is required to meet drivers varying expectations/requirements under different driving conditions. For IW, the congruence between warning strategy and the users' perception of their level of drowsiness/inattentiveness seems to be high since user ratings were highly positive and stable over time.

3. For the **Speed Regulation System** (SRS = SL+CC), it was observed that over-speeding and harsh braking events were reduced when SL is active. The effect of CC on over-speeding was a strong increase while strong jerk, critical time gap, and harsh braking occurrences were reduced. These findings highlight the relationships between system usage and driving condition showing that the level of traffic is likely to be an important factor for system use.

   The analysis showed that no safety effects occur when using that system, despite an increase in speed. The a priori acceptance of the SRS was very positive and the use of the systems confirmed this tendency. Ratings of usefulness and satisfaction generally increased over time, except for a slight decrease experienced with SL. Drivers used the cruise control (CC) mostly in free flow conditions and on motorways (40% and 66%). Usage of the SL was lower (about one third of the driven km, less on motorways).

   Since a change in the mobility behaviour was found neither in the objective nor in the subjective data, the effect of the SRS on traffic efficiency is related to the change in average speed which increased about 2.4% on motorways. Environmental aspects of the system showed also a positive trend towards reduced fuel consumption through the system use (approximately 1%).

4. The analysis shows that **navigation systems**, as part of SafeHMI, are highly accepted and also widely used. Results indicate that route choice while driving with an active navigation system is more time efficient than the baseline condition when no navigation system was available. As expected significant positive effects on travel distance and travel time were found in the analysis. The system was used mostly on long trips on unfamiliar routes. For the tested built-in device the usage rate on these trips (long, unfamiliar) reached almost 100%. Furthermore, navigation systems seemed to support a fuel efficient route choice, depending on their routing algorithm.

   It was observed that on urban roads potentially safety relevant indicators show a positive effect if the system is activated. This positive effect is reflected in positive changes in lane keeping behaviour, distance to the lead vehicle and harsh braking events.

5. Overall, drivers indicated that **BLIS** is highly appreciated. The acceptance rating remains high over time, which indicates that drivers continue to perceive it as useful as they experience interacting with it over an increasingly large variety of conditions. Approximately 80% of drivers felt that BLIS increases safety. It is perceived as most useful on urban roads in heavy traffic and does not increase workload. However, most hypotheses tested showed no significant effect on safety or driver behaviour based on the objective data. An exception was the frequency of turn indicator use, which slightly decreased when the system was available and drivers were not simultaneously using LDW+IW. However, from the free text comments it is clear that

most drivers consider BLIS as an important complement to visual checks, rather than as a primary source of information.

6. **CSW** has also a good overall evaluation in terms of usability and acceptance. The values for satisfaction and useful categories increased significantly while using the system. According to the survey, around 75% of the drivers felt that safety is increased thanks to CSW. They also found it most useful while driving on rural roads. Some participants stated that they used CSW as indicator or for practising a more cautious driving style. Moreover, participants trust in the system increased after usage. The trustworthy and reliability scores were higher after some experience with the system.

7. **FEA** is specifically designed to help reduce fuel consumption. The analysis of the data therefore focused on the environmental impact and did not consider possible side effects that may impact traffic efficiency and safety. The treatment phase showed a reduction in fuel consumption of 1.9% based on 3.6 million kilometers from 50 trucks.

In order to supplement the above findings also some final considerations and several **lessons learned and good practices** for future FOTs are included in this report.

FOTs can definitely contribute to the evaluation of Intelligent Vehicles, but are not the unique solution for all investigations concerning new automotive systems. The methodologies that are available need to be adapted to the specific systems, also taking into account existing constraints in time and resources. Conditions that are necessary for a successful implementation of FOTs include a large variety of aspects: in particular various aspects pertaining to the industrial, technical, organisational and methodological viewpoints.

1 Introduction

There is often a perception that Active Safety and Driver Assistance Systems have a potential for improving road safety. In this way they contribute to a key element for sustainability of transport, and a major objective for car industries, authorities and the society in general. Several advanced technologies have been developed in recent years by vehicle manufacturers, providing improved vehicle controllability in difficult situations and effective support to the driver.

In particular, Advanced Driver Assistance Systems (ADAS) are now available in several modern vehicles, with the purpose of supporting the driver to help avoid road accidents or mitigate their severity. Significant positive effects are expected from their increased penetration into the market [1,2].

However, the majority of work on ADAS has so far concentrated on system development, and demonstrations were often restricted to a set of controlled conditions or to small scale investigations. Therefore, a limited amount of data is currently available on the influence of these functions in ordinary traffic. In particular, there is a need to improve our knowledge on how normal drivers react to the system interventions, whether they accept the received support in the intended way, and whether the systems provide any real world benefits.

This is particularly important, considering the fact that driver behaviour is often cited as the primary factor leading to road accidents. Moreover, further information in this respect will help make decisions for deployment at both the technical and political level [3].

With these considerations in mind, the euroFOT project has implemented **a large scale Field Operational Test (FOT) in Europe**, evaluating several ADAS applications with normal drivers on production passenger cars and trucks; a sufficient period of time was considered, in order to collect statistically sound data in real world conditions.

The following developments have been key drivers for reaching the goals of the project:

- Applying a **general methodology** at European level, which takes as a basis the results provided by the previous European project FESTA. [4,5]
- Elaborating an **experimental design**, starting from a set of research questions relevant for the applications, and defining a chain of hypotheses, indicators, metrics and signals.
- Improving the instrumentation and SW for **data acquisition** on vehicles, by means of reliable and cost-effective components.
- Developing and applying procedures for the **execution** of a FOT (e.g.: driver recruitment, driver support, vehicle management, data management, etc.).
- Developing and applying **methodologies for the analysis** of the very large amount of collected data, including new procedures and routines for automated data processing.
- **Assessing the different ADAS**, particularly in term of performance, use by the drivers, impacts on road safety and traffic efficiency.

1.1 Experiments

The central phase of the project has been the **execution of the experiments**.

The FOT was organised by four Vehicle Management Centres (VMC) across Europe: in France, Germany, Italy and Sweden. Eight ADAS applications have been considered, representing solutions already present in the market, or sufficiently mature for a commercial application (see chapter 2.1 for a description of the applications). This choice was also motivated by the need of high system reliability in a large variety of traffic situations.

Participating drivers used the vehicles during their daily routines, and large amount of data have been gathered by sensors capturing driver and vehicle behaviour, as well as the driving context: in the course of the project, data have been collected from more than 1000 drivers on 980 vehicles, travelling a total of 32 mio. km during several months.

When implementing these measures, a number of challenges have been encountered. This was not surprising because euroFOT is the first Field Operational Test with such a large scale in Europe. In particular researchers had to address the variety of vehicle types and ADAS systems, the wide geographical area covered, and also the presence of subjects with non-uniform characteristics (for example professional and ordinary drivers).

In synthesis, the following main **results regarding the experimental phase** can be highlighted:

- The FOT assured compliance with the planned procedures, customised information to the subjects, and assistance during the tests in most cases.
- Legal aspects, including privacy and security for the drivers, were properly addressed.
- The different challenges have been solved by establishing an harmonised approach for all the VMCs, which provided a general framework, but was also open to specific adaptations depending on the function, equipment, type of drivers, etc.
- Reliable techniques for data acquisition have been implemented without disturbing the driver, and normally without interventions by the experimenters.
- Data quality has been guaranteed by an appropriate monitoring and control process.
- Instrumentation of the FOT comprised questionnaires, continuous loggers for CAN-data, video, and extra sensors (e.g. eyetracking, distance radar).

1.2 Data analysis

The next important development within euroFOT has been the **data analysis.**

Five partners - universities or research centres in collaboration with vehicle manufacturers - were in charge of storing and analysing the recorded data. Ad-hoc procedures for the management of the data bases have been developed. A significant effort has been dedicated to data pre-processing, which included for instance formatting and storage into the data base, filtering, adding map references, synchronising the videos, etc.

The analysis of the FOT results for the selected functions has been focused on the following three main areas:

- User related aspects, regarding the effects on driver behaviour and performance, and in particular user acceptance and workload.

- The expected impacts on driving safety, traffic efficiency and the environment ("green" driving).

- A socio-economic cost benefit analysis.

The major achievements regarding the data analysis are summarized as follows:

- A suitable approach has been established for the impact assessment regarding road safety, traffic and fuel efficiency.

- Techniques for the identification and extraction of the required driving scenarios have been produced, mostly with the application of automatic tools.

- The feasibility of storing, pre-processing and analysing the collected data in the available time has been shown, for both subjective and objective information.

- A validity check proved that conclusions could be derived for most of the hypotheses (amounting to about 100)[1].

1.3 Objectives

The euroFOT project aims to demonstrate the benefits and to encourage the deployment of Intelligent Vehicle Systems on European roads, focusing on Advanced Driver Assistance Systems which have now reached a good level of maturity.

Previous experience worldwide has shown that Field Operational Tests are an excellent method to collect real data, evaluate the impacts and enhance the take-up of new solutions. These tests have also proved to be a powerful tool for gaining insight into the usability of functions, when operated in the real environment, and for a sufficiently long time to reach the daily operational and behavioural level.

At the same time it is clear that such experiments require considerable resources and efforts, and therefore the project is also aiming to develop and validate new advanced methods, able to facilitate the operation of FOTs and to ensure a high quality of scientific results.

In this context, the main objectives of euroFOT are the following:

- To perform multiple coordinated tests of several Intelligent Vehicle Systems with ordinary drivers in real traffic;

- To investigate performance, driver behaviour, and user acceptance;

- To assess the impacts on safety, efficiency, and the environment using data in a variety of driving scenarios.

A specific target of testing 1200 drivers on 1000 vehicles was defined, with a plan to collect data for travel in different traffic situations, and reaching a total exposure of around 18 million km.

[1] A few exceptions are due to driving situations occurring very rarely, or to the difficulty in obtaining statistically significant indicators

EuroFOT also intends to contribute to the on-going efforts for the diffusion of advanced Vehicle Technologies and the promotion of a general European perspective. The following two additional goals are therefore particularly important: consolidating a common approach for the implementation of FOTs, and raising awareness in the general public regarding the potentialities of driver support functions.

1.4 Methodology

One of the main goals of euroFOT was to develop a suitable methodology for conducting the experiments, so that the study would be as useful as possible, i.e. enabling scientifically sound conclusions to be drawn from the resulting data.

In order to properly evaluate the effects of the driver assistance functions, a general work plan was therefore defined, starting from the methodology developed by the previous FESTA project, consisting of three phases for the preparation, execution and analysis of the experiment.

However, a number of adaptations and practical details were needed for the specific purposes of euroFOT. In particular, the following areas required new developments: the elaboration and proper use of pilot studies, the iterations in the sequence of phases, the analysis of combinations of functions, legal and ethical issues, methods for handling a large data-set, and finally simulation methods for scaling up the socio-economic impact analysis.

A clear understanding of the systems under evaluation has been the first step for the euroFOT experiment. Hence, starting from a definition of several research questions and related hypotheses, a set of more than 100 performance indicators has been specified, describing driving behaviour, driver workload and acceptability, traffic safety, traffic efficiency, and impacts on the environment. This goes down to the level of a detailed prescription of the actual metrics to be applied for each of these aspects. In addition, a set of situation variables has been developed to take into account different conditions related e.g. to weather, traffic conditions, road types or vehicle occupancy. A measure matrix was also implemented to represent all the measurements to be performed in the FOTs: this was especially useful for defining the sensor set-up on board and the data acquisition system.

A typical experimental design was based on the following approach:

- The FOT is conducted for a period of 12 months
- The first three months serve as a baseline period, when the functions to be tested are deactivated
- The next period involves the so called treatment phase, when drivers are free to activate the assistance functions according to their needs and habits
- The final testing of hypotheses is built on a comparison between the baseline and the treatment periods.

Not all the tests could follow this ideal scheme, and other experimental designs have been used depending on the application, time available and ADAS set-up. The supply of questionnaires was also regulated by well-defined procedures at selected time intervals.

1.5 Outline of the report

This final report aims to summarize the work done and incorporates the main findings of euroFOT. It is expected that the results of the project will be used by the European automotive industries and by the society in general: technology providers, research community, final users, public authorities, etc. to influence the deployment of ADAS in vehicles, and to indicate the best routes for future developments.

More details can be found in the project deliverables, for which a selection is listed in the chapter “References”. The name and subdivision of subprojects (SP1 to SP6) into workpackages (WP1100, etc.) can be resolved in the overview on the project structure given in the chapter 6.1.

The structure of the report follows the phases of the project plan, schematically shown in Figure 1: In chapter 2 *Specification and Piloting* the functions are described, the experimental plan is given and the pilot experiments are explained. At the end of chapters 2 to 5 *Lessons Learned* summarize the experiences gathered. Chapter 3 *Execution* refers to the implementation of the tests, operational aspects and the data flow. Chapter 4 *Methodology* focuses on the methodologies for data analysis and specific tools for data evaluation, including: user related aspects, expected impacts, and the cost-benefit analysis. Chapter 5 *Results* describes the major outcomes from the experiments within euroFOT. Chapter 6 *Management* provides general information on the project, with considerations on procedures and legal aspects. Finally Chapter 7 *Conclusions* reports general comments derived from the experiences, and shows directions for further work.

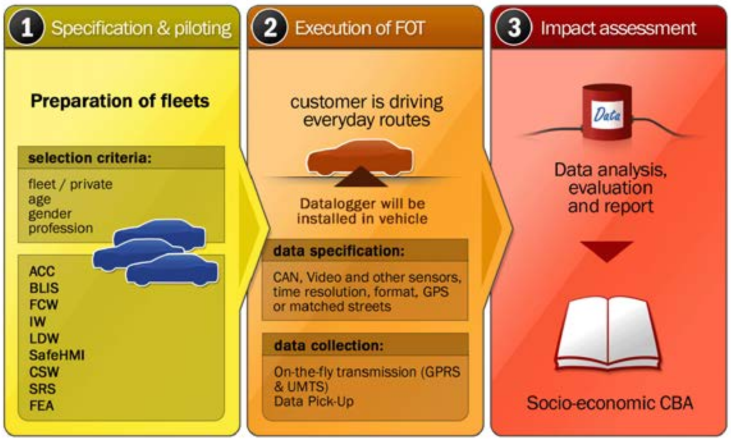

Figure 1: Scheme of the project plan

2 Specification & Piloting

The first part of an FOT is to prepare the theoretical base of the experiment. After the functions have been selected the experiment can be designed. The general research questions are unravelled to from hypotheses, use cases and scenarios.

Further on the piloting procedures need to be described, which take into account the complexity of the data acquisition systems that are developed in the second step “Execution” and the integration and operation of the DAS inside the vehicle. Thus, a parallel approach is necessary with a lively communication between S&P and Execution.

2.1 Functions

Based on the recommendations of the eSafety Working Group, on existing roadmaps, and on the availability of well-developed systems, nine functions developed by leading European OEMs have been selected for testing, as shown in Table 1.

Table 1: Overview of brands, test vehicles, and functions of interest in the euroFOT project

	RENAULT / ceesar	Ford	MAN			BMW	DAIMLER	FIAT GROUP / CRF CENTRO RICERCHE FIAT	VOLVO	VOLVO Volvo Technology Corporation
FCW		✓	✓						✓	✓
ACC		✓	✓	✓	✓				✓	✓
SRS	✓									
LDW			✓	✓	✓			✓	✓	✓
BLIS									✓	
Nav sys						✓	✓			
CSW		✓								
FEA										✓

The functions are specified as follows:

2.1.1 Forward collision warning FCW)

A forward collision warning system provides an alert to assist drivers in taking the necessary action to help avoid or reduce the severity of crashes involving the equipped vehicle striking the rear of another vehicle.

This function detects and tracks obstacles in front of the vehicle and provides a warning to the driver in case the evaluation of trajectories and relative speed of the subject vehicle and the obstacle show a high probability of a collision. This function is intended to decrease drivers reaction time in case of potential rear-end accidents.

2.1.2 Adaptive Cruise Control (ACC)

The ACC function supports the driver in selecting (and then automatically maintaining) an appropriate speed and distance to the vehicle in front depending on his/her preferences and the current traffic situation.

The ACC function actively controls the vehicle speed to adapt to drivers selected speed and following distance. This function is designed to detect and track if a vehicle is in front and adjusts the speed accordingly (e. g. by controlling the throttle or braking). If the leading vehicle accelerates, the function accelerates up to the target speed and keeps in the pre-selected following distance. The system is disengaged when the driver acts on the brake or when the driver pushes the related disengage button (in case of manual gear, changing gear will also disengage the function). The function is not active below a certain speed and when the vehicle is started.

2.1.3 Speed Regulation System (SRS)

The Speed Limiter function limits the speed of the car in order to prevent the driver from exceeding a selected speed limit value. This speed limit value is preset by the driver during the system activation. The minimum value of this speed limit is 30 km/h. This function can be OFF or ON. When it is ON, it can be active or inactive. When it is ON and active, it is only restrictive when the speed of the car reaches the selected value. When it is restrictive, it can be temporarily overridden.

The Cruise Control function maintains a constant speed without any manual control by the driver. This speed is set by the driver. This function can be OFF or ON. When it is ON, it can be active or inactive. When it is active, it can be temporarily overridden. The system can only be activated when the speed is above 30 km/h and the last gear box positions (position 4 and 5) are engaged.

2.1.4 Lane Departure Warning (LDW)

The LDW function is designed to warn the driver when the vehicle is unintentionally departing from its current lane of travel. LDW uses lane markings to monitor the vehicle’s position on the lane. LDW continuously assess the vehicle’s relative position on the road and a warning is issued to the driver if the vehicle is unintentionally (e. g. not using the turn indicators when crossing the lane marker) venturing from its intended lane.

2.1.5 Blind Spot Information System (BLIS)

A blind spot warning system is designed to provide feedback to the driver in case an object has been detected in one of the blind spots of the vehicle. This function continuously monitors the rear blind spots on both sides of the vehicle. In case an obstacle is detected in the blind spot an information/ warning is issued to the driver.

2.1.6 Safe Human Machine Interaction (SafeHMI)

Safe design and use aspects for the human machine interface are essential for all in-vehicle information and communication systems. These aspects are considered for different HMI types of navigation systems which provide location and route guidance information to the driver. Several types of systems (e. g. OEM fitment, after-market solution) with different display positions and technologies are already on the market. The BMW system uses a head-up display to put selected information directly in the driver's line of sight, Daimler’s navigation system offers the possibility to select route guidance by speech as well as route guidance by manual control.

2.1.7 Curve Speed Warning (CSW)

CSW technology has been developed with the goal of identifying potential dangerous situations (such hidden curves, and sharp turns) and warn the driver in advance. The system is currently in development status.

2.1.8 Fuel Efficiency Advisor (FEA)

The FEA supports the driver in maintaining the engine speed in the "green area" towards optimal usage of the vehicle with respect to fuel efficiency. The driver is also advised when the engine speed is outside the "green area" longer than a pre-set limit. Further FEA warns the driver when a certain speed threshold is reached and when the engine is on idle for an extended time. Green area limits can be administrated by the Dynafleet back office system. When the driver is driving outside the "green area" longer then the allowed time limit, a warning message will be shown at the display to alert the driver. Driver performance (percentage of time within the "green area") will be logged for every hour and sent to the back office for post trip follow up.

2.2 Design of the FOT

Experiments come in different shapes and sizes but they all have a common goal: to show an effect of an experimental treatment. The experimental treatment can differ from the hopefully beneficial effect of a medicine, to a positive effect of friendliness on someone's well- being to the positive safety effect of an advanced driver assistant system. The scale of the experiment can be small (one or two participants) or very large (hundreds). Whether big or small or whatever the experimental treatment a good experimental method is needed to ensure that any effect can be ascribed to the experimental treatment alone. Although large in size, a field operational test such as euroFOT is still an experiment or quasi-experiment [41]. And given the often chaotic circumstances in an FOT (the uncontrollable daily driving conditions) a good experimental method is especially important.

In euroFOT eight different functions were tested. The different functions investigated in euroFOT can be divided in three groups (see also Figure 2):

- longitudinal control functions: FCW, ACC, SRS
- lateral control functions: BLIS, LDW, IW
- advanced applications: CSW, FEA, SafeHMI

This figure also shows to some extent the difficulty in developing the experimental design that euroFOT faced. Table 1 (above, Section 2.1) shows the distribution of functions over car manufacturers. Six of them have vehicles with multiple functions. The aim of euroFOT is to show the benefit of individual functions and therefore the individual contributions of different functions

The individual effects of these functions preferably need to be isolated through the experimental design. However certain functions in certain tests are always operating together. For example, when ACC is switched on so is (in some cases) the FCW system. The individual effects of these functions cannot be isolated through the experimental design. They need to be unravelled, if necessary, in the analyses phase.

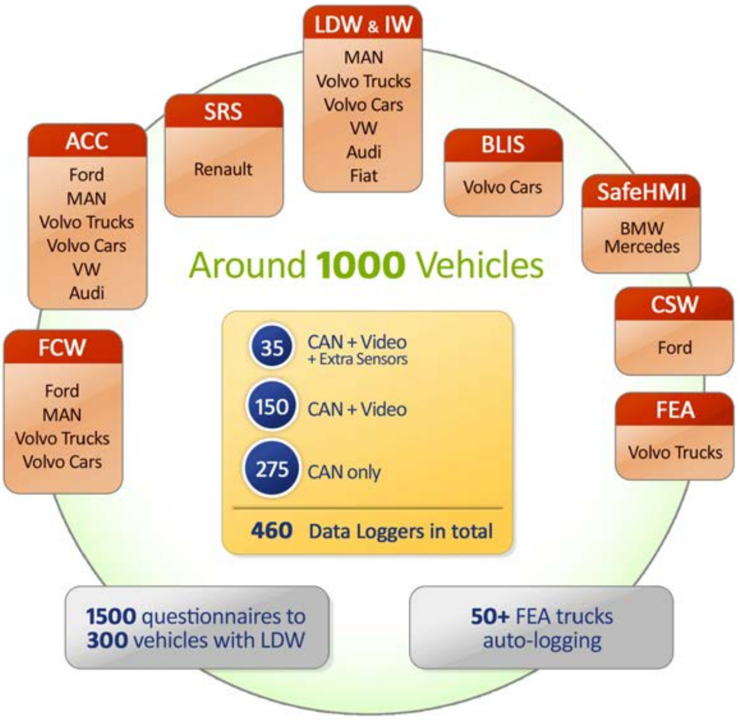

Figure 2: Overview of systems and data loggers in euroFOT

2.2.1 The experimental method

The experimental method describes how the experiment is conducted. It describes the sample (the drivers' age and experience), the design and the procedure (when was what done during the experiment). We will describe what the initial ideas were in the following three paragraphs.

2.2.2 Participants

In SP4 a power analyses was performed to assess the number of drivers to be involved, the duration of baseline and experimental treatment and the kilometres driven by the drivers in euroFOT. The recommendations were as follows (deliverable D4.2):

"The simulation results suggest that as effect sizes become more modest, and hence provide a safer bet in terms of achieving results, the number of required cars increases substantially. Simulations have shown that when at least 120 participants are included, who drive 15000 kilometres per year, sufficient power will be reached for even finding the small effect sizes that can be expected in a FOT. Including more cars or more unique participants should take precedence over measuring for longer periods. For example, measuring for a year with 60 participants could fail where measuring for half a year with 120 participants could provide a significant result."

Another way of improving the power would be to reduce the variance measured between participants. This can be achieved by choosing a homogenous group of drivers for example male drivers between 30-40 years of age with similar mileage. This would improve the power, but at the cost to generalise (external validity) the results.

The proportion of time spent with and without the system or systems should be roughly equal. However, the simulations do not take into account a phase in which the participant makes the transition to driving with the system(s), which may need to be analyzed separately. Ideally, the order of systems and baseline periods would be counterbalanced.

The simulations have shown that up to 15000 km per year the power increases rapidly for almost all effect sizes but a higher distance driven is beneficial if the effect of systems on/off is small."

In summary:

- The baseline period should be as long as the treatment period
- approximately 120 drivers needed per test site
- approximately 15000 kilometres
- homogeneous group of drivers

These recommendations on basis of the power analysis could not be completely followed. The number of cars was fixed and therefore to some degree so was the number of drivers. One can of course have multiple drivers in a vehicle but then the duration of both baseline and treatment can be too short especially for certain systems (e.g. an FCW).

Since the systems tested by euroFOT are factory installed most drivers are recruited that buy a new vehicle with the systems. These drivers are asked to participate in the tests. The baseline should therefore be as short as possible given the fact that the customers paid for these systems and it seems unreasonable to restrict their use for a long time. Fortunately normal driving can be assumed to be more stable as it has usually been practised over years. It was decided to have a shorter baseline period then the treatment period; the **baseline would last 25%** of the time and the **treatment phase would last 75%** of the time. This has the additional advantage of capturing possible changes over time in the usage of the system during the treatment phase.

The age of the drivers should be between **30 and 50 years** to ensure some homogeneity, the recruited driver should drive around **15000 km a year** and all drivers should have a **full driver license**. They should have **no experience** with the systems tested.

In deliverable D4.2 it was also described how to deal with drivers who declined participation and drivers who drop out. Drivers who decline to take part may be inherently different from those who do. For example, unintentionally only those drivers may be sampled who have positive attitudes to road safety or to new technologies or those who feel they have the time to fill in the questionnaires and these variables might be correlated with their driving behaviour during the test phases and hence bias the results. The best that can be done is to attempt to quantify in very broad terms the characteristics of the non-participating sample, that is, main socio-demographic variables as age, gender or socio-economic status with empirically established correlations in driving and traffic behaviour.

If a participant or a vehicle has to drop out of the FOT, a decision has to be made as to whether to "replace" them in the data set. A "drop out" is defined as a participant or vehicle whose data collection ceases for a specific amount of time. In this sense "replacement" means the recruitment of another driver for the participant who dropped out of the sample. First of all, it is recommended that the baseline period of driving should be mandatory for all drivers. Furthermore:

- Drivers who have to drop out before the baseline period is completed should be excluded from the analyses and replaced with a driver who meets the demographic criteria outlined in the previous section.
- If the cessation of driving occurs before the baseline and treatment periods are of equal length the driver is also replaced.

2.2.3 Experimental design

An experimental design can either be a “between subjects design” or a “within subjects design”. In a between subjects design there are minimally two groups of participants. One group of participants receives no experimental treatment (baseline condition) and one group of participants receives the experimental treatment. In the within subjects design the participants participate in both the baseline condition and the experimental treatment condition(s).[2]

In euroFOT both within and between experimental designs were considered. A between subjects design looks as follows:

Month											
M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12
Baseline											

Month											
M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12
Driving with system X											

One of the advantages of this design is that there is no need for a control condition. Such a control condition can reveal changes in driving due to circumstances that are not under control. A major disadvantage is that two groups of drivers are needed preferably both of the same size and more participants than in within subjects design due to the higher variation in a between subjects design

A within subjects design looks as follows:

Month											
M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12
Baseline						Driving with system X					

[2] For more information on experimental design for FOTs the reader is referred to euroFOT D4.2 and to the FESTA deliverable ‘Primer on experimental design’ [5].

A disadvantage of this design for an FOT is that there is no control condition during the treatment period. This can pose a problem when the baseline period is for example during spring and summer and the treatment period is during autumn and spring. Differences in driving due to seasonal changes are then not captured. In case of a within subject design a control condition is required.

In euroFOT we recommended to adopt a within subjects design with a small control condition. However both designs have been applied.

2.2.4 Procedure

The procedure of an experiment describes what was done at different points in time during the experiment. It describes for example what participants need to do, what kind of information is provided, etc.deliverables D5.4 and D4.3 [16] describe in more detail how participants of euroFOT could be informed on the goal of the test and what they could expect. They should also receive information on the usage of the systems. One of the more challenging parts of the procedure was the timing of the questionnaires. An extensive questionnaire was developed within euroFOT for measuring acceptance, usage, workload, etc. The timing of the questionnaires, the instructions and the debriefing are presented in Figure 3.

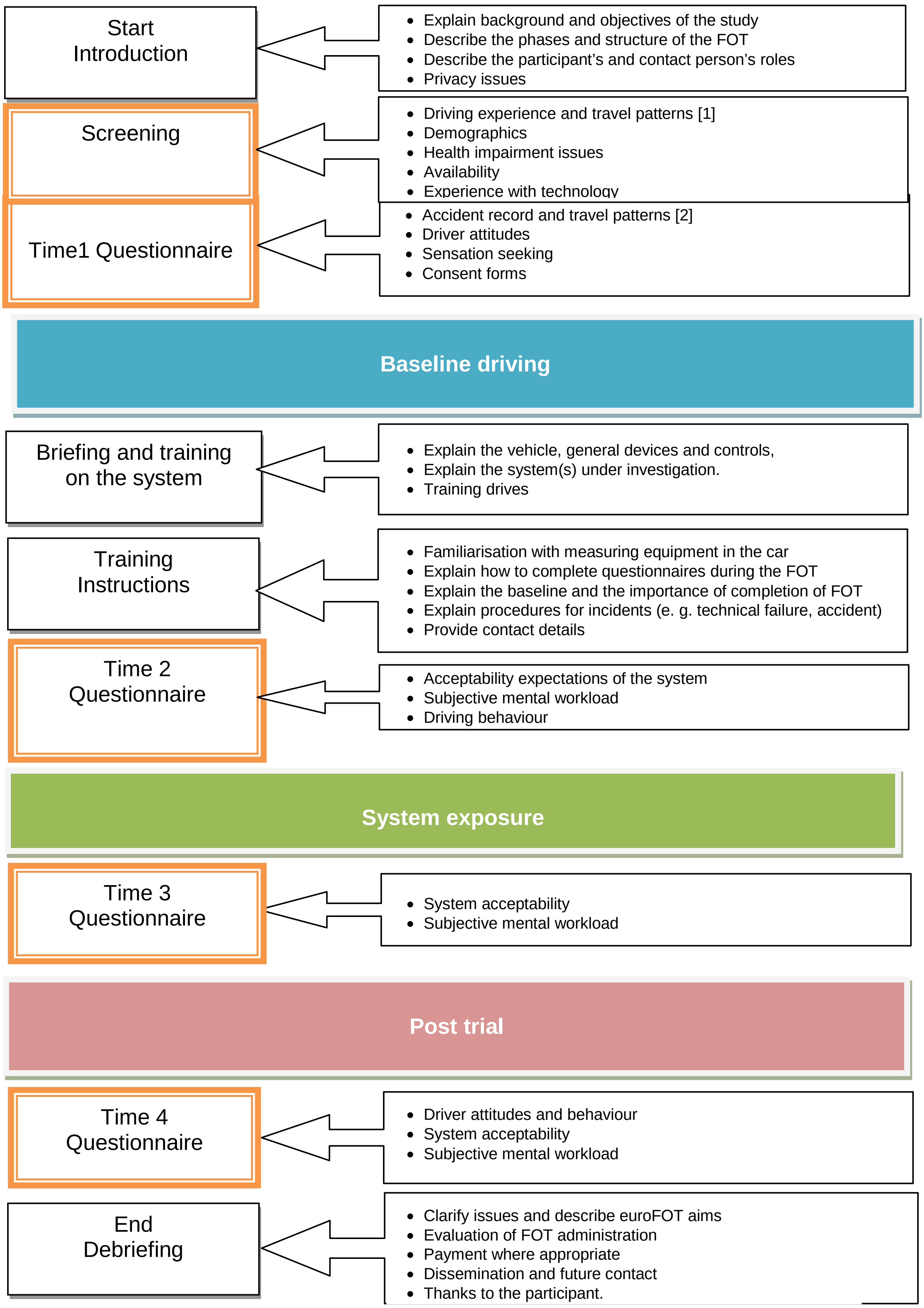

Figure 3: The different interactions with the drivers in euroFOT.

2.3 Specification

2.3.1 Combination of functions

Once the functions to be tested and the research questions (e.g. Does the function increase safety?) are defined. The next step is to specify the hypotheses which can be verified or otherwise. This has been further structured into five main tasks (Figure 4). A necessary intermediate step during specification is to check if combination of functions exist which cannot be easily separated. An example is longitudinal control through ACC and FCW: During ACC usage a potential collision will be announced by braking initiated through the ACC followed by an FCW-event. The combination of ACC and FCW can thus mask effects and needs to be considered in the formulation of hypotheses.

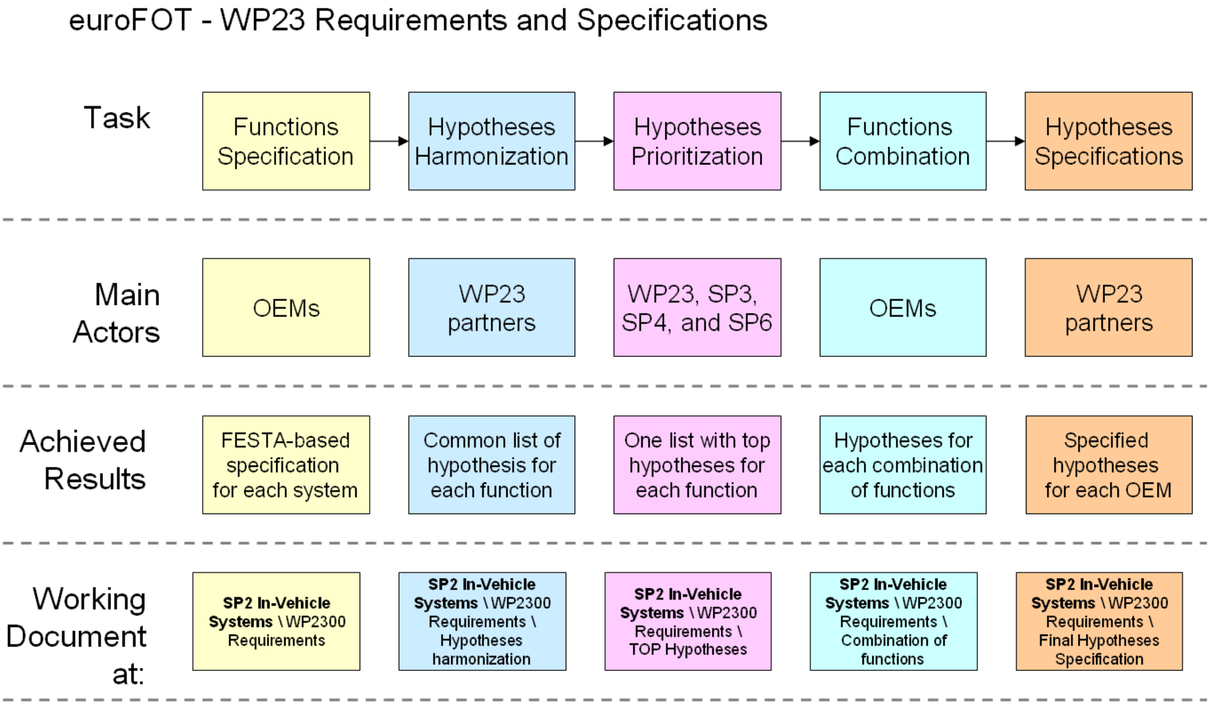

Figure 4: Breakdown of tasks to collect and harmonize hypotheses during specification.

Drivers may not use in-vehicle systems as development engineer thought they would or as the drivers' manual advises them to do. For instance, a particularly lazy driver may set ACC speed very high to perform an overtake manoeuvre without pressing the throttle pedal.

Some drivers might temporarily override ACC by accelerating in overtaking situations leading to a significantly reduced time gap to the vehicle in front before changing lanes to overtake the vehicle. Some "aggressive" drivers might also accelerate to close the gap to the vehicle in front (in case preceding vehicle is driving on the "faster" lane) thereby signalling an overtaking "intention". These situations are hard to predict but may influence and potentially confound analysis if not recognized and filtered out from the data. Further, a specific analysis of such situations may furnish very precious insights for product development and help understand driver behaviour. For these reasons, hypotheses specification has also considered misuse and abuse of functions

To identify the effects of the systems under investigation on driving behaviour, traffic safety, environmental impact and traffic efficiency performance indicators are needed.

2.3.2 Performance Indicators

Performance Indicators are defined quantitative or qualitative measurements, agreed beforehand, expressed as a percentage, index, rate or other value, which are monitored at regular or irregular intervals and can be compared to one or more criteria. Performance indicators are directly measured (e.g. speed profile) or derived from a measurement (e.g. speed is used to calculate an average speed, to calculate standard deviation of speed, the deviation from speed limit, the acceleration, etc; see Figure 5). The performance indicators had to be defined to test the hypotheses but also to investigate the usage, acceptance and satisfaction with the systems. In order to ensure that possible effects can only be ascribed to the systems under investigation, other influences (e.g. weather conditions) need to be excluded. Within euroFOT these 'situational variables' were identified and a description on how to measure them was provided. An extensive overview has been produced that describes different performance indicators and situational variables which assisted the standardization of these variables over different test sites and which can be used for standardization across FOTs in Europe (see deliverable D4.1).

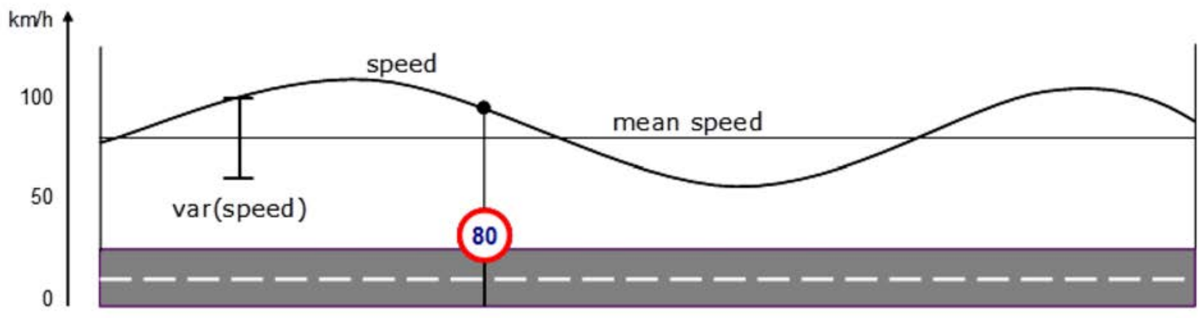

Figure 5: An example of different performance indicators based on speed

2.3.3 From hypotheses to experimental design and performance indicators

The importance of the preparation phase and especially the development of the hypotheses is often underestimated. The relevance is that a well formulated hypothesis identifies different relevant aspects of the experimental design and the performance indicator. The relevance can easily be demonstrated. Preceding a hypothesis is a general research question. A research question comes from a general idea about the effect of a system. An example of a research question for a Forward Collision Warning (FCW) system may be:

Will an FCW system influence traffic safety?

Such a research question can only be investigated when '*influence*' and '*traffic safety*' are better defined. A simple definition of '*Traffic safety*' is the number of accidents and we assume that the number of accidents decreases. So '*influence*' is translated as a 'decrease'. So, a first hypothesis could be:

Hv1: An FCW system decreases the number of accidents.

Fortunately for an individual driver there is a small risk of being involved in an accident. However that makes testing the hypothesis rather difficult. It would mean, for example, that data need to be collected for a very long time with many vehicles. Since accidents seldom occur 'surrogate measures' have been proposed. These are measures that can more easily be collected and have a potential relationship to accidents. So an alternative for *Hv1* would be:

Hv2: An FCW system decreases the number of incidents

Incidents happen more often than accidents. However there is no general agreement on what an incident is. Therefore we need to define incidents ourselves. For example we can define an incident as every occurrence in which the vehicle decelerates with more then 4 m/s^2 (and for ease of reading and writing we call this hard braking)

Hv3: An FCW system decreases the occurrences of hard-braking

In this hypothesis '*incident*' is defined as '*hard braking*' (this is defined as decelerations of more than 4 m/s2). However, '*decreases*' indicates that a comparison needs to be made. There can be no reduction without some kind of comparison.

Hv4: An FCW system decreases the occurrences of hard-braking, relative to a baseline condition

So the comparison that will be made is with driving in a baseline condition which in this case means driving without an FCW.

Hard braking can occur for many reasons and some reasons may be of more interest than others. For example hard braking in a car-following situation in which the leading vehicle suddenly brakes can be more of interest (since this is a situation where the FCW is designed for) than hard braking at a crossing just because the light turns red. So we can add to *Hv4*

Hv5: An FCW decreases the occurrences of hard-braking, relative to a baseline condition, in a lead-vehicle braking scenario

We have come from a general (not easily testable) research question to a testable hypothesis. The importance of the process of developing hypotheses is that the final hypothesis has important ingredients that are needed in the development of the experimental design, namely:

- A baseline condition which is driving without an FCW
- An experimental condition which is driving with an FCW
- Measuring speed such that decelerations can be calculated, or
- Acceleration sensors for the calculation of decelerations
- Calculation of maximum deceleration
- The event of a lead vehicle braking
- FCW decreases

So in the experimental set-up we need to ensure that we have a baseline (driving without the FCW), that we have an experimental condition, that we can identify a lead vehicle that brakes, that we can measure hard braking and finally that we ensure that any effect found can be ascribed to the FCW system. The hypothesis has indicated to a large extend what needs to be done in the experimental design. And in euroFOT the most challenging aspect is to ensure that any effect can be ascribed to the system under investigation and to that system alone. For that reason the situational variables developed in WP4300 were very important.

2.4 Piloting

During the process of designing a large scale FOT experiment, many steps are needed to assess reliability of the whole data acquisition chain from a technical point of view (devices are working as expected) but also from a purely organisational aspect (chain of processes are working as expected). Piloting consists of a theoretical framework and the realization of the required test procedures. Three work packages have therefore been scheduled through the whole preparation process.

- Small scale testing and validation (WP3600)
  - The “small scale test” can be regarded as something like a “factory acceptance test”, a laboratory test of the whole data management chain covered in SP3. It has, from a purely technical point of view, tested the data chain from data acquisition in the data acquisition system (DAS), through the data upload and quality assurance procedures, to the database, to the analysis tool. It has validated the technical requirements, and the aim was not to test the systems’ suitability for study of the functions in “real life” as this is a matter for the formal pilot tests.
- Adaptation and in-vehicle implementation (WP2400)
  - This work package concerns the technical installation and adaptation of the chosen DAS. The goal was to supervise the installation of the data loggers in vehicle prototypes with the help of the OEM and suppliers. Suppliers and OEM gave information about the CAN-bus and other function-specific needs. WP2400 also helped solving some problems encountered during the pilot test experimentation.
  - These tests covered in-vehicle technical implementation of the material needed to collect data, and also the correct functioning of the system once integrated in the vehicles.
  - Each partner was responsible for creating a detailed installation guide for equipping their fleet of vehicles with specified logging and sensing equipment. The guide describes in a structured way what and how to install so that a person responsible for the installation (e.g. a mechanic) can manage to install all the equipment with little or no help at all.
- Piloting (WP2500)
  - These tests have addressed real situations using the specific equipment. Based on technical specifications/documentation given by the DAS provider (SP3) and also on the SP4 inputs on the final experimental design, various verification procedures have been identified and tested in a small scale experiment.

2.4.1 General considerations

The piloting activity is of much importance in the process of designing a large scale FOT since it is the final step before the real experiment begins. A wide range of objectives are covered by such a task:

- Test the entire data acquisition chain installed in a vehicle prototype driven in real conditions.

- Use verification procedures to identify the problems and solve them with the help of WP2400.

- Test subjective data collection, participant briefing and experimental design. More generally, test every aspect not directly related to objective data collection.

- Give feedback to the partners involved in the methodological aspects of the FOT in order to modify a potentially unfeasible method (research questions not precise enough, performance indicators (PI) not computable, lack of data or needed frequency not available, etc.)

- Identify and disseminate the best practices for piloting.

- Authorize the start of the large scale experiment after every necessary condition has been fulfilled.

Although various OEM's with different systems and data acquisition equipment are involved in the euroFOT project, a common approach is needed to ensure the same quality level of the collected data and therefore the analysis to be conducted. This was one the major challenges of the pilot test activity. As described in Figure 1, the diversity of the partners involved made this goal very difficult to reach for both cultural and economical reasons.

The following elements represent the basis of the euroFOT approach to piloting and are described in this section:

- Use cases: The most representative use cases are evaluated during the pilot.

- Scenarios: A tested driving situation defined by a set of verification procedures.

- Verification procedures: Step by step description of the tests.

- Incident: Description of the circumstances and part of the equipment concerned by an incident or a non-conforming operation.

2.4.2 Identification of the most representative use cases for each VMC

Each VMC needs to identify what are the most representative use cases for each system tested. Identified use cases should be listed by each VMC using a unique ID. The proposed use cases template can be found in deliverable D2.1 Annex 2 [. Table 2 gives a real use case in the proposed template.

Table 2: Example for the description of a use case for piloting.

ID	Description General description of the use case			
USE_SafeHMI_02	Driving on route with congestions			
System and vehicle requirements				
System status Possible states of activation of the system	System activity Possible activity of the system	Vehicle characteristics Test vehicle	Interaction between systems Possible interaction with other in-vehicle systems	Activity occurrence Define how often the system intervenes

ON-OFF	ACTING (giving route information) NOT ACTING CONSIDERING TRAFFIC INFO (dynamic routing) NOT CONSIDERING TRAFFIC INFO			Depends on route

Environmental requirements			
Traffic conditions Required traffic condition	Environmental situation Required lightning, weather, visibility conditions	Road characteristics Required road characteristics	Geographical characteristics Required geographical characteristics relevant for testing the systems
High traffic density	No specifications	Different ways to destination are possible	No specifications
Comments			
Familiarity is assessed via button press for Daimler and via trip diary for BMW			

2.4.3 Identification of representative scenarios

In addition to the use cases, it has been suggested to describe scenarios that are representative of the activities scheduled in the VMC. Writing down representative scenarios may be useful to assess if an activity is described from the beginning to the end. For example, if the data analyst wants to use one of the performance indicators (PI) to answer one research hypothesis, it is necessary to check the possibility of computing the PI from the variables, then to check if those variables are present in the database and in the correct format and precision, and if the raw values are accurate enough. Doing this, it may be easier to discover some gaps in the complete process of trying to perform a single action.

Each scenario will therefore be associated to a flow of necessary verification procedures (VP). In order to check if the scenario is a success, each VP needs to be checked separately. The proposed template to describe briefly the scenarios is given in deliverable D2.1 Annex 3. The following graph gives an idea of what is meant.

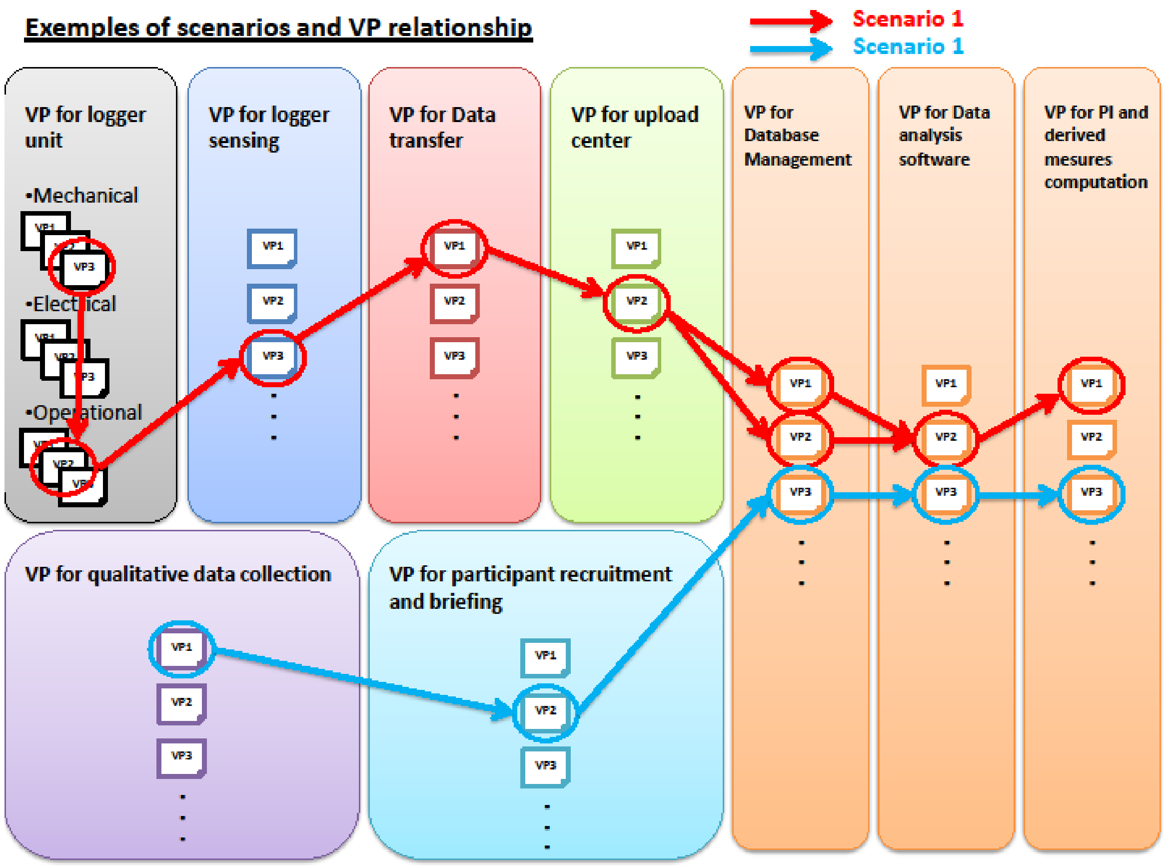

Figure 6: Example of a sequence of verification procedures (VP).

2.4.4 Verification procedures

The verification procedures (VP) describe a single test, for a single element of the data acquisition chain. There is a need for these tests to be described step by step. Conditions of acceptation and use cases also need to be detailed precisely. In order to facilitate the work and to ensure harmonised presentation a common template has been proposed

From the detailed list of procedure, it has been proposed to summarize them in the form of a validation. This validation plan is in the form of a checklist that should include the main information of the procedures, and can be used for reporting the results of the testing procedures and the envisaged solutions (see example in Table 3). An importance level has been specified to indicate “Procedure of such an importance that leads to a GO/NOGO condition.” (Level 1) and “Procedure that can be modified in the following way: delete, simplify or improve.” (Level 2).

Table 3: Verification procedure template

Unique id	Creation date	Creator	Summary (name)
VP_01	2008-03-22	G. Saint Pierre	File creation when the vehicle is started
Importance level	Version	Stakeholder	Type of material concerned
1	0.1	French VMC	Data logger

Description of the use case
The vehicle engine is started as usual, and the driver starts to drive,
Description of the functionality to be tested
The Data logger should create a file containing all the measures in the correct format (this particular aspect will be the subject of another verification procedure). The file should be created if the engine is on during more than 30 seconds, or if the car is going further than 500 meters.
Description of the verification procedure
- Start the engine less than 30 sec. - Shut down the engine - Check if the file exist on the system - Start the engine more than 30 sec. - Shut down the engine - Check if the file exists and if the collected parameters are correct - Start the engine and drive more than 500 meters and less than 30 sec. - Shut down the engine - Check if the file exists and if the collected parameters are correct - Start the engine and drive less than 500 meters and more than 30 sec. - Shut down the engine - Check if the file exists and if the collected parameters are correct.

2.4.5 Incidents

All detected incidents or non-conform events need to be saved in a traceability tool to track the solutions that took place until the problem is solved. Two types of documents are necessary for that:

- Incident definition
- Failure report and tracking

An example of an incident definition tracking is given in the next Figure:

Reference	Title	Status	Date opened	Date closed	Link
INC_001	CMS stopped by system reboot after automatic update	closed	11/02/2010	11/02/2010	INC_001
INC_002	Short data acquisition before engine start	open	11/02/2010		INC_002
INC_003	Data acquisition continues after KL15 Off	closed	11/02/2010	22/02/2010	INC_003
INC_004	Single trip split into several files (DL reboot)	open	11/02/2010		INC_004
INC_005	Interpolation in 'empty' time ranges by sync. Tool	open	12/02/2010		INC_005

Figure 7: Example for an incident definition and tracking scheme for piloting.

2.4.6 Conclusion on methodology

This common framework has been constructed by the partners in order to satisfy all VMC's needs. However, the work has been organised so that each of the VMC could maintain its own specifics, either due to the functions tested, the materials used, or the database structure chosen. This proposed framework was used by each of the VMCs as a basis to develop their own verification and testing procedures and templates.

2.4.7 Technical functioning of the data acquisition systems in real driving situations

Some of the equipment installed in the vehicles is common among the partners. That includes the central logging unit (data acquisition system), and positioning system. In addition to this each partner has its own set of sensing systems, that is, sensors already available in the vehicles or even extra sensors and video recording. Below is a list with top-level requirements for the common equipment setup:

- Data acquisition system (DAS): Logging unit that will collect and synchronize data delivered both by factory and additionally installed sensing systems. DAS shall also have wireless communication capabilities for uploading data to a central server as well as having the possibility to remotely control the unit. DAS shall be specified for automotive usage to manage any expected environmental impact.

- Video logging system (where applicable): Several video cameras will be installed covering both internal and external views.

- Positioning system: A GPS system will log the position of the vehicle whenever it is in use.

- Extra sensors: Some of the vehicle fleets will be equipped with special sensors.

    - Eye/head tracking system (for monitoring the state of the driver's head and eyes concerning e.g. position, gaze direction and eyelid opening)

    - Lane tracking system (lane position, lane width etc.)

    - Radar/lidar sensors (for monitoring the area around the vehicle)

    - Digital map data for the Electronic Horizon

In this section all the identified main categories of verification procedures are described. Some of them are common, and some are VMC specific. This section's aim is to list all the possible categories of data acquisition chain parts that may eventually be relevant for the VMC's.

Verification procedures are grouped so that there is a clear division between CAN only (CO) loggers, CAN+Video (CV) loggers, and CAN+Video+Extra sensors (CVE) loggers.

In addition, to the processes and procedures identified in this report, there might be other procedures, requirements or methods that are VMC dependent. For different reasons, some of the processes and procedures listed here might not be needed for some VMC's.

Below are the identified categories that have been addressed.

a) Verification procedures for logger unit (CO,CV,CVE)
    a. Mechanical
    b. Electrical
    c. Operational
        i. Start-up of operational verification application
        ii. Vehicle data
        iii. Audio logging (CVE)
        iv. Positioning systems (CO,CV,CVE)
        v. Driver/vehicle metadata (CO,CV,CVE)
        vi. Video logging (CV,CVE)
        vii. Driver monitoring (CVE)

viii. Head/eye tracking equipment
ix. Feet proximity sensors
x. Non-video environment data sources (CVE)
xi. Extra sensors (CVE)
d. Driver interaction with the logger
e. Acquisition of vehicle bus data

b) Verification procedures for logger sensing
a. Functionality verification
i. Functionality verification - Standing still
ii. Functionality verification - Driving
iii. Specification of common needs for data logging
b. Data transfer
i. Wireless
ii. Manual
c. Upload centre
i. Formatting
ii. Synchronisation and timing
iii. Data quality
iv. Data storage

c) Post processing related aspects
a. Database management
b. Data analysis software
c. Performance indicators and derived measures computation
d. Answers to research questions (function specific)

The verification procedures made use of tracking tools and traceability tools. For example, to know what type of problem is encountered for certain vehicle, to know the frequency of those problems, to know the associated software or system configuration, etc..

The installation handbook was tested. That means one of the first verification procedure was to install the Data acquisition chain following the handbook in order to check if everything is understandable by the final user.

2.4.8 Conclusions

Pilot tests are of primary importance during the process of building an FOT experiment. They can be viewed as a small scale FOT whose goal is to acquire data in real driving situations, with ordinary drivers. In contrast, the WP2500 experience was a bit different as the complete data collection process was not precisely defined at the time of the pilot tests.

One example were the pilot tests of the subjective data collection process for which the last version of the questionnaires was not available. Moreover, due to the difficulties of the participant recruitment process, it had not been possible to use ordinary drivers to test the equipped cars. Therefore, drivers from the VMC staff that were not directly involved in the euroFOT project have been driving the cars instead. A small drift from the FESTA methodology had, therefore, to be accepted during the piloting process to ensure on time and efficient tests.

The work related to data treatment and analysis had to be put aside from the pilot tests. Therefore, some items of the data analysis chain such as PI computation scripts, statistical data analysis methods, or software tools implementation (event detection, visualisation etc.) could not be tested extensively during the pilot tests as most of them were continuously being developed and modified until the very last moment.

During the euroFOT pilot tests, a large part of the analysis process was still under construction, and it is likely to get the same situation in future FOTs. New methods are implemented during the long duration of an FOT experiment (more than one year in the euroFOT case), and a large part of the necessary methodology was neither well known nor detailed in the literature. Therefore, further research and discussions between the partners were necessary.

Considering the above comments, an FOT pilot test is not a truly small scale FOT in the sense that the final step of analysing the data and answering the research questions is not performed. It is of course important and very useful to collect that data in the form as close as possible to the final format for the following purposes:

- Validate all sensors, logging devices, and transmission process.
- Validate the quality of the collected information (missing data proportion, veracity of measured values, etc.).
- Provide an estimate of the final amount of data to be collected.
- Provide a first data set in real conditions to test, validate, and fine tune all the software and methods necessary for the final analysis (Database format conversion tools, PI computation scripts, scripts for statistical methods, etc.).

During the euroFOT project the whole data acquisition chain has been tested by all VMCs. This ranged from the technical implementation and adaptation of the systems in the cars to the data transfer to the final database server in raw format. It has been considered that generating the final database structures are only relevant after the start of the analysis part of the FOT process. The main reason is that suitable statistical analyses are closely dependent on database format and structure. Coming back to the FOT chain, as described by the FESTA methodology and adopted by the euroFOT project, Figure 8 describes the topics covered by the euroFOT pilot tests.

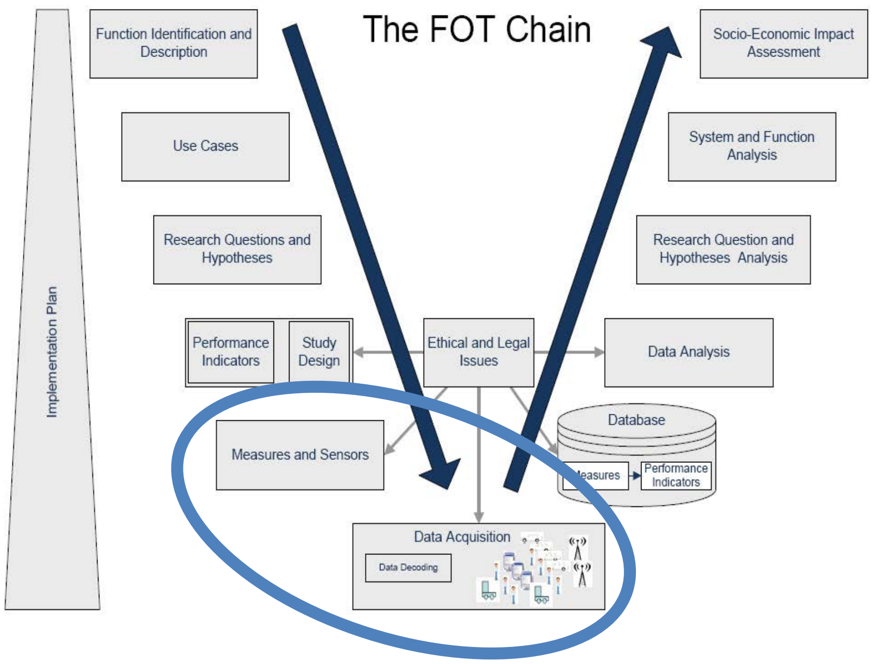

Figure 8: The steps that typically have to be considered when conducting an FOT, as explained in the FESTA methodology. The large arrows indicate the time line, and the ellipse indicates the steps validated by the pilot tests.

During the euroFOT pilot tests a collaborative approach has been employed setting out various recommendations and assumptions. They were used as a basis for discussions in numerous conference calls involving all the Vehicle Management Centres (VMCs). More than ten conference calls between partner together with a physical meeting (Stockholm 23/09/2009) were necessary to follow the process of piloting different equipped cars for different purposes in different countries. This work was closely related to the WP2400 devoted to in-vehicle implementation of the equipment (Measures and sensors cell in Figure 8).

Piloting the euroFOT project was a great challenge due the variety of the teams and the differences in the implementation of the equipment. A common approach has emerged from the discussions, mainly in the form of templates for use cases, verification procedures, and test scenarios. Due to the different OEM technical teams involved in the pilot tests, their internal processes, as well as the specifics of the vehicles equipped and the different sensors that were used, it has proven to be very difficult to adopt a common methodology in the pure sense. All teams, though, have greatly benefitted from the ideas and procedures generated with this common approach.

2.5 Technical problems encountered and solutions

Technical issues are more relevant for the objective data collection rather than for the subjective one (e.g. questionnaires). Important technical problems during the subjective data collection piloting tests have not been observed for any of the VMC. It was very important to test each functionality of the online survey tool, project participants' details database, data entry and data export in order fulfil the FOT specifications and to ensure data consistency. As for all the VMCs' classical online survey tools were considered (LimeSurvey for example). The process is already well established and no surprises were observed. The main aspect to control is the process to fill the final database with subject's data. This is far less complicated than the same issue for the numerous objective signals collected in real time from hundreds of cars.

The main issue in piloting is to ensure the technical functioning of the data logging systems under real conditions. As many OEM were partners of the euroFOT project, different kind of vehicles and different levels of equipment were used and tested using different methods. Some differences were also observed depending on the devices. For example, were they already widely in use to equip vehicles, or were they non-standard and firmware dependent.

During the pilot phase all components and processes which are necessary for the final FOT were tested; starting with the production process, configuration process of the data logger, data storing and transmission and ending with the processes for vehicle service.

The first technical aspect to be tested by the pilot tests were related to the practical aspects of vehicles equipment: cables, powering of the equipment, connectors, size of the sensors, compact flash cards compatibility, and many other items. During piloting, special emphasis was given to EMC and connector quality and the OEM experts have developed professional electric harness samples for DL-connection and also professional installation documentation. All cable harnesses where carefully tested before installation.

2.5.1 Example from FORD (German 1 VMC)

A problem occurred with continuous power supply to the logger. The initial solution to pick-up clamp 30 (non-interrupted 12V power) from an internal control unit was changed to the on-board-diagnostic plug. The power to the internal control unit was controlled through another unit and switched off some 10 minutes after ignition off. This problem did not occur during testing as the collected signals (hour-long drives on several runs), as well as GPRS connection, showed no errors. The error manifested itself as a wrong start-up-sequence of the logger due to an incomplete shut-down procedure.

2.5.2 Example from MAN (German 1 VMC)

During the first piloting phase (production pilot) the cabling of the data logger of different lots were manufactured with wires of different colours and diameters. In order to maintain a reliable production of the harness, the specification of the cabling was updated appropriately.

The piloting process is important and may lead to a change in the electrical cabling and additions to the installation handbook. Due to the complexity of the installation itself, a final testing procedure is recommended.

2.5.3 Example from FORD (German 1 VMC)

During the final test of the installation procedure, each set (DL together with vehicle cabling and antenna, SIM-card, SD-card, programming for CAN-logging and GPRS-access) was installed in a test car and taken on a 5 km test drive. Correct channel reception and connection to the server (maintained by partner ika) were necessary to receive clearance for each set. During this final test on 100 sets three non-functional SIM-card, one corrupt SD-card and two broken cables were found and had to be repaired or replaced.

Issues were more frequent with some new equipment: data logger, embedded PC for videologging, eyetracker or other specific sensors. The piloting phase has, however, been long enough to identify many non-blocking but really annoying issues, some of which only happened in really specific use cases.

2.5.4 Example from CEESAR (French VMC)

An issue occurred with a power-off problem with the embedded PC, which did not always shut down when ignition is turned off. These has been solved by not relying on the hardware generated event, but replace it by a change in the videologging software which now continuously monitors the ignition signal and correctly stops itself and the rest of the PC when it goes off.

Another technical issue is the availability of the signals necessary for the statistical analysis of the hypotheses. Data may not be collected in a correct format, or at the expected frequency. Some unexpected work may be needed in case suitable signal is not directly available. In this case, a solution needs to be found to provide the information through other ways and/or to adapt the equipment in order collect data that fits the needs of the planned analysis. Enough time should be planned in advance for searching and testing all needed signals in detail.

2.5.5 Example from Daimler and BMW (German 2 VMC)

Both Daimler and BMW collect data from several CANs, and BMW also collects data from MOST and Flexray. In the piloting process, it took a lot of effort to find and verify all the signals needed.

For Daimler, the status of the navigation system was not directly available. After modifications, the text message is sent from the on-board system to the dash-board. It is then decoded and used as an indirect signal providing information on the status of the navigation system.

For BMW, the system status that is needed is available on MOST. However, it is not sent as a cyclical signal but only whenever system status changes. To avoid missing relevant information at the beginning of a drive when the logger is still booting, the status signal had to be changed into a cyclical one.

2.5.6 Example for Volkswagen (German 1 VMC)

The series production cars ACC sensors don’t deliver the signals required for some euroFOT hypothesis. In order to retrieve these signals a special parameter set for the ACC engine control unit (ECU) was developed. A python script for the data logger was created to send stimulation messages for the ACC ECU. This setup was proven out in the endurance test.

Volkswagen ACC and LA development departments have supported the euroFOT team by making the piloting cars available and with the definition of CAN-messages lists needed for testing the defined hypothesis.

Finally, piloting tests provide the occasion to test integration of equipments together and observe problems that are caused by unexpected interactions.

2.5.7 Example from CEESAR (French VMC)

Eyetracker and videologging system perform well separately. However, the eyetracker generates a constant IR filtering in video cameras. This issue has been greatly attenuated by new filters and complementary IR lighting for the standard cameras.

Interactions are not always physical. They can also happen during signal transmissions. Also, communications signals can be corrupted.

2.5.8 Example from MAN (German 1 VMC)

The message generation for LDW data is realized by a python script which initiates the transmission of a CAN message to request additional data from the LDW System. When an update process for vehicle software or a diagnosis process was started, the message corrupted the communication between LDW system and the diagnostic tool. The script was amended by a function which stops the transmission of the CAN Message in this specific case.

Even if such issues are quite common, nobody knows exactly and in advance which one of them will surface. All these examples described above show how uncertain the combined system behaviour can be in real driving conditions and how much effort may be necessary when an unexpected technical issue is found. The most important lesson learned from the technical euroFOT pilot tests is that for future FOT enough time should be planned in advance for this activity.

Each VMC has successfully tested the proposed data transmission chain for a time period that was long enough to ensure reliability. They have learned how to optimally install the data acquisition system into the car, and how to control the intrinsic system functions. At last, also the effort needed for the installation and de-installation activities was estimated.

During the piloting phase not only the efficient functionality of the data loggers was demonstrated, but also the correct vehicle data recording and correct data transmission to the central data storage centre (CDS).

2.5.9 VMC internal organization

The perhaps most difficult part of the pilot is to make sure that all conditions, in which problems are met, are really understood and classified. Those conditions include all environmental factors that may and actually do affect equipment behavior. They also include all the successive modifications to the instrumentation. A positive effect is expected from all modifications, but regressions and side effects can (and do) occur.

Therefore, a thorough traceability process must be implemented during the pilot. This includes using other tools than the data logger (for example, manual logs were used by the French VMC) to keep track of all trips done, and the conditions in which they were done. The logs were useful to have a reference to which the recordings can be compared. It is also important to properly keep track of all problems, the efforts to solve them, and the solutions or workarounds found. Here again, a proper traceability is mandatory. This can be enforced by using collaborative tools such as Bugzilla or Redmine, which, although designed for software development, fit this kind of need.

The collaborative issue is particularly relevant when several departments within a specific OEM are involved to solve problems encountered (for example CAN-signals, regulations for installations, warranty issues). The departments of a single OEM are cooperative and result-oriented. Nevertheless, the high workload requires good communication to explain the out-of-the-ordinary needs of an FOT.

2.5.10 Recruitment

From early on in the project, it became clear that trying to search customers buying a new car as FOT-participants will be very difficult.

In the French VMC, it has been decided to only recruit drivers using a car they just bought to avoid pre-learning of system’s usage. In the German 2 VMC it was decided to assess safety impact with new customers to avoid prior knowledge.

Put together, these conditions lead to important difficulties to find and recruit the subjects; especially during the economic crisis in 2009, in which vehicles sales dropped significantly.

Additionally, the installation of data loggers into costumer cars is difficult because the installation is time consuming and requires adaptations of the vehicle. It also limits the availability of the drivers. Another recruitment issue was due to privacy which is still something not widely considered. A lot of effort was invested to find solutions in which the FOT-vehicles are actually owned by the OEM, drivers recruited from a fleet or increased incentives. The solutions found, were in the end OEM specific and cannot be generalized easily to other FOTs.

2.5.11 Subjective data collection

It was very important and useful to test every organisational procedure in detail. in order to adequately become familiar with them. Development of recruiting materials, operational procedures for administering questionnaires to drivers, collecting paper-based questionnaires filled in and driver liaison centre activities are procedures that have been tested in piloting.

The development of common questionnaires started too late in the project. For this reason some VMCs were not able to pilot the final version of the questionnaires and had to use preliminary releases. Furthermore, shortly before the start of some of the pilots, the main component that was still missing were the questionnaires to be provided by another project partner not involved in any VMC data collection process. In general, it seems preferable if instruction material, questionnaires etc. are prepared by partners who are actually involved in conducting the FOT. This makes adaptations to problems appearing in the interaction with the participants easier and quicker.

2.6 Lessons learned

2.6.1 Experimental design

For all VMCs it has been very difficult to get drivers involved in euroFOT. The original requirements with respect to **participant characteristics** needed to be loosened in order to get enough drivers involved. The age boundaries were changed from 30 – 50 to 30 – 59 but most VMCs even had drivers outside those boundaries. The annual mileage (above 15000 km) was generally kept although in some VMCs there were participants who indicated to drive less than 15000 km or even below 10000 km. Preferably drivers had no experience with the systems under investigation. It's not clear from the results presented whether and to what extent drivers had experience with the systems. However drivers have been recruited who already owned a relevant vehicle which makes it only reasonable to assume that they already had (some) experience with the systems in that vehicle. Although these changes may be needed to get the number of drivers on the road they do not contribute to a homogenous group of drivers. Consequently it may prove to be more difficult to find effects of the systems under investigation.

Related to the participant characteristics is the **representativeness of the samples**. In all VMC's there were hardly any women participating (except for Sample 2 of VCC). For truck drivers this was expected. It is not clear whether women more often refused to participate or that they are simply not in the pool of drivers from which the VMCs sampled. For a European FOT the sample should preferably reflect the European distribution of driving licence holders. However it has to be noted that due to the specific requirements on participants' characteristics (age, mileage) already the characteristics of the population from which the samples are recruited changed. Not every driving licence holder drives 15000 km a year.

The **questionnaires** were judged as being overly extensive (too long, too detailed). In general the experience with web based questionnaires was good. It has to be noted that an important aspect of euroFOT is to assess subjective experiences as usage, acceptance, willingness to pay etc. Clearly for operation centres with multiple systems the questionnaire is bulky indeed. But that is due to the number of systems under investigation and the fact that the aim is to identify the effects of individual systems.

Web-based questionnaires were preferred although participants also liked to discuss the paper version with their family. In general with subjective 'paper-based' data collection different ways should be offered to the participants ensuring the commitment of participants to fill out the questionnaires in a meaningful way.

The **experimental design** that was proposed in D4.2 was not completely implemented. With the exception for the Italian test-site the control group was sacrificed for having more instrumented vehicles with systems on the road. Also the experimental designs proposed for the different operation centres were abandoned for mainly practical reasons. For example switching individual systems on and off was proven to be more cumbersome than initially assumed and was not appealing to customers who bought the vehicle. Therefore, complex experimental designs were not followed (except for MAN).

From a pure experimental design point of view full access to all systems involved in the FOT including the possibility to switch each of them on and off is preferred. Such access would give full control over whether the systems are used or not and gives full control which systems can be used. Then complex experimental designs are possible to unravel the individual effects of systems. For some VMCs this was possible. However it has to be noted that this technically and logistically (e.g. something is changed on the vehicle so checks need to be made whether it is still suitable to drive) not a simple task

Whenever you choose to do something there is also something that you don't do. This also applies to experiments even as large as FOTs. In euroFOT the choice was made to test **systems that were already on the market** and sold to customers. The drivers in euroFOT were mainly customers who bought a vehicle and were asked to participate. In, for example, the IVBSS FOT (Integrated Vehicle-Based Safety Systems; see http://www.its.dot.gov/ivbss/) the integrated systems were developed by the research team. They were tested and tried on closed circuits until they were allowed to be on the road and used by unescorted drivers. The vehicles were then lent to drivers that wanted to participate in the FOT. IVBSS had therefore in principle full access to the systems. The benefit of this approach is that you give something to a participant. You give them a full instrumented vehicle they can use. Within euroFOT to ensure the participation of drivers the **incentive** has to be high (as high for example as in IVBSS where participants were lent a vehicle). In euroFOT MAN offered an especially high incentive: 'A euroFOT sales package has been defined, which allowed the sales people to offer an attractive participation package to potential customers. This resulted in the acquisition of 21 customers in Germany with a total of 57 vehicles for euroFOT.' Although it was still difficult for MAN to get participants due to economic crisis this sales package was appealing for customers.

The incentive must be appealing for the participants and should be at a level that the researcher does not have to make too many compromises with respect to the experimental method.

At some test sites it was not possible to switch the system(s) completely off (without the possibility for drivers to switch it on again) and drivers were simply asked not to use their systems. However, there was no direct control whether they did or did not. And with multiple systems in a single vehicle it was practically not feasible to switch just one system on. One reason for this was also that customers paid for their vehicles and were already asked not to use them in the baseline condition. This again shows a compromise of the researcher in order to get drivers involved.

Given the economic downturn it was quite difficult to get enough drivers that were willing to participate. In order to get a substantial **number of drivers** on the road some of the requirements of drivers were loosened (e.g. age, experience with systems, annual mileage). As described above the risk exits that the group of drivers became too heterogeneous to show statistical significant effects (whether or not this is true will be shown by the analyses). It's questionable whether the pressure within a project of having vehicles on the road should prevail over the selection of participants. It's the difference between being a research project or a demonstration project with research involved.

For a research FOT the selection of the right sample of participants (age, mileage, and experience) should prevail over having the promised numbers of drivers on the road. Preferably more time should be spent on getting the right sample of drivers.

In euroFOT there were minimal requirements on the drivers' characteristics (age, mileage, experience, etc). However such requirements were not identified for **professional drivers**. Requirements such as experience as truck driver, experience with the truck, should be developed.

Requirements for professional drivers need to be developed

Finally, as stated IVBSS had in principle full access to the systems and the data. However this comes at a price. The development and tests of the systems and lending vehicles to participants in IVBSS resulted in a lower number of vehicles in IVBSS than in euroFOT. If euroFOT also had provided participants with a vehicle then the costs of the project would have been considerably higher.

2.6.2 Piloting

The euroFOT project is very ambitious in many aspects, but the piloting tests were clearly underestimated. All the partners learned a lot from that experience. Some of them (if not all of them) were involved for the first time in such a large scale naturalistic driving study, and many problems were encountered and solved for the first time. In order to help the future FOT in reaching a successful data collection process, some lessons learned have been identified and detailed by the partners involved in euroFOT. Here is a list of the main ones, together with a short explanation of the reasons that lead us to reach that point.

Do not underestimate the time required for comprehensive pilot testing.

This advice was already stressed by the FESTA project. Piloting is a complicated task and therefore needs a finely detailed planning to ensure timely resolution of new issues. All FOTs must have an extensive pre-test, desktop test and pilot to perform well in a larger scale. Piloting must start well in advance, to be able to validate the system before the large scale FOT.

Pilot testing's importance should never be underestimated and should reflect real use of the FOT systems

Installations that work individually in the hand of research engineers still need a lot work for care-free operation in unknown real world conditions.

Do not underestimate the resources needed for incentives / vehicle costs.

Participant recruitment is a difficult task that needs to be achieved while the pilot is still running. Difficulties to recruit may have an impact on the experimental plan, and also the type of instrumented cars (OEM or subjects own cars). Incentives are very important to convince the few persons interested and suitable to enter the experiment, while budget is crucial to quickly adopt alternative solutions.

Piloting should always be done with the vehicle type to be used in the FOT and enough time for adapting the data logging equipment has to be calculated.

Do not use different car types during the pilot tests than the ones scheduled for large scale FOT, even if they are very similar. Due to differences in vehicle design and vehicle electronics, a lot of further adaptations can be required before the set-up of the data logger could be finalized.

Never start a large scale FOT before all components of the system are validated in production condition.

The whole chain should be put into motion, including the data retrieving and processing platform, which may be somewhat overlooked in a DAS-centered pilot. The pilot should seamlessly become the FOT ramping-up phase, without any technical change at that point. It would be an error, for instance, to validate software on a ‘pilot’ platform, transfer it to an ‘FOT’ platform which is slightly different, and expect the transition to be seamless, while the FOT really starts, i.e. when everybody should be focusing on organizational matters and forget about technical difficulties.

Data analysis related aspects cannot be tested during the pilot tests

The euroFOT partners used the FESTA methodology but realised that it was not feasible to strictly follow the guidelines. The pilot activity starts early in a large scale FOT workflow, and it is unrealistic to expect ready to work software tools at the time the data collection begins. Unless the partners of an FOT have developed such tools in a previous project, it is more realistic to consider that piloting data will help to build them instead of testing them.

Piloting tests cannot be made with external subjects

It proves to be very difficult to recruit the drivers for an FOT. People from the OEM not aware of the FOT project can be asked to drive the cars during the pilot tests. They can help diagnose questionnaire issues and briefing quality. There is sufficient variance between internally recruited subjects to ensure a pilot test that is representative. We recommend accepting this small bias in order to save time, money and gaining efficiency by improving the feedback quality.

Piloting tests should continue during the ramping up of the FOT

Once the main technical aspects are secured and tested, ensuring that the driving data is collected, the FOT can begin. Other parts of the data collection and analysis process can still be tested even once the large scale experiment begins. Strictly speaking, the post processing part of data treatment can be tested and validated during the experiment. This is particularly true for database structure, data model, new variables or PI computation, event detection or situational variables detection scripts.

Pilot is extremely useful to improve recruitment procedures

It is very important to highlight the research aims of the whole project while recruiting drivers. Piloting phase is very important to improve these procedures and highlight the most effective contact protocol, therein including wording.

Offer multiple options for filling in questionnaires

It is very good to duplicate or offer multiple options for filling in questionnaires (i.e. hard copy and electronic copy) in order to reach more drivers and improve response rate. It is also very important to develop and improve a web-based survey tool for data collection in order to save time and prevent data transcription mistakes.

Keep in touch frequently with the drivers

It is very important to get in touch frequently with drivers, in order to ensure data collection, but this has to be done on case by case basis, since some drivers might react differently to this approach.

3 Execution

The execution stage of an FOT relies heavily on the availability of the followings major points:

- Data acquisition systems to be installed in the vehicles and continuously collect vehicle data, driver's behavior and environment of the vehicle. Data acquisition equipments are summarized in Table 4.
- Tools and procedures for the rest of the data management (i.e., data transfer, data processing, storage and access management) to make the data ready for analysis. The summary of these are given in Section 3.1.
- Vehicles and drivers participating in the FOT, as well as the management of the FOT daily operation. See Section 3.2.

3.1 Data acquisition equipment

Here, DAS is defined as the in-vehicle hardware and software that constitutes a complete logging system, including external sensors used for the purpose of logging necessary signals for FOT. In the deployment of field operational tests, data-logging is a key issue, and euroFOT is no exception. Success of euroFOT depends on the DAS performance, it has to record all the information planned to be logged with no data-loss or corruption, be proven to work in different conditions, and be reliable for use in a long period (at least one year) by a large number of vehicles.

It was a big challenge for SP3 to provide the best suited data acquisition systems for this project. This is mainly due to very high requirements (in terms of technical, safety, ethical, legal issues, as well as cost) from the euroFOT project. Furthermore, none of the data acquisition systems available on the market at that time could be used directly for large scale FOT purpose. Therefore, the euroFOT partners have had to develop suitable data acquisition systems from commercial off-the-shelf components to suit the euroFOT purpose.

Only an overview of the data acquisition systems used in euroFOT is given here. But, details about the requirements and the whole process that was followed to provide the best suited data acquisition systems can be read in deliverable D3.3.

Overall, five state-of-the-art data acquisition systems have been successfully used in the euroFOT project (see Table 4):

- CTAG Data logger II (CAN-only logger)
- CEESAR videologging system (a combined CTAG Data logger II and video logger based on Nexcom VTC3300, with radar, eye tracker and lane tracker)
- SAFER-euroFOT DAS (an integrated CAN and video logger based on Nexcom VTC6100, with external accelerometer and eye tracker)
- BMW DAS (a CAN and video logger on network attached storage, with two radar sensors)
- Daimler DAS (a CAN, video and audio logger).

These data acquisition systems are accompanied with some diagnostic functions to help detect problems with different components of the DAS that might happen during the FOT. In addition to that, web-based monitoring tools were also built to accommodate DAS-quality check during FOT. Before they were used for FOT data collection, each of these DAS has had to go through extensive tests and quality check. The details of the diagnostic functions and the web-based monitoring tools available for each of the systems, as well as the tests that were performed are described in deliverable D3.3

Some photos showing the data acquisition systems used and how they are placed in the vehicles are provided in Figure 11 to Figure 15. German-1 operation centre is presented first then French VMC and then Swedish VMC; this sequence was chosen as there is a gradual increase in complexity of the data acquisition systems to be used in these VMCs. German-2 Operation Centre (OC) is presented after the previous three; this OC is the only OC in euroFOT that studied SafeHMI. The Italian VMC is presented last, as this VMC does not actually use sophisticated DAS as the other VMCs.

Table 4: Details of Data Acquisition system in the VMCs

Category	**German-1 OC**	**German-2 OC**		**French VMC**		**Swedish VMC**
		Daimler	**BMW**	**Light instrumen-tation**	**Heavy instrumen-tation**	
Main component	CTAG Datalogger II	PROVEtech:VA	ReadyNAS 2100	CTAG Datalogger II	CTAG Data logger II and Nexcom PC (VTC3300)	Nexcom PC (VTC6100)
Number of CAN channels collected	4	4	4	2	4	4
Operating System	Embedded Linux	Linux	Embedded Linux	Embedded Linux	Embedded Linux, Windows XP	Windows XP Embedded
Storage device	SDHC (up to 8 GB)	3 USB-disks	4 x 1TB hard disks	SDHC(up to 8 GB)	SDHC (up to 8 GB) and 80GB automotive hard disks (for video)	compact flash card (16 GB) + 500GB office hard disks
Cameras	-	4 (driver face, forward view, rear view, navigation system)	4 (driver face, forward view, rear view, and centre stack to see driver hands and legs)	-	4 (dashboard view, forward, rear, feet)	For cars: 4 (forward view, rear view, driver, feet); For trucks: 4 (forward view, blind spot, driver/cabin, feet)
Radar	Existing vehicle radar (via CAN)	Existing vehicle radar (via CAN)	2 Short range radar sensors	Long range ACC radar (200m)	Long range ACC radar (200m)	Existing vehicle radar (via CAN)

Category	German-1 OC	German-2 OC		French VMC		Swedish VMC
		Daimler	BMW	Light instrumen-tation	Heavy instrumen-tation	
GPS	Built-in-the data logger	External GPS and existing vehicle GPS	Existing vehicle GPS	Built-in-the data logger	Built-in-the data logger	External GPS
Lane tracker	Existing vehicle information (via CAN)	Existing vehicle information (via CAN)	Existing vehicle information	-	Mobileye	Existing vehicle information (via CAN)
Accelerometer	Existing vehicle information (via CAN)			Existing vehicle information (via CAN)	-	External two-axis accelerometer with additional yaw sensor
Eye tracker	-	-	-	-	One camera system	One camera system
Audio functionality, route identification	-	Driver interaction box (for audio comments, and route identification) and microphone	-	-	-	-
External Navigation device	-	Added for certain period	Added for certain period	-	-	-
Data transfer	Wireless (CAN data)	Wireless (status information), disk pick-up (CAN+video data)	Wireless (status information), disk pick-up (CAN+video data)	Wireless (CAN data)	Local Network transfer	Wireless (status information), disk pick-up (CAN+video data)
Encryption	Yes	-	-	Yes	Yes	Yes
Compression	Yes	Yes	-	Yes	Yes	Yes

3.1.1 Installation

The integration of equipment is described in deliverable D2.2: Report of the results of adaptation, in-vehicle implementations and piloting. The following figures give examples of real installation of the DAS into the vehicles.

Unlike the other VMCs, the Italian VMC does not log any CAN-data and therefore does not use any sophisticated data acquisition system. The data collection in the Italian VMC relies only on questionnaires. The drivers provided their feedback by answering specific questionnaires or filling in forms related to specific events that happened during the FOT.

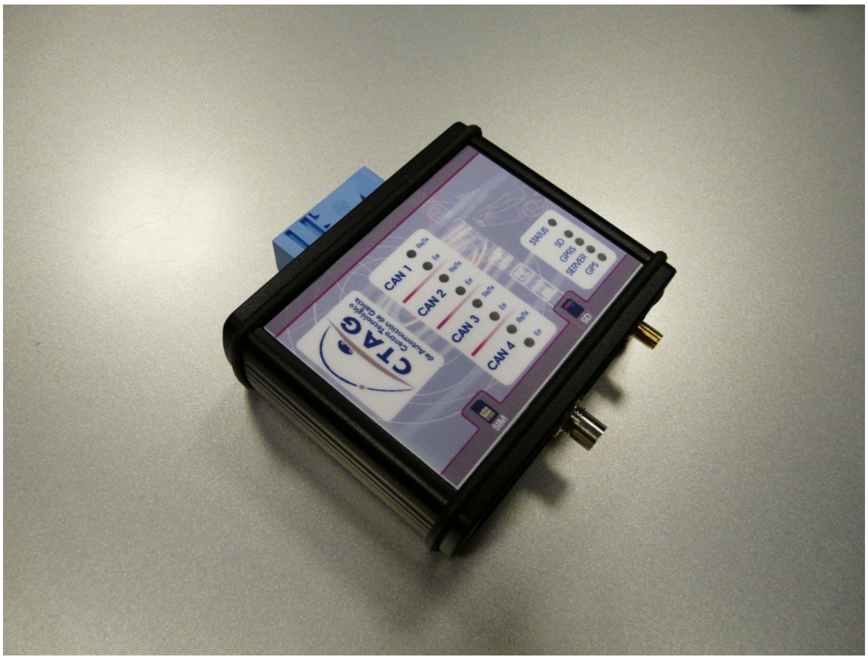

Figure 9: CTAG's Data logger II. Each CAN-channel has two LED, further information is given with five status LED.

Figure 10: German-1 VMC - Placement of DAS inside glove compartment (Ford) and inside truck cabin (MAN).

Figure 11: Video logging system at French VMC with four cameras, video capture and car pc.

Figure 12: French VMC - Placement of extra sensors and cameras.

Figure 13: Swedish VMC. Placement of DAS including camera and extra sensors.

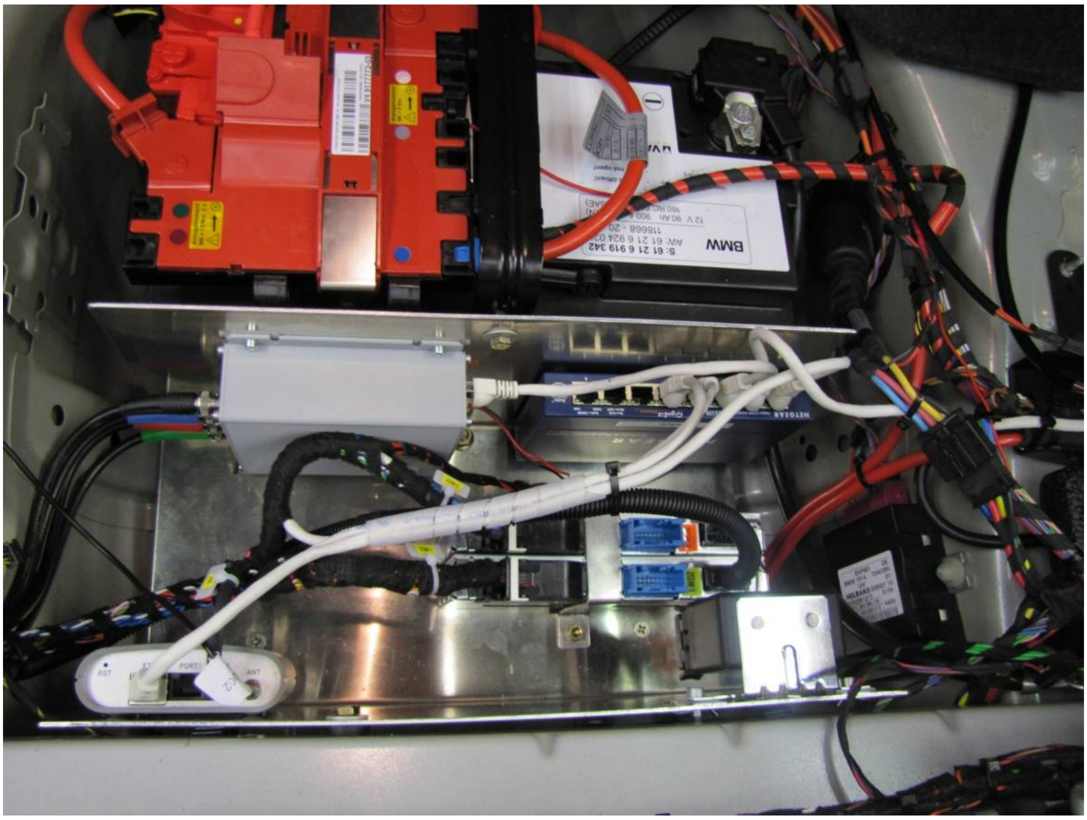

Figure 14: Placement of DAS in a BMW vehicle.

Figure 15: Placement of DAS in a Daimler vehicle.

3.1 Data management

From the time the data is collected by the DAS in the vehicle, the data will go through extensive processes before finally it is ready at the VMC centres for analysis. First, the data needs to be moved from the vehicle to the VMC centres. The choice of data transfer method depends on the size of the data collected in the vehicle. If only CAN-data is collected, then the data is transferred wirelessly using GPRS. Since wireless transfer is often not reliable, the CAN-data is also stored temporarily in an SD card in the DAS. When video data is also collected, then disk pick-up is the chosen method to transfer the bulk data. In such case, 1) the data is first saved into a hard-disk in the DAS and the disk will be collected by OEM personnel at certain times (for example, when the disk is near full); 2) status data is sent wirelessly after each trip to the VMC centre for FOT monitoring purpose.

Generally the flow of data can be summarised as shown in Figure 16. Once the data arrived at the VMC centre, the data is processed rigorously (see Pre- and Post-Processing in Figure 16). Basically, all predictable computations are done at the pre-processing stage. The quality of the data is checked several times throughout the pre-processing stage. These checks include checking for missing data, values out of range, and wrong dynamic behaviour. Generally, three levels of quality checks are performed, i.e., per sample, per measure and per trip.

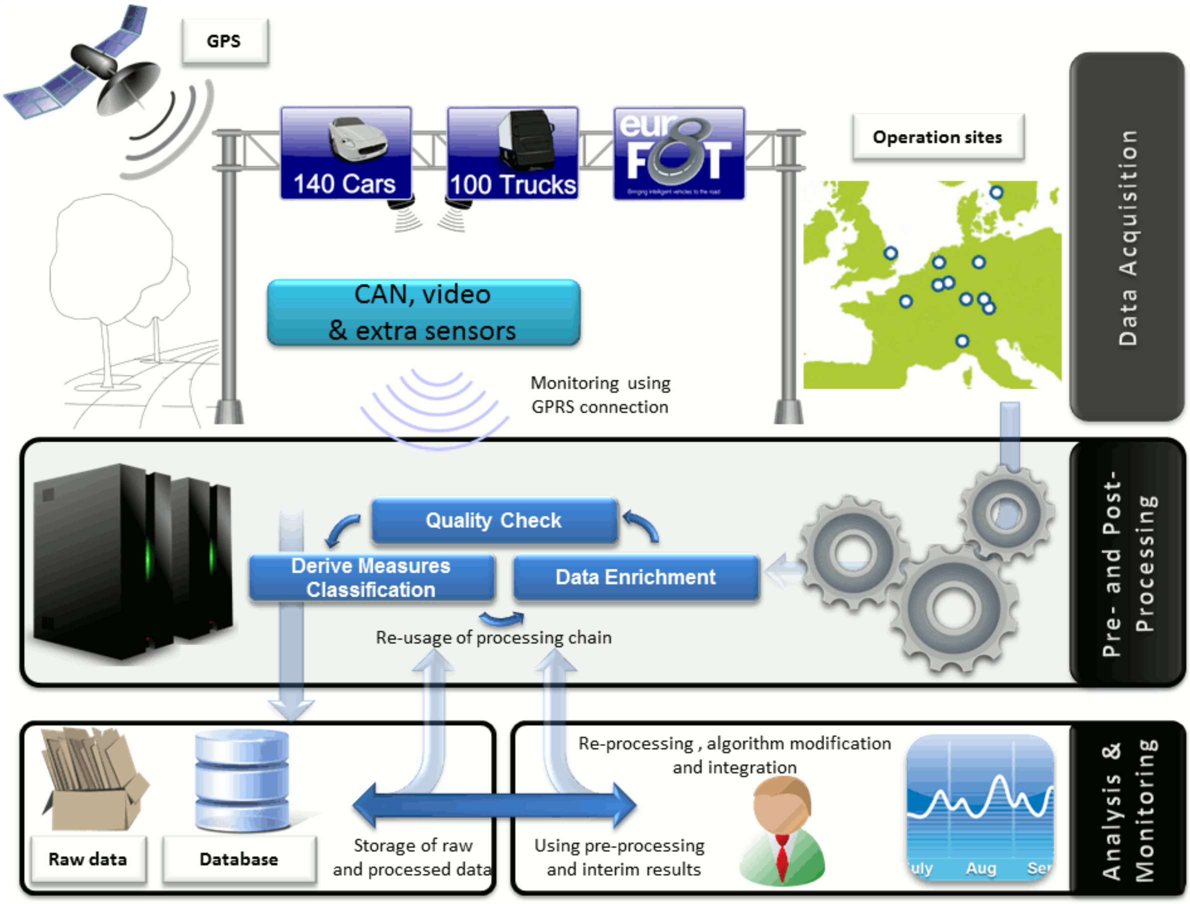

Figure 16: Flow of data from vehicles to database and storage.

A number of checks and comparisons are also done between the processed data and the corresponding raw data. This is to ensure that the pre-processing does not cause any significant alteration of the measure data. The post-processing includes adding/updating driver ids, manual annotations, and new derived measures. The details of the pre-processing and post-processing as well as the quality checks done can be consulted from deliverable D3.3.

3.2 Field test

The execution of the euroFOT's field operation test was the responsibility of the SP5 with the following objectives:

- Organize, and coordinate the vehicle management centres (VMCs) and guarantee a smooth operation of the tests.

- Secure the provision of the logging and sensing equipment and data quality management defined by SP3. Include SP2 for the development of the installation handbook

- Guarantee that the methodologies and the experimental procedures defined in SP4 are applied during the tests i.e. test matrix and experimental design which will include both a Baseline and a Treatment period for each function.

- Provide the operational platform to SP6 in order to perform the necessary study to evaluate the functions defined in SP2 i.e. verify their respective hypothesis (e.g. traffic safety, efficiency, acceptance) as well as other aspects (e.g. practical constraints in recruitment of vehicles with the functionality needed).On the operational point of view SP5 will:

    - Guarantee all drivers, equipment and quality assurance procedures in place for efficient, scientifically rigorous data collection.
    - Organize for all driver interactions throughout the studies
    - Provide guidelines for incentives and vehicles handover
    - Operate Hotline for drivers and link any requests to the respective VMC
    - Organize and handle driver focus groups, questionnaires and workshops workshops to obtain data about safety and efficiency
    - Organize user workshops (defined together with SP4 and SP6) to obtain acceptance and usage feedback

The preparation of the field tests revealed a complex test environment with different challenges in every operation site of the project. The different aspects of the challenges can be listed as follows:

- The multitude of tested functions and their various configurations in the test vehicles

- The multitude of test vehicles including personal cars and heavy duty trucks

- The different selected data acquisition solutions adapted to the requirements of the tested functions

- The different instrumentation level of the test vehicles

- The geographical coverage of the project including 6 countries of the European Union

- The different economical and legal challenges impacting the driver acquisition

During the piloting phase of the project it became obvious that the project had to deal with the above described test environment and needed to adapt its operational structure according to the challenges

3.2.1 The test vehicles and their instrumentation level

The recruited and used vehicles in the project had to comply with the requirements of the project: to have the tested functions. In addition to this the selected recruitment method also impacted the composition of the global euroFOT fleet.

All of the vehicles (except in the case of the Italian VMC where a control group was recruited with no LDW in the car) were equipped with advanced driver assistance systems (ADAS).

CEESAR, French VMC: SRS (Cruise control and Speed Limiter)

Renault Clio  Renault Laguna 

Ford, Operation Center 1, German VMC: ACC, FCW and 2 CSW vehicles

Ford Mondeo  Ford Galaxy  Ford S-Max 

MAN, Operation Center 1, German VMC: ACC, FCW and LDW

MAN TGX 

VW, Operation Center 1, German VMC: ACC and LDW

Volkswagen Passat 

BMW, Operation Center 2, German VMC: SafeHMI

BMW 530 

DAG, Operation Center 2, German VMC: SafeHMI

Mercedes E-Class 

CRF, Italian VMC: LDW

Lancia Delta 

Volvo AB, Swedish VMC: ACC, FCW and LDW plus 50 vehicles with FEA

Volvo FH12 

VCC, Swedish VMC: ACC, FCW, LDW and BLIS

Volvo V70  Volvo XC70 

Figure 17 Overview of deployed FOT vehicles in each VMC.

The fact to have so many different types of vehicles meant to have many different data logger installation procedures for all vehicle types participating the project's field test. Appropriate procedures were developed and tested in the pilot phase. Installation handbooks were developed by the project partners for the large scale installation of the data acquisition systems.

During this procedure the project had to respect the development/testing rules of the automotive OEMs in terms of quality insurance and testing procedures.

In addition to the above situation different instrumentation levels were planned. In the French VMC CEESAR has equipped 5 vehicles with video recording capability. The equipment, as well as installation and calibration procedures were thoroughly tested with the 5 CEESAR owned vehicles, before starting to instrument participants' vehicles.

After this validation phase, the 5 same vehicles were fitted with the video logger, cameras, eyetracker and lane tracker. During the experiment, those five vehicles were rotated between drivers, during three two-week periods. This allowed collecting an equivalent amount of richer data for each participant

The German VMC operation center 1 vehicles were equipped with CAN only data loggers while the German VMC operation center 2 and the Swedish VMC vehicles were all equipped with video recording devices (Figure 18).

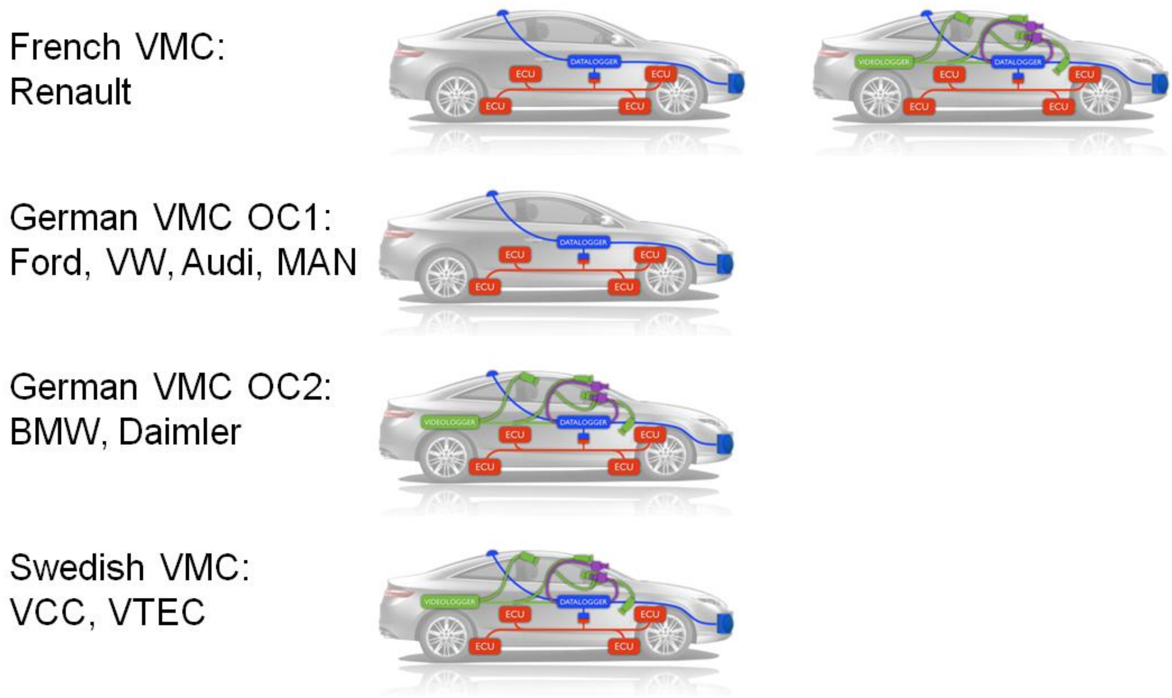

Figure 18: Overview on instrumentation for each VMC

3.2.2 Geographical extension of the project

To recall the geographical extent of the project the following details are added and updated from deliverable D5.1 Description of the VMCs and common guidelines:

Vehicle Management Center: The VMC is an entity which can be seen as the **overall management** entity. Each VMC represents a country in Europe. There are 4 VMCs in the euroFOT project: French VMC, German VMC, Italian VMC and Swedish VMC.

Operation Center: An operation center is a sub-entity of a VMC dealing with the **operational management**. One operation center groups together different operation sites because they are connected through one or several aspects of the organization (common data servers, common data acquisition system, common organization). There are 5 Operation Centers: France, Germany Operation Center 1, Germany Operation Center 2, Italy and Sweden.

Operation Site: An operation site is directly linked to an OEM. It deals with the **site management** and the OEM specific aspects (Maintenance, Dealers…). One operation center may have one or more operation sites depending on the location of the most involved dealerships. Driver acquisition was, however, not restricted to the operation site. There are at least 10 Operation Sites identified: CEESAR-West Paris, Ford-Aachen-Köln, MAN-Munich, VW-Wolfsburg, BMW-Munich, Daimler-Sindelfingen, CRF-Torino, Volvo-The UK, Volvo-The Netherlands, VCC-Gothenburg.

3.2.3 Driver acquisition

The address of the buyer of a series car is often not available to the manufacturer or its use to directly inform and acquire the buyer is prohibited. Several different approaches were considered and used:

Direct customer contact through car dealers (Ford, Italian VMC, French VMC). The dealer receives information material and speaks to customers. Direct customer contact through fleet operators (VTEC, MAN). The fleet operator owns the vehicles and employees drive the vehicle. Direct customer contact proposing car leasing with reduced leasing fee as incentives (Daimler, BMW). OEMs own fleet with employees (VCC, VW)

Each solution had advantages and disadvantages: the dealerships of the different car manufacturers provided an efficient way to deal with the recruitment as they are in contact with the customers having the necessary means and experience to talk to them. Nevertheless, it proved extremely difficult to recruit drivers during the sales process (i..e new car buyers which would then add the functions to their order). The reason might be that the flow of sales talk is not improved by the many questions which participation in an FOT will bring up.

The driver acquisition provided a heterogeneous pool of drivers: ordinary drivers recruited usually in the dealerships using their vehicles in their daily life and professional drivers recruited mainly using heavy duty truck fleets using the vehicle for work.

The take-rate of vehicles equipped with the functions was low due to the weak economic situation, which resulted in less vehicles sold and with less options installed.

3.2.4 The solution to address all the challenges

During the preparation phase of the project the following organizational structure has been developed and presented in the deliverable D5.1 VMC description and common guidelines:

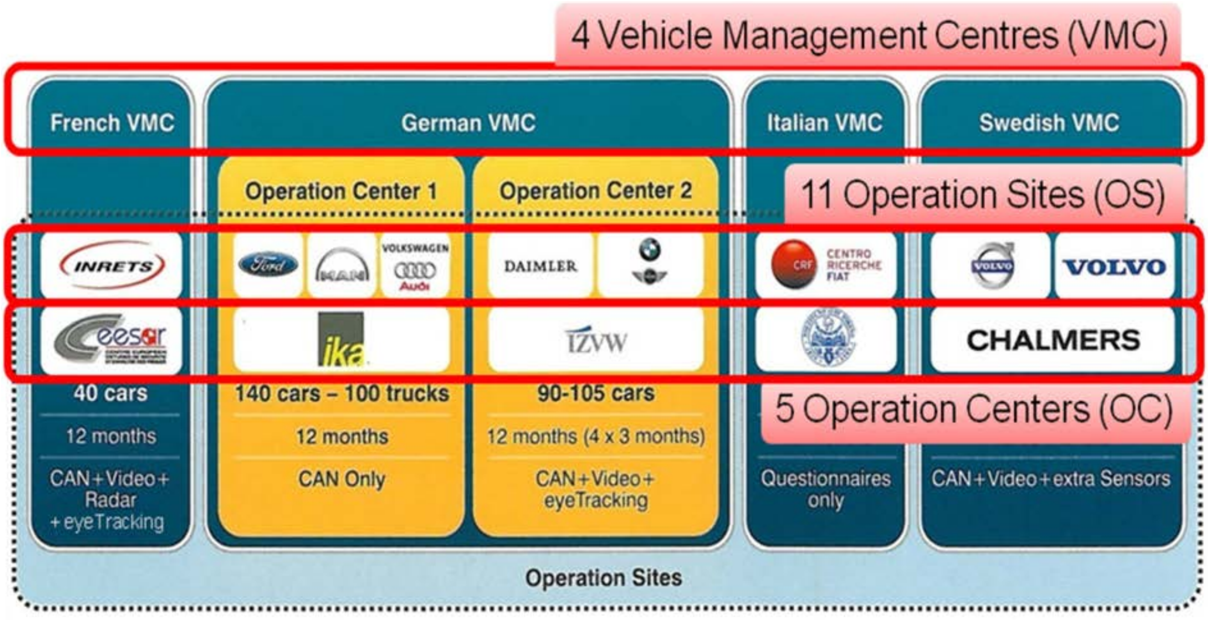

Figure 19: Operational structure of euroFOT

The four Vehicle Management Centres are divided as follows:

French VMC: The cars from Renault are instrumented by CEESAR and data analysis is in cooperation with IFFSTAR.

German VMC: In Operation Centre 1 vehicles from the brands VW/AUDI , FORD and MAN , are instrumented. Data is received and analysed through IKA. Operation Centre 2 vehicles from Daimler (DAI) + BMW are analysed through IZVW.

Italian VMC: The vehicles from Lancia are coordinated through Centro Ricerche Fiat (CRF) and results are analysed through Politecnico di Torino (POLI).

Swedish VMC: Vehicles are coordinated by Volvo Technology (VTEC) and Volvo Cars (VCC). The data is analysed through Chalmers.

At each of the 4 VMCs one to four OEMs contribute to the FOT by providing test vehicles. Due to several organisational reasons (e.g. functions tested, DAS used, geographical location), the German VMC with its five contributing OEMs is divided into two operation centres. At each of the VMCs and operation centres, different ADAS are of special interest. They are thus evaluated in the given operation sites (see above).

The scheme in Figure 20 describes a common structure that the Vehicle Management Centers adopted in order to facilitate the operation of the FOT.

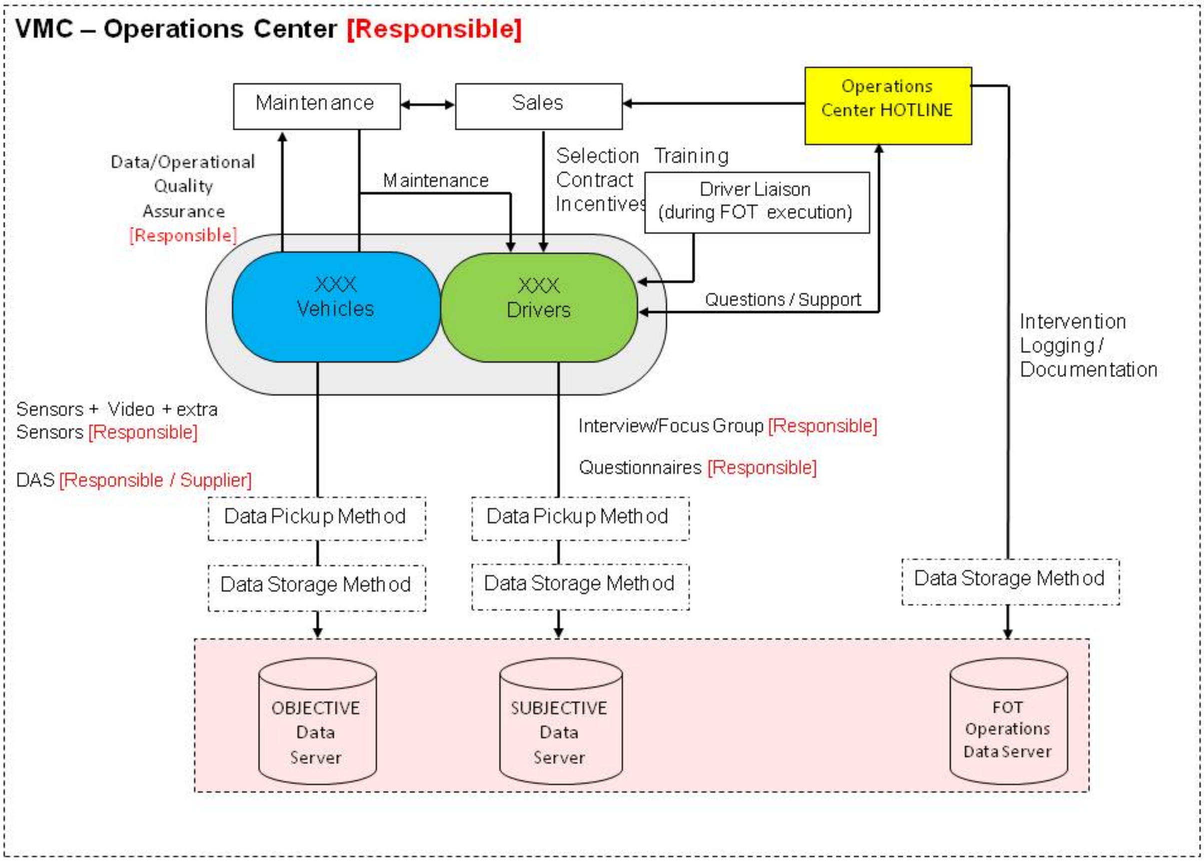

Figure 20: Common structure for Vehicle Management Center

Depending on the specifics of each VMC, there may be some variations but the main structure is about the same. This representation also helped to check that each part of the organization is properly organized with a responsible unit for each task.

3.3 Lessons learned

3.3.1 Development, planning

Tested functions with low market penetration need strong marketing effort since the beginning of the project. This involves direct involvement of the car manufacturer's marketing department: with their direct support the identification of the users is quicker, simpler and more efficient.

OEMs participation is needed while establishing the experimental design and the recruiting policy. As the OEMs and their marketing department have direct information about their costumers it is obvious that they can help the Field Operational Tests' management to identify and to develop the most appropriate methods.

3.3.2 Organisation

The VMC has to be the central element in the project organisation. The VMCs structure gives considerable independence on the execution of the FOT operation. This may also help splitting the project into smaller FOTs (subprojects) which could make sense in cases where each subproject studies one function only or use different methodology or apply different acquisition method.

Regular face-to-face meetings at the VMCs turned out to be a very effective tool to synchronize development and management efforts. In addition to this the meetings help to understand the difficulties in a very heterogenic test environment and provide a common platform of discussion for the consortium members.

3.3.3 Legal and ethical issues

The automotive manufacturer's product warranty can be made void if a customer connects unauthorized equipment to the vehicle. The extent of installation work and clearing of the DAS should be discussed prior to the recruitment process to avoid later issues with the participant's vehicle. Several months should be planned to clear all issues involving the service department of the vehicle manufacturers.

Because of the privacy issue of the video recordings, many drivers had doubts about giving their permission in consent forms. Clear communication is necessary with the recruited drivers explaining the restrictions on the use of recorded videos and the publication of the results.

The ownership and the future use policy of the gathered data should be defined in very early stage of the project.

Filming in areas where filming is restricted (e.g. company yards, military areas) needs to be considered. In euroFOT the use of the videos is restricted (e.g. in contrast to large scale web-publication of area scanners) and no instance of forbidden filming has been found.

3.3.4 Piloting related issues

The piloting phase as one of the most critical part of the project should be scheduled much longer in order to test all aspects of the tests including recruitment procedures, hardware and software issues. It should consider the testing of all FOT related procedures: driver and vehicle recruitment, incentives, questionnaire handling, data gathering and uploading, data pre-processing, equipment installation and de-installation.

Pilot tests must start after most of the technical issues were settled and resolved as well as after having all organizational procedures have been developed and deployed. It is important to understand that the pilot phase should test the whole experimental design of the FOT and not only the physical implementation of the test equipment. Thus the pilots should provide feedback concerning the whole experimental design and the practical issues.

3.3.5 Driver recruitment and vehicle acquisition

The driver recruitment procedure should be made in close cooperation with the marketing and sales departments of the OEMs. They have all the necessary tools and experience to identify and to handle vehicle customers. In addition to these contact materials such as leaflets, brochures, videos for drivers recruiting are very important to maximize response rate and to keep the participants motivated to follow the project until the end of the tests. For this it is very important to communicate the goal of the whole project during the recruitment campaign using the available dissemination materials, communication channels (newspaper articles, TV spots, interactive web sites with news flash service, etc.).

The driver recruitment and the necessary technical modifications on the vehicles should be defined prior the pilot tests as these have big impact on the availability of participants. The pilot will then provide feedback on the efficiency of the selected recruitment methods and the complexity and time consumption need of the technical modification. With the result of the pilot these may be adjusted to better answer to the needs and requirements of the field operational tests. A solution is to use vehicles owned by the manufacturer and used by the their employees which simplifies the installation of the test equipment much easier than in case of the customer cars recruited either by the dealerships or by the marketing department of the OEM.

The use of customer vehicles is acceptable only for CAN data logging, because for additional sensors and high-end data logging either the effort for designing and ensuring a traceless de-installation is considerable or the modifications to the vehicle are unacceptable for normal customers. A new vehicle model for field operational test, recently released by OEMs, should only be considered if the production vehicles are already available before the pilot phase.

In case of low take rates of features to be tested large efforts are needed to reach geographically widespread participants. Because of the large effort need a multiple approach may be applicable: OEM's market department, fleet operators, dealerships. Contacting big dealerships may be useful to reach lots of customers reducing time of communication and facilitating installation efforts while small dealerships might have more direct contact to motivated customers which may results fewer dropouts.

Common international driver recruitment coordination can however improve sample representativeness for European driving population as a whole.

Financial issues: incentives and the related taxation represent also an important point. Because of the heterogenic legal (and taxation) environment all VMCs had to deal with the incentives payment independently. There was no common solution. Some operation site used reduced leasing fee, some used fuel vouchers, etc. Where incentives are given in cash, tax-regulations have to be considered. Also the amount of work involved to find the appropriate accounting procedure is considerable. In addition to this the planned amount of incentives was not sufficiently motivating in euroFOT and should be calculated much higher.

The cost of cabling can include connectors which are not available in small quantities, thus rather small parts of a cabling cost much.

3.3.6 Subjective data gathering

The subjective data gathering of the FOT is also an essential part providing the vision of the participants on the tested systems. To achieve high answer rate it is very good to duplicate or offer multiple options for filling in questionnaires (i.e. hard copy and electronic copy). It is very important to develop and improve a web-based survey tool for data collection in order to save time and prevent data transcription mistakes. Another way to increase the answer rate is to centralize the subjective data gathering implementing a multilingual data collecting tool which can be is very useful in order to simplify global subjective data collection.

During the preparation of the questionnaires which are to be used in different countries, the translations and related validation should be done from the very first draft.

3.3.7 Hotline, driver liaison

During the FOT operation phase the driver contact turned to be crucial: hardware and software issues can be solved quickly if the VMC has proper driver liaison procedure and an organisation implementing these procedures: introducing the “customer care centre” of the dealerships is recommendable, due to the professional customer handling capabilities.

It is very important to get in touch frequently with drivers, in order to ensure data collection, but this has to be done on case by case, since some driver could be disturbed by this.In euroFOT the project was managing a very heterogeneous test environment with many vehicle types, several selected data loggers, many tested functions and their various combinations. The geographical distribution of the operation sites and as consequence the different legal environments led to a very complex situation to deal with.

4 Methodological approach

The goal of SP6 was to assess the potential impacts of the Advanced Driver Assistance Systems (ADAS) tested in the euroFOT project. Data analysis methods were described for each of the work packages to be undertaken within SP6 of the project: User Acceptance and User-Related Aspects Evaluation (WP6300); Impact Assessment (WP6400) and socio-economic cost-benefit analysis (WP6500).

The impact assessment translated the effects found in the trips made by the equipped fleets in the FOT to the EU level. In other words, the effects found in the FOT data for certain situations or for certain groups of drivers were scaled up to both a larger population and geographical scope. This led to an assessment of the potential effects of the evaluated ADAS if they would be widely deployed in the vehicle fleets across Europe. Not all functions have impacts on safety, efficiency and the environment. For a number of systems no significant impacts were found and therefore such assessments were not performed for those systems.

During the FOT both subjective data (derived from questionnaires and driver interviews) and objective data (derived from the vehicle CAN and video recordings) was gathered. Data processing and data analysis were carried out and performance indicators and situational variables were calculated.

In the impact assessment, hypotheses were tested and research questions were answered. This was done using performance indicators, situational variables and events. These hypotheses, research questions, performance indicators, situational variables and events were defined in earlier work packages in SP2 and SP4. The impact assessment methodology was based on the approach defined in the EC-funded Field opErational teSt supporT Action (FESTA) project, which was adapted to the specific conditions and needs in euroFOT.

Based on research questions, derived in SP2, hypotheses for each aspect of the analysis (e.g. safety, traffic efficiency, driver-related aspects) were developed in SP2 and revised and prioritised in SP6. The hypotheses were used as a basis for identifying the required performance indicators and the corresponding signals to be collected by the data acquisition systems installed in the vehicles. For example, to test the hypothesis “ACC decreases the number of incidents”, the performance indicator “number of incidents” is needed. Depending on the definition of the incident, the signals to identify incidents have to be defined (e.g. vehicle speed, distance to forward vehicle, deceleration). This process is depicted in Figure 21. In practice, the process was more iterative than the figure suggests.

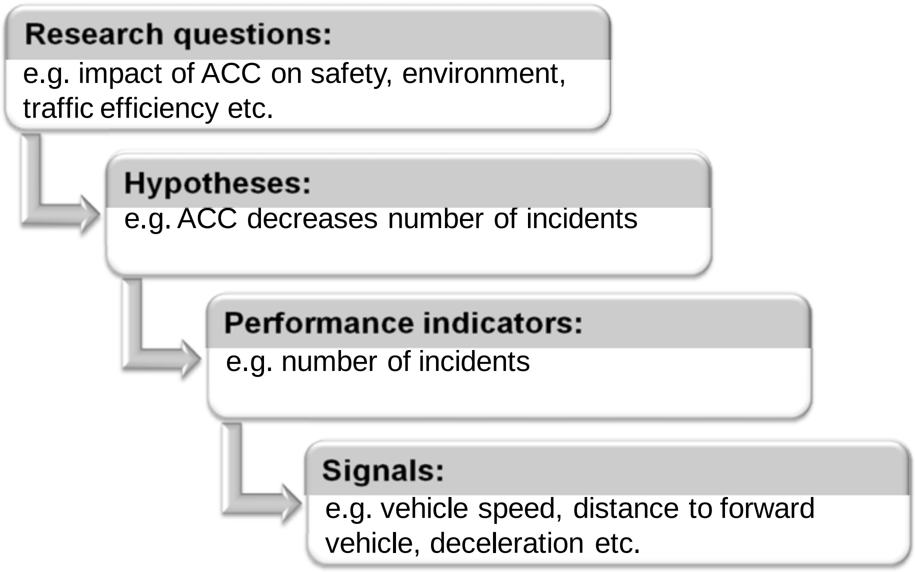

Figure 21: Definition of hypotheses and required signals to be collected from the vehicles

Relevant events and situational variables were also taken into account when defining and testing the hypotheses.

4.1 User acceptance and user related aspects

It is necessary to evaluate the impact of the ADAS on driver behaviour and to evaluate how acceptable they are to drivers. These evaluation activities were undertaken in WP6300 (User Acceptance and User-Related Aspects Evaluation) of the euroFOT project, and therefore information on the final results can be found within D6.3.

The assessment of user related aspect is widely based on hypotheses testing, including also data processing and PI calculation. Harmonised data analysis methodologies for the evaluation of the impact of the functions on driver behaviour and acceptance were developed by a Harmonisation Task Force created in WP6300.

The aims of the user acceptance and user related aspects analysis were:

- Impact on driver behaviour
- Impact on driver workload
- Driver acceptance of the function (defined as usability, usefulness, and social acceptability)
- Trust in the function
- Function usage
- Exposure

All euroFOT functions were tested in relation to user acceptance and user related aspects.

The VMCs performed the implementation and analysis of all the functions: that is, testing of hypotheses using the methodology described in this chapter. Global result integration and harmonisation was carried out by CTAG as WP6300 leader.

4.1.1 Research questions and hypotheses

For each function, the following aspects were analysed:

- Acceptance
- Trust
- Driver workload
- Usage of function
- Usability
- Usefulness
- Social acceptability
- User practices

Moreover, the following research questions were answered per each function:

- What features of the function, in terms of usability (e.g. accessibility, readability, controllability, compatibility while driving) influence acceptance?
- What features of the function, in terms of usefulness, influence user acceptance?
- Does acceptance change with experience?
- Does trust in the function change with experience?
- Do drivers find the function more usable with experience?
- Does usage of the function change with experience?

The hypotheses to be tested rely both on subjective data and objective data. The subjective data was collected via a harmonised, purpose-designed, questionnaire developed in Definition of performance indicators (WP4300), with input from WP6300. Some additional questions were developed by some VMCs, to answer some of their own specific research questions. The Italian VMC only used subjective data to test its hypotheses; the other VMCs relied on both objective and subjective data to test their hypotheses.

Some of the outputs of the behavioural impact assessment were used in WP6400 and WP6500 to support the assessment of safety impacts, traffic efficiency impacts and environmental impacts.

4.1.2 Testing of hypotheses - Objective Data

The chosen approach to analyse the objective data within euroFOT was based on the FESTA methodology. The data collected during the collection phase was stored in a database. In that database the collected measures as well as the additional information that have been determined during the processing phase were stored and could be used for performing the analysis. The additional information was generated by means of different processes. For instance, information on road type, speed limit and curvature were not available on the vehicle’s CAN bus. This information was determined by means of the collected GPS signal. Furthermore, relevant events (e.g. incidents, lane change manoeuvres) as well as situational variables (e.g. weather condition, lighting condition or traffic density) were collected by means of processes, which search the data with respect to defined patterns (e.g. exceeding a certain deceleration value as a trigger for detection of incidents).

The approach for data processing varied between the VMCs – especially the processes for detection of relevant events and situational variables, due to the different data acquisition strategies. The key steps required for testing the user acceptance and user related hypotheses were the following ones:

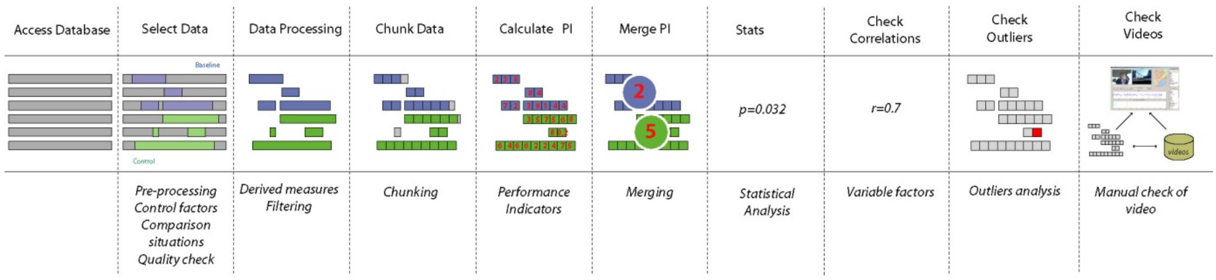

Figure 22: Process for hypothesis testing using objective data

Four general ways of preparing the data for the analysis can be distinguished, depending on the focus of the analysis:

- Comparing the average state of some variable like speed or headway in baseline (no function available) and treatment (function available)
- Comparing how often a particular type of event or condition occurs in baseline and treatment (like the frequency of near crashes)
- Studying if function usage changes over time, e.g. if the driver uses a function more often once it has become available
- Studying whether function presence influences some other aspect of driver behaviour, such as the proportion of time spent doing secondary tasks

The steps described in Figure 22 are:

- **Pre-Processing:** Pre-processing all calculations which need to be carried out on the raw data level before one can start selecting data for the hypothesis testing. This involves procedures for deriving measures, applying frequency filtering on signals, etc.
- **Comparison situations:** Hypothesis testing principally involves some form of condition comparison. Therefore one must decide on the conditions which are to be compared.
- **Controlled factors:** After the conditions have been defined, one needs to define the controlled factors for which the dependent measures, or performance indicators (PIs), are compared. For example, to analyse whether the average speed changes when LDW becomes available, one must decide which treatment data to include. Should everything in the treatment phase be considered (i.e., is it enough that the driver has the function in the vehicle) or should one look only at the portion of the data where LDW actually can be activated and used (i.e., speeds above 60 km/h); and, if so, should one be even more restrictive and only select data where the road markings are sufficiently visible for reliable lane tracking? Moreover, it should be discussed whether data should be organised *per-driver* or *per-vehicle* if this information is available.
- **Quality checking and filtering:** Once these definitions are in place, it is needed to check the data quality according to some criteria. Before computing the performance indicator some filtering may also be needed, for example to eliminate high-frequency components from a signal.

- **Chunking:** When definitions and quality measures have been addressed, and if the hypothesis to be tested is looking for some average difference, then it is time to chunk the data. Chunking means that the identified segments (i.e., those selected by applying control factors to the dataset) are divided into chunks of data of equal size. Chunking is applicable to all PIs which are based on time series data. Chunking guarantees that PIs are calculated on samples of equal size, which reduces variability. It also provides a simple way of keeping track of how much data per condition was included in the analysis.
- **Performance indicators:** When the above has been completed, it is time to calculate the PIs, such as average speeds or event frequencies. Also, if the data was chunked, it is necessary to decide if and how it should be merged (so far there's one data point per chunk). Merging depends on the different hypotheses and dependent measures, and more than one merging procedure can be specified, e.g. by averaging over different time windows.
- **Statistical analysis:** The final step is statistical analysis of whether there is a significant difference between the comparison conditions. Depending on the setup of the procedure above, various statistical models may be validly applied (ANOVA, t-test, Mixed Generalised Linear Models, etc).

More detailed information on this methodology can be found on D6.2.

4.1.3 Testing of hypothesis - Subjective Data

In the case of analysing subjective data, the following steps were executed:

- **Data needs:** A purpose-designed questionnaire was developed in WP4400, with specialist input from Work Packages 6300 and 6400. The steps involved in developing the questionnaire are described in D4.2. The questionnaire contains all questions necessary to yield data that can be used to test those hypotheses.

  The questionnaire was divided into five parts (referred to as Screening and Times 1, 2, 3 and 4), each part of which was administered at different times during the FOT. The Screening, Time 1 and Time 2 questionnaires were administered during the baseline period, prior to function's activation (Screening and T1 at the beginning of the baseline period, and T2 at the end of the baseline period). The Time 3 and 4 questionnaires were administered during the treatment period, after function activation (at the mid and end points of the treatment period, respectively). More information on these questionnaires can be found on D6.2
- **Preparatory activities:** Special preparatory activities must be undertaken prior to analysis of the questionnaire data. These preparatory activities consist of:
  - **Data quality:** The questionnaire must be checked to ensure it contains all questions required to test all hypothesis with no missing data, while being clear and understandable. Moreover, data should be in a format suitable to be analysed. Administration of the questionnaire must be standard among all VMCs, and it must be checked that the drivers responded appropriately to the questions and that there are no errors in coding these answers, as well as in the database entry. Poor quality data should be cleaned.
  - **Data coding:** For open-ended questions, harmonised categories for the answers must be derived so that, as for closed response questions, the data is coded in the same way across VMCs.
- **Analysis plan:** A common analysis plan for the subjective data was developed to ensure that VMCs were able to analyse and report the data derived from the questionnaires in a coherent and consistent manner.

Within an Excel spread sheet, each hypothesis was assigned a separate worksheet and, using dummy data, each VMC was provided with guidance as to which items in the questionnaire should be "dropped into the worksheet" to test each hypothesis. Where necessary, guidance on the coding of responses was provided, and an explanation regarding the items was provided in an associated document.

With regards to statistical testing, the coding and analysis of the standardised scales within the questionnaires replicated that reported by the original authors.

Given that the questionnaires were administered at various time-points throughout the trial, some items could be analysed using repeated measures techniques, using time as an independent factor.

Non-standardised items were typically analysed descriptively.

- **Reporting plan:** A harmonised approach has been developed for the reporting of the findings that derive from the testing of hypotheses relying on questionnaire data. In effect, the VMCs can copy and paste their data into the Excel spread sheets developed and the graphs and summary statistics are then automatically generated. This ensures that all the graphs across functions and VMCs are in the same format, for ease of interpretation.

Again, for a more detailed definition of the methodology followed for subjective data, deliverable D6.2 should be consulted.

4.2 Safety

The purpose of the safety impact analysis was to assess the extent to which the functions being evaluated in euroFOT can potentially be expected to alter the current crash populations in the EU-27 region in terms of accidents, injuries and fatalities. The CBA in WP6500 estimated the costs and benefits of deploying these functions on the EU-27 level.

At the core of the safety impact analysis was a set of hypotheses that were tested on the empirical data. (D6.2, Annex 2A). However, as these hypotheses were generally of the type "does X change in treatment compared to baseline?" they needed to be placed in the wider context of a full benefit analysis in order to help answer the question of whether safety would be improved if the evaluated functions were widely deployed in the EU-27 region. Therefore, in this chapter, the process for calculating safety benefit estimates for the safety functions in euroFOT is described. The process has three main steps:

(1) **Defining the target crash population:** This involves finding out from national crash data in the countries where the functions are being evaluated, how many function-relevant crashes (i.e. crashes that the function could potentially help address) occur on an annual basis.

(2) **Identifying changes in safety related measures between baseline and treatment:** This involves testing a number of hypotheses on how various safety related metrics might change between baseline and treatment in the collected data for the functions evaluated.

(3) **Interpreting what any significant changes in these metrics mean, in terms of a generalised safety impact estimate on the EU-27 level:** This means estimating, based on the identified changes in step 2, of the influence on accidents, injuries and fatalities in the national crash population if the evaluated functions were to be nationally deployed, and then extrapolate those results to the rest of the EU, i.e. trying to project what would happen in the full EU-27 driving population if the functions were deployed EU-wide.

The general process is illustrated in Figure 23. Note that while the details on how this process is applied varied somewhat between VMCs. They depend on the function being evaluated and how data collection was set up. However, the general procedure was intended to be similar across VMCs. In the sections following below, a general description of what the three steps above mean will be given. After that follows a section on concerns with respect to this methodology.

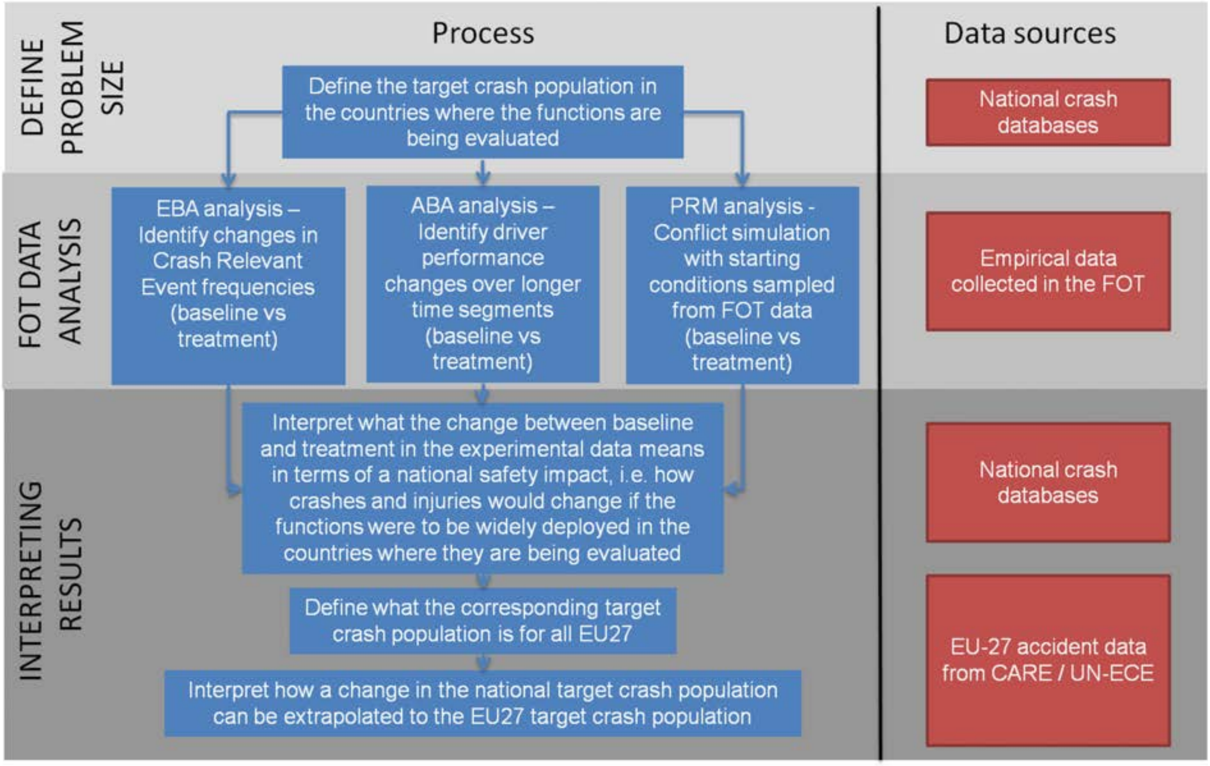

Figure 23: Overview of the safety impact process and the data sources used

4.2.1 Step 1: Defining the target crash population

The first part of the safety benefit analysis is fairly straightforward. It involves defining the target crash population, i.e. the set of crashes (including the associated set of injuries and fatalities) which a particular function may be capable of, addressing in some way. For example, for Forward Collision Warning (FCW) this is the set of rear end accidents that occur within the function's operational scope (certain ego vehicle speeds and approaching speeds to lead vehicle). For Lane Departure Warning (LDW), this is the set of crashes which start with an unintentional lane departure, and again, which are within the functions operational scope (visible lane markers, above certain ego vehicle speed, etc).

Once the target crashes have been identified, data describing the crash circumstances was cross tabulated to identify the most typical conditions under which these crashes occur. The rationale for this step is to provide a set of filters that can be used to exclude some of the empirical data from the analysis. In principle, any ADAS driven change which occurs outside the envelope of crash typical circumstances will not affect the safety impact, since by definition no relevant crashes occur outside those conditions. Leaving those data portions out of the analysis thus strengthens the link between the empirical data and the target crash population.

4.2.2 Step 2: Identifying changes in safety related measures between baseline and treatment

The second part of the methodology is quantifying any potential safety relevant difference that the presence of an ADAS generates in the empirical data, i.e. quantifying any changes in crash risk between baseline (no ADAS) and treatment (ADAS present). For this purpose, a partially new methodology was developed in euroFOT.

In terms of how the methodology is set up, it is first important to recognize that it was known before the project started that the number of actual crashes that would occur with the euroFOT fleet of vehicles would be very limited. Even when hundreds of drivers are being observed during a full year or more, the statistical likelihood of a crash occurring is so low that it is uncertain whether there would be any police reported crash events in the data at all.

This meant that the most direct measure of change in crash risk, i.e. the number of crashes which occur with and without the ADAS, is not available, at least not in sufficient numbers to reliably quantify a difference between baseline and treatment. Hence other indicators of change in crash risk had to be defined, such as the frequency of safety critical events or changes in driver behaviours that are known to be related to crash causation. In other words, to determine whether a particular ADAS is successful in influencing a certain crash type, one must first have an understanding of why that crash type occurs. When that understanding is in place, a measure of change that captures the function's impact on that particular crash causation mechanism(s) can be defined.

For example, many rear end crashes are thought to occur due to unexpected lead vehicle braking while the driver is visually distracted from the forward roadway [34]. In relation to this crash causation mechanism, FCW can be understood as a tool for interrupting the driver's state of visual distraction and redirecting his/her attention to the forward roadway and the braking of the lead vehicle. If FCW is successful in this regard, one would expect e.g. a decrease in the number of panic braking events when drivers are using FCW. The frequency of panic braking events can therefore be used as an indicator of change in crash risk due to the presence of FCW.

To facilitate the process of identifying changes between baseline and treatment in the collected euroFOT data, a number of hypotheses on how various safety related measures may be impacted by ADAS presence were formulated (D6.2, Annex 2A). Testing whether these hypotheses hold (or not) essentially forms the core of this second step of the methodology.

Depending on the function analysed and the hypothesis to be answered (i.e. how the function's influence on some crash causation mechanisms is conceived), three principal ways of doing the analysis were applied accordingly.

1. Events Based Analysis (EBA)

The first analysis method is Event Based Analysis (EBA). Here the aim is to find out whether the frequency of safety critical driving situations changes when a safety function is made available to the driver. The basic principle of EBA in a FOT context is to identify relatively short time segments (events), thought to be predictive of crash involvement, and then compare the frequency of these in baseline (no ADAS present) and treatment (ADAS present). Examples of events are actual crashes, as well as situations where the driver performs an evasive manoeuvre, i.e. where the distance in time and/or space from an actual crash is very small (near crashes/incidents). These events can be identified retrospectively in the driving data, together with interaction/confounding factors such as road type, speed limit, traffic conditions, other functions etc.

2. Aggregation Based Analysis (ABA)

The second is Aggregation Based Analysis (ABA). Here the aim is to identify any significant changes between baseline and treatment in aggregated continuous data, such as mean speed or average time headway. In other words, it captures the change between baseline and treatment in terms of how driving performance changes over longer periods of time. This type of analysis is primarily relevant for answering hypotheses on e.g. whether the average following distance or travel speed decreases as a function of ADAS presence.

3. Physical Risk Modelling (PRM)

The third is Physical Risk Modelling (PRM), which is a computer simulation based analysis of vehicle conflicts. In PRM the starting conditions for the simulations are sampled from the empirical data, and Monte Carlo simulations are then performed to explore a wide range of possible situation outcomes given those starting conditions. PRM can thus be said to be a simulation version of EBA, and in euroFOT it was applied to ACC, which often is hypothesised to have a positive effect on reducing lead vehicle conflicts.

Choice of method

Note that EBA, ABA and PRM are complementary forms of analysis which explore the impact of an ADAS from different angles, based on how the ADAS safety impact is conceptualised in terms of potential influence on crash causation (i.e. which hypothesis are selected to being tested for the function). For ACC/FCW for example, a potential increase in average time headway is best investigated with an ABA analysis, while a potential decrease in the number of lead vehicle conflicts is best investigated with an EBA or PRM type of analysis.

4.2.3 Step 3: Interpreting what any identified change between baseline and treatment means in terms of a generalised safety impact on the EU-27 level

The third step of the methodology is taking the quantified differences between baseline and treatment and calculating what it can mean in terms of reducing the full crash population on an EU-27 wide level. This is done in three steps: The first is to decide which of the identified differences are to be used for the actual prediction. The second is to calculate the reduction in crashes based on accident data, filtered as detailed as it is possible on national level. The third is to extrapolate those results to the EU-27 wide crash population level. Below, these parts are addressed in turn.

4.2.3.1 Step 3a: Statistically testing size and significance of identified effects

First, there is the issue of size and significance of an identified difference between baseline and treatment, or between some other comparison conditions (like at various times during treatment). To test this, many different methods are available, depending on the data analysed and the hypothesis to be addressed. For ABA data analysis, which typically becomes a comparison of means in baseline and treatment, or changes in means over time as drivers use a particular function more and more, various types of variance and regression analysis can be applied, such as ANOVA and linear regression models. The main challenges for statistical ABA data analysis are to define how baseline segments should be selected, as well as to understand what should be considered covariates and confounders.

When it comes to analysing EBA data, i.e. to compare event frequencies in baseline and treatment, also many different methods are available. The simplest form of comparison is to make a contingency table by counting the frequency of events in baseline and treatment conditions (based on some form of exposure normalisation, such as the number of events per driving hour) for each driver, to understand whether ADAS presence causes a change in event frequency. For example, consider the following contingency Table 5:

Table 5: Event frequencies

	Baseline	Treatment
Number of crashes (or safety events)	N1	N2
Km's driven (or duration)	T1	T2
Crash (or events) rate	$\Pi1 = N1/T1$	$\Pi2 = N2/T2$

The risk change due to function presence can then be quantified and statistically tested using both relative risk ($RR = \Pi1/\Pi2$) and/or the odds ratio ($OR = (\Pi1/(1\text{-}\Pi1))/(\Pi2/(1\text{-}\Pi2))$). The odds ratio approximates the relative risk when Π1 and Π2 are small.

However, a drawback of contingency tables and ordinary logistic regression is that they assume observations to be independent of each other. This assumption does not suit FOT data well in all instances, as it may contain driver-specific correlations (i.e. some drivers will experience more events than others). To study interacting/confounding factors and to account for these driver specific correlations, more sophisticated statistical models need to be applied. These models are generalizations of the linear model which have been adapted to a binary outcome, something which suits the EBA analysis division of events into baseline and treatment events well. These models include additional parameters to deal with correlations, and confounding factors are viewed as explicative variables that can be used to predict event probability.

4.2.3.2 Step 3b: Impact on the national level

If a significant difference between baseline and treatment has been established for an ADAS in terms of a risk indicator, the next step is to make an impact estimation based on that difference. This means that one has to interpret what the identified difference actually means in terms of how the target crash population can be expected to change if the function is widely deployed.

This impact estimate should first be carried out for the national level, i.e. for the country in which the function is being evaluated. Next, the national impact should be projected onto a wider EU-27 scale (Step 3c below). This two-step approach was chosen because when selecting which part of the FOT data to look for changes in (see step 1 above), crash conditions are a very relevant input. For a best fit with the data collected, it makes sense to use crash conditions from the country where each respective system is being evaluated. Once the national analysis is done, the effect of a comparable relative change in other countries can be assessed.

The national impact estimation can be carried out at various levels of detail. The least detailed approach is to apply the identified change to the whole target crash population. For example, if the frequency of FCW relevant near crashes turns out to be 20% lower in the treatment phase, one might use this to predict a 20% decrease in FCW relevant crashes and injuries if all vehicles were equipped with FCW.

A more detailed approach would be to first calculate the impact for individual conditions in the target crash population, and then sum up the total impact. For example, if there are 2000 rural and 3000 urban FCW relevant crashes in the target crash population, the analysis might find that while the near crash reduction ratio is 25% for urban environments its only 17% for rural roads. In this case, the total safety impact would be calculated for urban and rural roads individually before summing up, i.e. the potential reduction in crashes would be $(0.17*2000 + 0.25*3000)/5000 = 22\%$.

While the more detailed approach naturally is preferable, it also requires larger significant effects to be meaningful, since differences quickly shrink when the list of dividing conditions grows longer. Thus one has to balance the desired level of detail in the impact assessment with minimum requirements on how many events per condition are needed to test for statistically significant differences between baseline and treatment.

4.2.3.3 Step 3c: Impact on the EU-27 level

In this step, the national impacts as identified above are projected onto the crash population in the EU-27. To extend a national impact to the EU level, it is first necessary to map the crash shares affected by the evaluated functions on the national level to an EU-27 crash population. For example, if the national impact predictions for a function evaluated in Sweden and Germany indicate that 5% of rear end crashes in Sweden and 7% of rear end crashes in Germany would be addressed if all vehicles were equipped with this feature in these countries, then one has to take a meaningful average of these two impacts (simplest form: 6%), and calculate what a 6% reduction of rear end crashes means in terms of accident and injury reduction for all EU-27 countries.

4.2.4 Methodological concerns

In relation to the different ways of estimating a safety impact for the euroFOT functions described above, there are a number of methodological concerns that can be raised. Some of the most important ones are discussed in turn below.

First, the relationship between changes in the evaluated crash predictors, when comparing baseline and treatment, and accident involvement is not straightforward. This is a problem for all three approaches (ABA, EBA and PRM) described above. Ideally, one would select and compare only events and/or aggregate measures which are known to be predictive of actual crash involvement, i.e. where it is legitimate to infer that a particular change in what is measured corresponds to a particular change in crash frequency. If this relationship is established, any identified reduction in the treatment phase could then justifiably be used to directly predict a reduction in future crash involvement.

Unfortunately, such relationships are not fully established, at least not for FOT type data. For example, in terms of events, while hard braking may seem a plausible candidate for event selection, in the VTTI 100 car study [34] they were not able to reliably identify near-crash events in lead vehicle following situations based on hard braking alone, i.e. such braking occurred also in many driving situations which they did not think were indicative of crash risk. Similarly but in terms of aggregate measures, while a reduction in mean speed could be indicative of a reduction in crash involvement, there is no empirical base available for estimating the importance of mean vehicle speed in FOT data in relation to crash involvement. It follows that insight into crash causation mechanisms is key to the selection and interpretation of relevant measures of change between baseline and treatment.

Second, in euroFOT, a number of hypotheses on change in safety related indicators are being tested. In an ideal world, it could be hoped that all tested indicators for a particular function would point in the same direction, whether it is toward a general increase of decrease in perceived safety in treatment. However, most likely there will be contrasting findings, as well as some statistically significant and other not significant results. This means that an important part of the impact assessment is to find a way to tell the overall story of each function’s potential impact, given how its particular set of indicators come out from the empirical data analysis. Presenting the results of the indicators only will not suffice for a full impact analysis; some form of an integrated narrative has to be constructed as part of the impact assessment. In case of a weak link to crash causation, any predicted change in the target crash population if the evaluated ADAS were to be widely introduced, should only be seen as a general indication of a positive influence of ADAS on the wider target population. This is true at both the national and EU levels.

In light of these methodological concerns, it is important to avoid misuse of this study. For example, it would be incorrect and scientifically unsupportable to use this study to assert that any individual accident would have been affected by the presence of any of the technologies on any particular vehicle. Similarly, this study does not support any claim that a vehicle not equipped with any of these technologies is unsafe or defective in any way.

4.3 Traffic efficiency and environment

The approach for efficiency impacts and environmental impacts were very similar. Therefore they are described together. Besides being able to analyse the function under research by testing hypotheses and answering research questions, the traffic efficiency and environmental impact assessment needed to provide input for the cost-benefit analysis (CBA). The CBA in euroFOT required information about the costs and benefits of the functions at the EU level. The traffic efficiency benefits (some possibly negative) were derived from the traffic efficiency impact assessment. The following quantified traffic efficiency impacts had to be provided to the CBA (at the EU level):

- **Travel time changes** (direct effect)
- **Changes in the amount of accident related congestion**, based on changes in number of accidents (indirect effects)
- **Homogenisation / reduction of congestion effects** for environmental impact assessment (direct effects)

The following quantified environmental impacts were provided to the CBA (at the EU level):

- **Direct effects**: change in **fuel consumption** and **CO_2 emissions** caused directly by a change in tactical driver behaviour (e.g. speed, acceleration)
- **Indirect effects**: change in **fuel consumption** and **CO_2 emissions** caused by a change in kilometres driven (for example less congestion due to less accidents)

Inputs for the traffic efficiency and environmental impact assessment were the performance indicators processed from the raw FOT data during the data analysis.

This section describes how to go from calculated performance indicators to tested hypotheses, answered research questions and quantified traffic efficiency impacts. This includes translating small scale results to the EU level.

Figure 24 gives a high level overview of the steps before, in, and after the traffic efficiency and environmental impact assessment. A certain function is tested in a FOT and produces results in terms of performance indicators and situational variables. Then the traffic efficiency impact assessment starts: direct effects and indirect effects are calculated. Direct effects can be calculated by both a linear and by a modelling approach. Indirect effects are calculated via the safety impact assessment results: number of avoided fatalities and injuries. Direct and indirect effects together form the total traffic efficiency or environmental impact. Results of the traffic efficiency and the environmental impact assessment provide input for the cost-benefit analysis. At several points in the overview figure, external data may serve as additional input. An example is the mileage distribution over road types on EU level. This is used to scale the FOT level effects to EU level. The external data sources are not displayed in the Figure 24.

The three red boxes in Figure 24 are explained in the sections below.

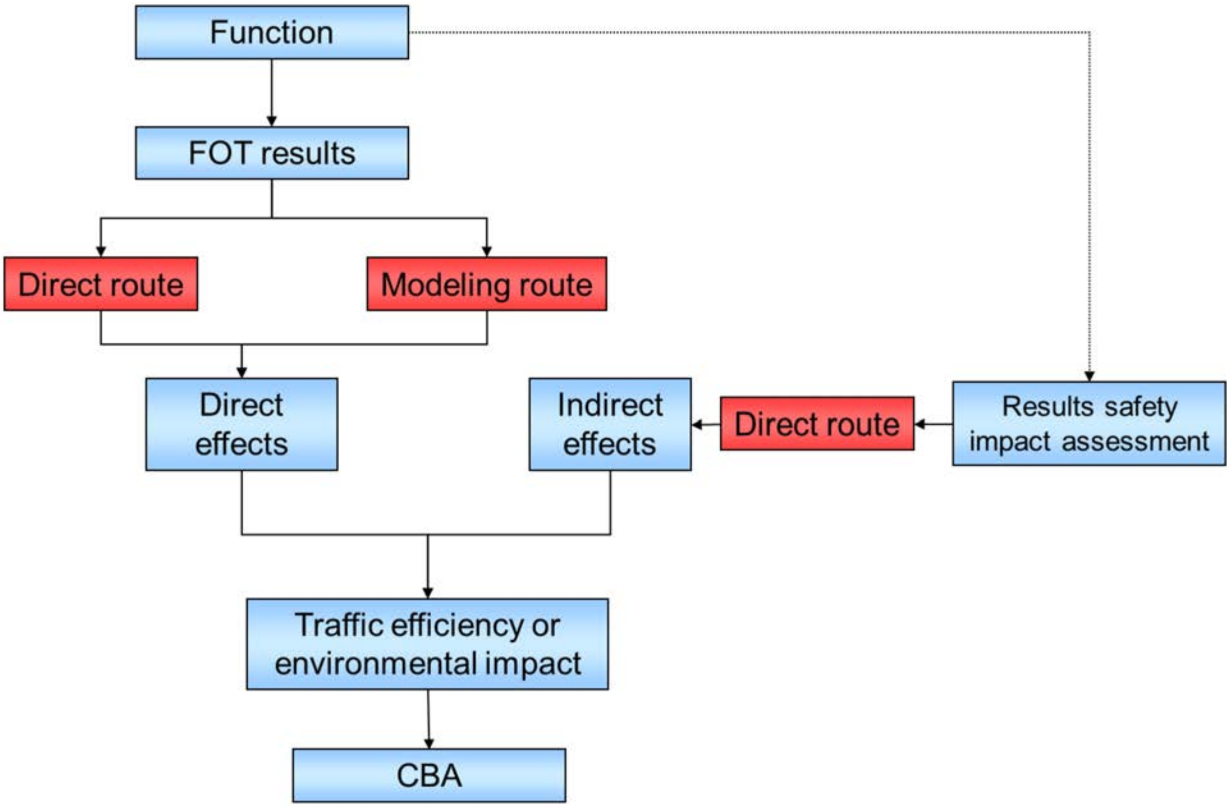

Figure 24: Overview traffic efficiency and environmental impact assessment

Overview traffic efficiency impact assessment per function

Table 6 presents how the different functions need to be assessed. For each function it shows what types of effects (direct and/or indirect) can be expected and need to be assessed. It also shows which assessment approach (linear or modelling) is the most applicable one. Check marks indicate the impacts that are relevant and therefore need to be assessed. Check marks between brackets indicate that this impact may need to be assessed if the FOT data shows an effect of the function on regular driving behaviour. In the FOT it turned out that these impacts were not assessed.

Table 6: Functions and types of assessment

Function	Indirect effects	Direct effects on efficiency	
		Linear approach	Modelling approach
ACC/FCW	✓	✓	✓
CSW[3]	Not enough data		
SL/CC	✓	✓	✓
LDW/IW	✓	(✓)	Not possible, because only change in lateral behaviour expected, which cannot be modelled
BLIS	✓	(✓)	-
SafeHMI	✓	✓	Not possible because of limited data loggings[4]

[3] No effects are expected, taking into account the limited number of instrumented vehicles in the set-up of the pilot with CSW

[4] There are no loggings of the suggested route and the compliance of the driver

FEA	✓	✓[5]	Not possible because of limited data loggings

In SP2, research questions and hypotheses were derived, for each aspect of the analysis (e.g. safety, traffic efficiency, driver-related aspects, and environment). These research questions and hypotheses were revised and prioritised in SP6. The research questions and hypotheses that are answered and tested in this deliverable are the following:

Research questions:

1. What is the impact of function X on travel time?
2. What is the impact of function X on journey speed?
3. What is the impact of function X on amount of delay?
4. What is the impact of function X on variation in speed?
5. What is the impact of function X on network performance per road category?
6. What is the impact of function X on fuel consumption per kilometre?
7. What is the impact of function X on CO_2 and regulated emissions per kilometre?

Different research questions are answered using different types of results. The results that were used were from hypothesis testing, subjective data from questionnaires, safety impacts, simulation results and some additional analyses.

Also, not all research questions were applicable to all functions, and some research questions could not be answered because the results were not significant. This does not mean that the function cannot have an effect on the indicator, but it means that we could not find a significant effect. Table 7 shows which research questions were answered.

Table 7: Research questions answered

No	Research question	ACC and FCW	SRS	Safe HMI	LDW and IW, BLIS, CSW	FEA
1	Impact on travel time	✓	✓	✓		
2	Average journey speed	✓	✓	✓		
3	Delay	✓	✓	✓	✓	
4	Variation in speed	✓	✓	✓		
5	Network performance	✓	✓	✓		
6	Fuel consumption	✓	✓	✓		✓
7	Emissions	✓	✓			

Travel time

The impact on travel time is a combination of the effect on average journey speed and the effect on trip distance. Only for navigation system an effect on trip length and road type was expected. For other functions a seasonal impact was expected to disturb the results on trip length, so subjective data on the impact on trip length and road type was collected to test is. If people did not indicate in the questionnaires that their mobility behaviour changed, then no

[5] Due to limited loggings only environmental impact assessment possible

effect on trip length was assumed. Incidental delay is not taken into account and is reported as a separate indicator. The effect on trip distance is determined.

Average journey speed

The average speed was directly tested from the FOT data using statistical tests. The effect is the difference between the average speed in the baseline period and the average speed in the treatment period. The treatment period includes driving with the system on and activated and also driving with the system off. This means that the usage as observed in the FOT is taken into account. The average speed was determined per road type or speed limit in chunks of one minute taking into account the variance between drivers. Generally ANOVA tests were used. The tests are described in the Annex of deliverable D6.5.

Delay

Delays were calculated with use of the speed distribution. As a starting point it was assumed that on motorways trucks driving slower than 80 km/h are delayed, on rural roads the boundary is also 80 km/h and on urban roads the boundary is 50 km/h. Vehicles that drive 5 km/h cause more delay than vehicles that drive 40 km/h. As a measure for delay vehicle loss hours were used. The indirect effect of a reduction in accidents and the incidental delay as a consequence of that was determined from the reduction in accident. These were only available for the ACC and FCW bundle.

Variation in speed

The variation in speed was analysed using the speed distribution. For a number of functions the expected effect was that people would drive closer to the speed limit. This was analysed in a qualitative way, except for the Navigation systems function which was statistically tested.

Network performance

The network performance was determined in terms of average network speed. The network performance was determined by traffic simulation for the functions ACC and SRS. For these functions the interaction between vehicles was expected to be different for different penetrations of equipped vehicles.

Fuel consumption

The fuel consumption was directly measured from the CAN data in the FOT and tested for significant effect using statistical tests. The effect is the difference between the average fuel consumption per kilometre in the baseline period and the average fuel consumption per kilometre in the treatment period when the function is active. It is then scaled for the usage based on mileage. This means that the usage as observed in the FOT is taken into account. The fuel consumption was determined per road type or speed, limit taking into account the variance between drivers. Generally ANOVA tests were used. The tests are described in the Annex of deliverable D6.5.

CO_2 and regulated emissions

Additional to the effect on fuel consumptions, the function for which the highest environmental benefits were expected, being ACC and SRS, an emission model was used to determine the CO_2 and the regulated emissions CO, NO_x, PM10 and HC. The model used speed-time profiles observed in the FOT to determine the emissions.

4.3.1 Direct route

Direct traffic efficiency and environmental effects (changes in travel times and fuel consumption) can be calculated by the linear approach: FOT data was used directly. The indicators measured in the FOT were tested for significant difference in baseline and treatment in different situations. Then the effects were scaled up to EU level using EU level data about how often these situations occur. This linear approach was applied to all functions, except for the lateral warning function LDW, IW, and BLIS.

4.3.2 Indirect route

Indirect traffic effects are the changes in amount of accident related congestion, based on changes in number of accidents. If accidents are prevented, this means a certain amount of accident related congestion is also avoided. This approach can be applied to all functions as long as safety impacts in terms of a reduction in accidents are available. Two other inputs are used for the indirect traffic effects (besides the safety impacts): (1) the distribution of accidents over the periods of the day (per accident type), which is obtained from GIDAS, and (2) an estimate of the delay per accident type and period of the day according to eIMPACT. The distribution of accidents over the day was applied to the euroFOT reduction of accidents as determined in the safety impact assessment. Assumption was that the impact is the same for all periods of the day. Using eIMPACT, the delay per accident was determined, based on assumptions about traffic conditions and capacity reduction depending on accident severity. This effect is scaled per road type.

The safety impact assessment produced outputs in terms of changes in accidents at the EU level for one function, being ACC/FCW (limitation see 5.1.2).

4.3.3 Modelling route

Three simulation tools were used in the traffic efficiency and environmental impact assessments, two microscopic traffic simulators and an emission modelling tool.

The microscopic traffic simulators are PELOPS and ITS modeller were used to test if vehicles influence each other when a larger share of the vehicles in the traffic flow is equipped with the euroFOT systems. These interaction effects were expected and tested with ACC/FCW and SRS. In PELOPS the interaction effects of ACC and FCW were tested. In the ITS Modeller the interaction effects of SRS were tested. Both traffic simulators implemented the driver behaviour and system usage as observed in the FOT.

Versit+ is a statistical tool that models “regulated emissions” (CO, NO_x, PM10, HC) as well as CO_2, based on a database of driving patterns and associated measured emissions for 3200 light duty vehicles (20.000 tests on 200 driving cycles) and 500 heavy duty vehicles. The Versit+ model can predict real world emissions per second, based on the driving behaviour, which is characterised by the speed and acceleration as a function of time and the specific characteristics of the vehicle(s) in question. The driver behaviour observed in the FOT was used to model the emissions. The characteristics of the vehicle are summarised in a Versit+ vehicle class which is based, amongst others, on the vehicle type (i.e. passenger vehicle, delivery van, etc.), fuel type (i.e. gasoline or diesel) and the Euro class determined by the date of admission. The Versit+ type of the vehicles used in this analysis is a light-weight Euro-5 passenger vehicle with a diesel engine. This designation best matches the properties of the vehicles used in the euroFOT.

4.4 Cost-Benefit-Analysis

Cost-benefit analysis (CBA) represents the most prominent economic assessment tool to prove the profitability of a measure on societal level. Building on the sound methodological guidance offered by the FESTA Handbook [4] and previous impact assessment studies such as eIMPACT [34], [35] and TRL AEBS [36] study the cost-benefit analysis for functions tested in euroFOT assesses the socio-economic impacts in terms of improved road safety, more efficient and more environmental friendly traffic. The monetised impacts (i.e. benefits) are compared to the costs of equipping vehicles (cars and Heavy Goods Vehicles). When benefits exceed the costs, using the tested function is profitable from the overall society point of view.

The CBA goal is hence to inform about the socio-economic dimension of the impacts derived from euroFOT and the costs associated with these technologies. This required not only information available from testing in the field, but also complementing information on safety and traffic performance in the EU-27 in order to provide the bigger picture of European scale effects. Figure 25 illustrates this difference in dimension. The overall cost-benefit study [6] made use of results stemming from other work packages and the related deliverables [2-5]. This information was mostly provided on micro level representing vehicles or vehicle test fleet data. D6.7 [6] on the contrary looked at potential socio-economic effects on European level. Obviously, this required some element of up-scaling from FOT results to EU level. This required up-scaling had to be performed very carefully in order to ensure the credibility of the results. What had to be avoided was to impair the measured results on micro level (FOT) due to necessary simplifications of high level modelling (EU-27).

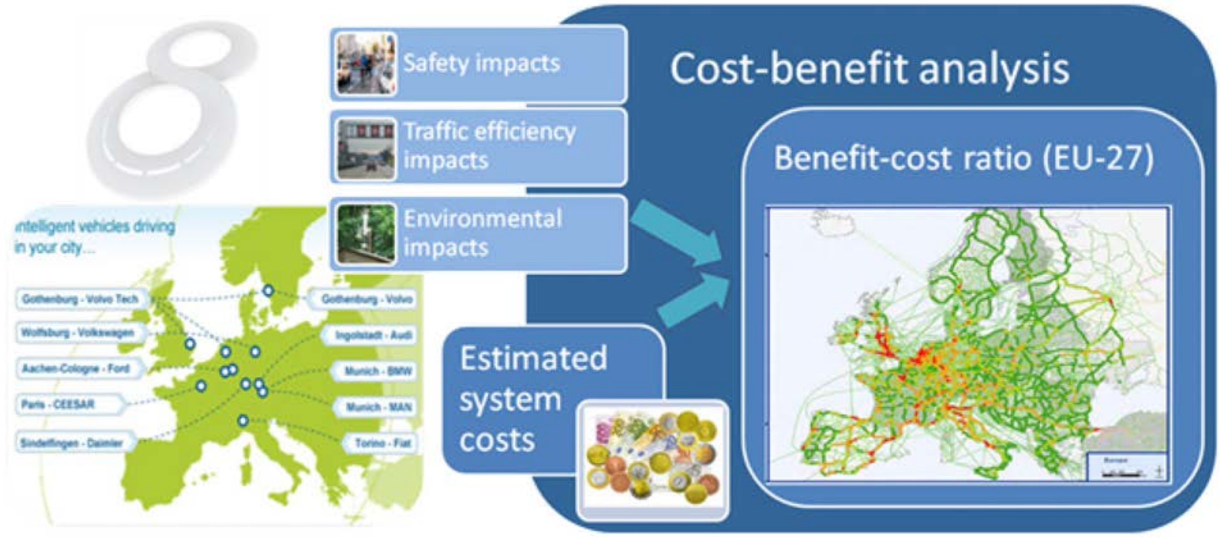

Figure 25: Cost-benefit assessment design within euroFOT

Besides the micro-macro-level (FOT vs. EU-27 impacts) consideration there were also other performance restrictions which limit the applicability of cost-benefit analysis to the euroFOT (impact) results. These restrictions are briefly commented below:

- Individual functions vs. bundles: The socio-economic benefit analysis of driver assistance systems in other studies (as in e.g. eIMPACT [34,35]) was intended to analyse functions that can be seen as independently equipped to vehicles ("optional features") with additional functionality benefits. Since during the trial no functional de-bundling was carried out, ACC+FCW and LDW+IW have to be treated as bundles in the CBA, i.e. looking at them as one comprehensive system with shared benefits and costs.

- Insufficient knowledge on EU-wide driver behaviour and network characteristics: For navigation systems (Safe HMI) and simple control functions (SRS) which impacts depend on the selection of driving periods and routes for which the system is used, determining the baseline – in terms of comparable mileage – is more complex. Due to the limited availability of results, cost-benefit assessment based only on direct fuel or time savings was not applicable for SRS and SafeHMI, since up-scaling these results would require excessive knowledge on EU-wide driver behaviour and network characteristics.
- Partial realisation of impacts due to past system deployment: In addition to monetising the impacts, re-modelling of impacts of already deployed systems such as navigation systems based on the results would be necessary. Identification of the baseline scenario without any deployment is hypothetical and hence represents a huge modelling challenge.

Hence, the most applicable scope for an FOT-based assessment covers driver assistance systems which consist of additional components, offer additional functionality that is a technological add-on to what the driver is capable of, and are marketed as optional features. Thereby, the CBA models on a large scale the meaning of FOT systems and identified impacts for society. The difference between baseline and treatment on FOT level matches for these systems with the socioeconomic state without and with the system.

The initial assessment framework (EU-27 up-scaling, market model, define links to impact assessment) according to FESTA [4] and eIMPACT [34], [35] was set up for functional bundles ACC+FCW and LDW+IW, and for the functions CSW and BLIS, since these features match the scope. This means for those systems the CBA was potentially applicable.

The feasibility of cost-benefit analysis was narrowed down due to non-applicable and / or significant impacts as well as performance restrictions in up-scaling to EU-27 level. A more detailed explanation of the restrictions per system can be found in the subsections of chapter 5.

Based on the above mentioned quality criteria and the limitations of the measured impacts, only ACC+FCW results for both cars and trucks could be taken into account to determine the socio-economic impacts of these systems on European level in a cost-benefit analysis.

4.5 Lessons learned

During the preparation and execution of SP6 ideas to improve an FOT have been collected and are summarized under relevant headers. Some topics in these 32 suggestions flow over into other fields (e.g. set up, piloting) and are one more indication for the interconnectedness of a Field Test.

4.5.1 Methodology

1. There is a need for further development of a methodology suitable for functions used on a voluntary basis. For such functions, it is not straightforward to use classical approaches like speed-accidents relationships, which merge all driving conditions in a single formula.
2. Data sharing among different test sites should be discussed at the beginning to have reliable information what can be shared and what are confidential information.
3. To be able to upscale the results from FOT data to a higher level, it is necessary to assume that the changes detected between conditions in the FOT are caused by the function. This requires that the function and its impact on driving are already widely studied in the literature based on experimental approaches. For other, less investigated

functions up-scaling is extremely difficult because the mechanisms behind the detected changes are not understood well enough.

4. When preparing questionnaires to be used in different countries, translations and related validation should be considered from the very first draft.
5. Piloting should provide feedback concerning the whole experimental design and the practical issues. The FOT could be deeply revised after the pilot results.
6. In order to reach more drivers and improve their response rate it is very good to duplicate or offer multiple options for filling questionnaires in (i.e. hard copy and on-line). It is very important to develop and improve a web-based survey tool for data collection in order to save time and prevent data transcription mistakes.

4.5.2 Data collection

7. For a function used on a voluntary basis, the data collection needs to be long enough to capture rare situations (using the function under a specific but rare driving condition). For example, with 500 000 km collected at the French FOT, some of the more complex hypotheses were answered using less than 1% of the data. For functions used on a driver's demand there is a higher risk that some factor combinations will not be present within the data. Having enough exposure and preferably a similar amount in baseline and treatment ensures a balanced analysis.

   An additional reason for a long enough data collection is to find comparable driving situations in baseline and treatment. When finding comparable situations various situational variables have to be considered. The consideration of additional situational variables shrinks the amount of available data.
8. The length of the FOT should ideally be able to prevent influences from seasonal effects. For example: A decrease in speed between baseline driving that was mainly done in summer and treatment driving that was mainly done in winter might be highly depending on the seasonal weather conditions.
9. Selecting drivers that drive very often the same trips (e.g. to work and back) might increase the number of highly comparable trips, giving precise and reliable results when comparing baseline and treatment. This is an advantage, even if the data could provide a reduced number of different situational variables. In general, data that are well suited for a comparison between baseline and treatment could be identified by using GPS based algorithms.
10. Without the use of video systems a reliable detection of the driver is complicated and might lead to effects that result from different vehicle users (e.g. family members).
11. The definition of road types should be detailed enough to consider differences within rural and urban roads. Especially within urban roads there might be big differences in driving conditions when considering in the same category small one-lane roads and bigger two-lane urban bypasses (where driving is more comparable to motorways). To overcome this problem detailed and reliable map information is needed.
12. Reliable traffic volume estimation is not yet possible using single front radar information and future FOT's should develop methods to better estimate the level of congestion.
13. The intrinsic correlation between different trips of a same driver needs to be taken into account in the analyses, especially if the number of participants is not large enough.
14. It should be guaranteed that no function is available within the baseline phase.
15. To extend information on situational variables information from external database sources can/should be considered (e.g. weather or traffic state information).

16. It is very important to improve and share driver liaison centre procedures in order to provide the best support possible to project participants. It is very important to get in touch with project participants, in order to ensure data collection, but this has to be done case by case, since some drivers could be bothered because of that. Any request has to be timely answered.

4.5.3 Experimental set up

17. The familiarity of the driver with the system should be considered when selecting the drivers.
18. Baseline and treatment periods should be equally filtered in order to focus the analysis to those conditions in which the function being studied was used the most (typical usage scenario in the dataset). Usage within these conditions must be taken into account before making the decision whether or not to consider function activation (e.g. where ACC is actively controlling the longitudinal vehicle movement) as a filtering criterion. This data selection step should be done as soon as possible prior to the start of any analysis or data annotation.
19. Pilot phase should be long enough to test the experimental setup. Especially the tool chain and the monitoring for data collection transfer and storage should be tested extensively to avoid problem during the FOT that cause missing data.
20. Considering different functions as bundles rather than analysing the individual effects helps to interpret the results in a clear way. As a rule, if there are two functions addressing the same crash type, they should be bundled.
21. It is very important to highlight the research aims of the whole project while recruiting drivers. The piloting phase is crucial to improve these procedures and highlight the most effective contact protocol, therein including wording.
22. Pre-screening of customers impacts the final recruitment and participation rate. Recent customers of new vehicles seem to be more sensitive to be involved in transport safety initiatives than experienced owners.
23. Withdrawals mainly appear before filling in the first questionnaire. This questionnaire is also the heaviest one in terms of number of questions. A more balanced distribution of questions, with a lighter first questionnaire, could help to reduce this number of withdrawals.
24. Continuous tasks, for example to register events during driving, are very difficult to be well performed, compared with the discrete task to fill questionnaires in.
25. A link to the OEMs customer services has to be established from the beginning, since customers could use the survey also to send questions that are not related to the project itself, but to the vehicle.
26. Different experimental setups (only CAN data collection, CAN+Video data collection, only questionnaires) prevent a good comparability among different test sites.

4.5.4 Data analysis

27. Enough time should be reserved for video annotation of incidents. Annotating events ranked by a certain severity measure (e.g., deceleration) from top to bottom helped us saving time while keeping those events that were more likely to be “true” incidents.
28. Using automatic incident detection based on kinematic pattern (as applied in German 1 VMC) is able to save time for event recognition but needs sufficient time for algorithms validation in a pre-test to avoid missing events or high false detection rates.

29. Incident detection is a crucial task that needs further research and validated algorithms to be used as an efficient safety indicator.
30. Analysis using both database and file system was beneficial, since different approaches offer different advantages. The database turned out to be extremely good for quick checks on overall trends and averages. The file system was better for more complex calculations (e.g., complex filtering), especially those made on-the-fly.
31. Traffic state estimation is a crucial task when studying longitudinal but not automated systems because drivers tends to use such systems when they feel confident about safety.
32. Harmonisation between different test sites widens the amount of situational variables (resp. data for these situational variables) that can be considered in the analysis. It gives additionally indication whether the results from one test site is reliable. The level of harmonisation should be as low as possible meaning ideally on a collected data level so that the data is put together and one analysis is done afterwards with the complete data set.

4.5.5 Cost-Benefit analysis

33. This study carried out – for the very first time – a cost-benefit analysis which is not based on ex-ante expert assessment of impacts but on results proven in the field. The FESTA methodology has proven its applicability to this type of research question.
34. When performance restrictions are present, socio-economic assessment as final assessment step of FESTA-V must lead to limited results, since only the most trustable and verifiable results can be used in quantitative terms for CBA. But for other functions, it could be possible to make further use of the FOT data, e .g. to test assumptions from ex-ante assessments or to improve simulation models. They could transfer intermediate results into benefit estimations which would reflect the real world impact on a larger scale. If this is not considered, the benefits and BCR results suffer from a “pessimism bias”. This must be considered in early phases of future projects e.g. by providing a contingency plan to make use of simulation or further expert assessments.
35. Upscaling from micro level (FOT) to macro level (EU-27 databases for accidents etc.) provides still considerable challenges, especially concerning the granularity of information. CBA makes typically use of averages of variables whereas distributions of variables would be valuable to keep the value added of FOT data. Research in this direction would help to solidify the derivation of socio-economic impacts from Field Test data.
36. Socio-economic impact assessment should allow for a wider scope of impacts, including those beyond transport, i.e. for the overall economy. Such impacts for productivity, growth and employment represent important results for policy making. There are concepts available to broaden the scope of CBA and to include macroeconomic / wider economic impacts in a “twin approach”. These figures have a different quality or nature than measured effects within a Field Operational Test.

5 Final evaluation results

In this chapter the final evaluation results of the tested functions are presented. Thereby for each function respectively bundle the final results of the different analysis steps are presented separately. At the end of the subchapters a final conclusion is provided for each function. The analysis is conducted by means of the elaborated methodology for the different data analysis step (e.g. safety, traffic efficiency).

Depending on the available data and the finding of the data analysis the following analysis steps were conducted for the tested functions (see Table 8).

Table 8: Overview of analysis steps conducted for the tested functions

	Safety	Traffic efficiency	Environment	User acceptance	Up-scaling to EU-27 (Safety)	Cost-benefit analysis
ACC + FCW	✓	✓	✓	✓	✓	✓
LDW + IW	✓	n/a	n/a	✓	n/a	n/a
Navigation Systems	✓	✓	✓	✓	n/a	n/a
SRS	✓	✓	✓	✓	n/a	n/a
BLIS	✓	n/a	n/a	✓	n/a	n/a
FEA	n/a	n/a	✓	n/a	n/a	n/a
CSW	n/a	n/a	n/a	✓	n/a	n/a

Not all of analysis steps were applicable for the eight tested functions. Due to limited data or not significant results certain analysis steps were not applicable (n/a) in order to conduct an up-scaling or a Cost-benefit analysis.

Altogether, about 35 million km were driven in euroFOT during the data collection phase. This data was used for the data processing and finally for the data analysis. The following table provides an overview of the collected data in euroFOT.

Table 9: Overview collected data within euroFOT

Operation site	Total Mileage [km]	Mileage used for statistical data analysis [km]	Total hours of driving [h]	Type of Collected data
CEESAR (French) OS	600.000	545.340	14.000	CAN and Video, CAN only
Ford (German OC1) OS	2.030.000	1.490.000	61.844	CAN only
MAN (German OC1) OS	7.500.000 (expected 10 mio)	180.000	182.467	CAN only
VW (German OC1) OS	300.000	130.000	6.315	CAN only

BMW (German OC2) OS	383.392	330.049	6.021	CAN and video
Daimler (German OC2) OS	629.870	610.858	9.108	CAN and video
Fiat (Italian) OS	>8.000.000	8.000.000	194.632	Questionnaires only
VCC (Swedish) OS	1.069.460	1.069.460	26.019	CAN and video
Volvo (Swedish) OS	14.356.000	4.000.000	97.316	CAN and video
Total	**34.868.722**	**16.355.707**	**597.722**	

After the data was collected it was processed for data analysis purposes. Within the data processing the raw data was reduced to the relevant data sets (e.g. only car-following situation for ACC, information on road type available, weather condition known, only drivers with mileage of 100 km in each comparison condition). This led finally to a reduction of the collected data, which in the end only contains the relevant data needed to perform the data analysis. In the following the final evaluation results are presented.

5.1 Adaptive Cruise Control (ACC) and Forward Collision Warning (FCW)

The Adaptive Cruise Control (ACC) was evaluated in a bundle with the Forward Collision Warning (FCW) at two test sites namely German1 and Swedish VMC. Data from three passenger car manufacturers (Ford, VW and Volvo cars) and two truck manufacturers (MAN and Volvo trucks) was gathered during the FOT (see Figure 26) which lasted at least six months for each manufacturer (three months of baseline and three months of treatment). During this time 174 passenger car and 53 truck drivers travelled more than two million kilometres of which more than 1.3 million km were used for the statistical analysis (see Table 10).

Table 10: Overview of used data

	Mileage		**Number of drivers**	
	Baseline	Treatment	$N_{passenger\ cars}$	N_{trucks}
Overall	727.114 km	623.615 km	174	53
Motorway	676.924 km	602.866 km	174	53
Rural	24.983 km	12.228 km	64	-
Urban	25.207 km	8.521 km	64	-

In addition to the collection of objective data subjective impacts of the drivers were assessed by three questionnaires that were filled in by the drivers during the progress of the FOT. Based on the possible distinction between ACC and FCW within the questionnaires the results for user acceptance and user related aspects can be presented separately while for the other subchapters ACC and FCW are treated as a bundle.

Ford	MAN	VOLVO	VCC	VW	CRF	IFSTTAR	BMW	DAG
✓	✓	✓	✓	✓				

Figure 26: Manufacturers that gathered ACC+FCW data during the FOT

With the help of the gathered data it was possible to answer the research questions on user acceptance and driver related aspects, safety, traffic efficiency, environment as well as the cost-benefit ratio which were defined at the beginning of the project. To that end the statistical analysis of the objective and subjective data aimed to test the hypotheses that were derived based on the research questions. In Table 11 and Table 12 the list of hypotheses to be answered is presented.

Table 11: ACC+FCW hypotheses list for objective data

	Hypothesis	Performance indicator
Safety	Using ACC+FCW, the number of forward crashes, near crashes, and incidents will decrease	Number of incidents per 100 km
	ACC+FCW decreases the number of critical time gaps to the leading vehicle	Number of THW<0.5s per 100 km
	ACC+FCW increases average time gap.	Average THW
	Using ACC+FCW, the number of harsh braking/ strong decelerations will decrease.	Number of harsh braking per 100 km
Traffic efficiency	ACC+FCW decreases average speed.	Average speed
Environment	ACC reduces the average fuel consumption	Average fuel consumption
User acceptance and driver related aspects	ACC+FCW use increases over time	Duration travelled with active ACC divided by total travel time Number of ACC activations divided by total travel time
	The driver changes the use of ACC over time by increasing the occurrence of overriding the ACC function by using the accelerator pedal.	Number of overriding the ACC by pushing the accelerator pedal divided by time travelled with active ACC
	Using ACC+FCW, frequency of drowsy driving will increase	
	Using ACC+FCW, driver’s reaction time (time to reach the brake pedal) will increase if ACC is used most of the time and decrease if only the FCW function is actually used	
	Using FCW+ACC, focus on primary task (time in which the driver looks straight ahead) will decrease over time on motorways.	
	Using ACC+FCW focus and level of engagement on secondary tasks will increase.	Subjective rating

Table 12: ACC+FCW hypotheses list for subjective data

	Hypothesis	Performance indicator
Driver behaviour	Using ACC+FCW focus and level of engagement on secondary tasks will increase.	Subjective rating
Acceptance	ACC+FCW increases driving perceived safety and comfort.	Subjective rating
	Certain features of the functions, in terms of usability, influence acceptance.	Subjective rating
	Certain features of the functions, in terms of usefulness, influence user acceptance?	Subjective rating
	Acceptance changes over time with function use.	Subjective rating
Trust	Trust in function changes over time with function use.	Subjective rating
Workload	Driver workload decreases over time with function use.	Subjective rating
User practices	User practices (heuristics, rules) will change over time during the FOT	Subjective rating
Abuse/Misuse	Drivers will not abuse or misuse ACC+FCW	Subjective rating

In the following the main results of the analyses are presented. Instead of answering all hypotheses listed in Table 11 and

Table 12 a more general description of the impacts of the ACC+FCW is given to highlight the system related influences. The results of the hypothesis testing can be found in the Annex of deliverable D6.5. For more detailed analyses the reader is referred to the deliverables D6.3 for user acceptance and driver related aspects, D6.4 for the safety impact assessment, D6.5 for the traffic efficiency and environment impact assessment and D6.7 for the cost-benefit analysis.

5.1.1 User acceptance and user related aspects

Since it is possible to distinguish between ACC and FCW within the questionnaires the results for user acceptance and user related aspects are presented separately while for the other subchapters ACC and FCW are treated as a bundle.

ACC

ACC related changes in the driving behaviour can be deduced from the same performance indicators as in chapter 5.1.2, 5.1.3 and 5.1.4. These are mainly based on objective data and indicate especially safer distance behaviour and a more homogenous speed distribution which lead to safety and environment benefits. The increase in average speed (between 2% and 4.5%) however gives in addition to the interpretations for the traffic efficiency insights on the conditions in which the ACC+FCW is used. This is that drivers tend to activate the systems in situations that allow higher average speeds and are rated as “safe”.

The questionnaire data indicates that the expectations of drivers to the system were fulfilled, i.e. the scores on satisfaction and usefulness that drivers gave before gaining access to the systems matched those given during and after the trial. In Figure 27 it can be seen that the acceptance rating on the Van der Laan scale (scaling from -2 to +2) shows very low variation. The acceptance is based on the average of the questionnaire items related to satisfaction and usefulness.

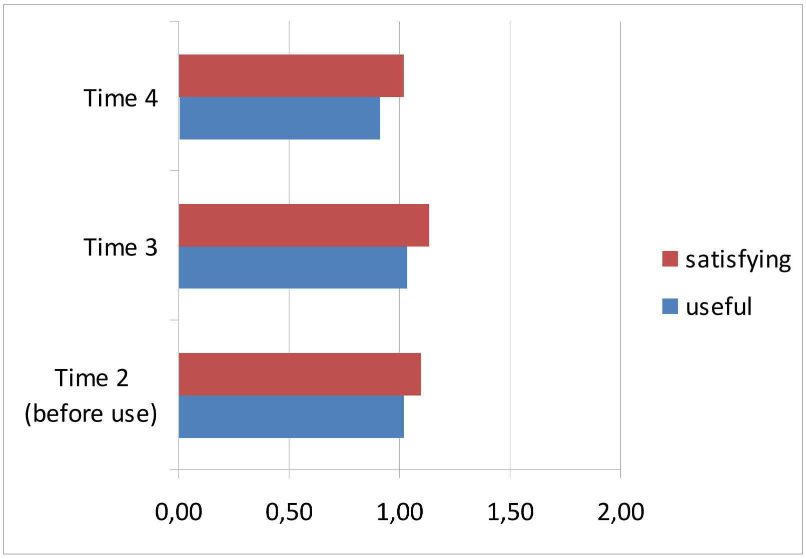

Figure 27: Acceptance rating in terms of usefulness and satisfaction

In terms of usage, it seems clear that drivers make the most use of ACC on motorways, which was quite expected. Also the hypothesised increase in the usage frequency could be confirmed by evaluating the travel time and distance with and without active ACC. Comparing the first months to the last months of the FOT there was a significant increase of ACC use in terms of travel time with active ACC (31%) and frequency of ACC activations (53%). The drivers seem to get used to the positive perception of the ACC system and use the system longer and more often over time even though they did not indicate a change in their usage behaviour within the questionnaires.

This increased use of ACC is in line with the perceived increase of safety and comfort which has been self-reported by the drivers. In contrast, self-reported ratings on trust did not change over time and thus did not reflect the positive perception related to safety and comfort. Confidence which is a sub-criterion of trust even decreased thus expressing that the drivers had higher expectations than the system could fulfill.

FCW

The results of the questionnaire answers specifically related to FCW can be summarised according to the following:

- Close to 70% of drivers feel that FCW increases safety.
- Before trying FCW, participants had very high expectations of the system. These were later somewhat devaluated based on their actual experience of the system (mainly for items effective, raises confidence and trustworthy).
- Despite this, the perceived usefulness and driver satisfaction are both very high and also stable, i.e. they do not increase or decrease over time.
- FCW is perceived as most useful on motorways in normal traffic.

An interesting finding is that confidence in FCW did decrease significantly when drivers started to actually use the system, as compared to before they had access to it. This means that drivers had higher expectations regarding the way in which FCW should “raise confidence” than the system could fulfil and, once confronted with the system’s limitations, these expectations had to be revised downwards. Despite this, most drivers perceive that the system increases safety and therefore satisfaction and usefulness remain high throughout the study. Another way of interpreting this is that driver expectations before interacting with the system were unrealistically high. This highlights the importance of managing driver expectations when these systems are introduced, in order to avoid levels of disappointment that might decrease system overall usage.

Drivers were not uniformly positive to the evaluated FCW's audio-visual interface, i.e. some reported that they perceived the timing of the warnings as too early and therefore annoying. This was expected, as many researchers have argued that drivers have individual comfort zones, and a following distance that is perceived as being too close for one driver may seem as a perfect distance to another. However, it reinforces the need for investigating new and creative ways of adapting warning timing to driver acceptance thresholds. A satisfied driver is more likely to respond as desired to a warning than an unsatisfied one.

5.1.2 Safety

It can be said that the combination of ACC and FCW is able to give a safety benefit for passenger cars as well as for trucks based on the data gathered in the FOT. This positive effect can be attributed to changes in the distance behaviour while driving with active ACC and FCW.

Passenger cars

The average time headway (THW) showed an increase of about 16% (see Figure 28) and leads therefore to bigger safety margins. Due to the predefined settings of the ACC time-headway the number of (intended or unintended) close approaching manoeuvres is highly reduced and prevents therefore critical driving situations. The analysis of critical time headways (< 0.5 s) revealed a reduction between 63% in urban areas and 81% on rural roads (73% on motorways). As a consequence of the safer distance behaviour the frequency of harsh braking manoeuvres is lower when driving with active ACC. On motorways two out of three extreme braking events (67%) happening in the baseline can be avoided by the use of ACC (in the treatment). However, the reduction on rural and urban roads is somewhat lower (45% and 32%).

Like for the harsh braking events the number of incidents is lower when using ACC+FCW. The incidents based on vehicle kinematics show more than 80% reduction while the video-based analysis evaluates more than 30% less incidents when driving on motorways. The decrease in video-based incidents which include individual rating based on subjective assessment was however not statistically significant. Details on the different ways of detecting incidents can be found in [37] and Annex 12 of D6.4.

Explanations for the increase in average time-headway and the reduction of critical time-headways, harsh braking events and incidents can be found in the selectable ACC settings that can never be lower than the legally prescribed value which is not always considered by drivers in baseline driving. Resulting from the increase in average time-headway the reaction time to avoid close approaching events is higher. If the driving situation exceeds the braking capacities of the ACC because of a highly decelerating vehicle in front the presented warnings (by the ACC and the FCW) give the driver appropriate time to react on the driving situation. It could be shown in the analysis that this effect can be mainly attributed to the ACC by comparing situations where only one of the functions was active.

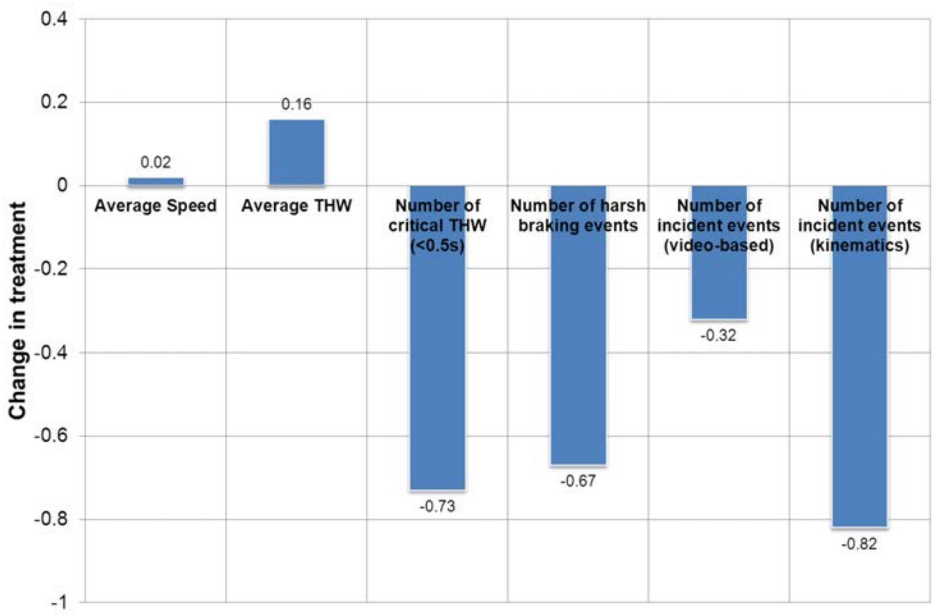

Figure 28: Overview safety indicators for motorway (passenger cars)

The found increase in average speed can be interpreted in different ways: On the one hand an increase in average speed was previously linked to decrease in safety (e.g. by [38]) but is on the other hand also an indicator that driver tend to use the system in situations that allow higher speeds and are therefore rated as safer (by the driver). It is therefore hard to decide where this increase indicates a decrease in safety or is only caused by the drivers' choice when to use the system.

As an additional safety indicator the engagement in secondary tasks was evaluated. Interestingly, it increased but only for non-critical driving episodes, not during actual critical events. This indicates that drivers do make use of the "freedom" to think and move that ACC provides when engaged, but do so in a selective and safe manner.

Trucks

Based on the overall results for ACC+FCW in trucks, displayed in Figure 29, it can be concluded that safety does improve when drivers use ACC+FCW. Although no decrease in average speed was observed (an indicator previously linked to increase in safety by [38]), the extended time headway and the reduced number of critical time-gaps significantly contribute to creating larger safety margins. Average time headway (THW) showed an overall increase of about 5% and the frequency of critical THW's reduced 54% on motorways.

As a consequence of the safer distance behaviour, the frequency of harsh braking manoeuvres is lower when driving with active ACC+FCW. In treatment, 37% of the events when drivers slammed on the brakes were avoided.

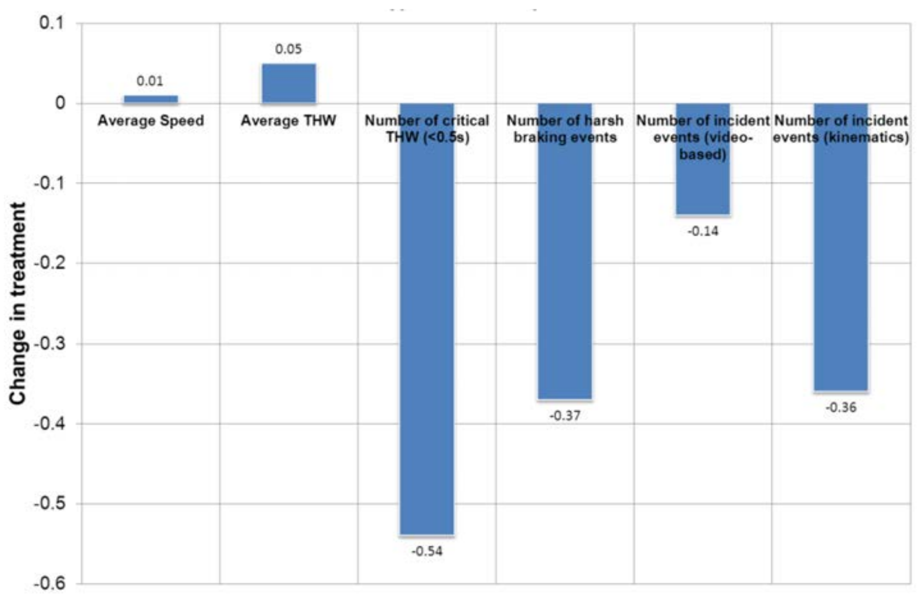

Figure 29: Overall benefit of ACC+FCW on motorways (trucks)

In addition, a reduction in both kinematically derived (-36%) and video annotated incidents (-14%) was observed, although the latter was not statistically significant. While the incidents based on the video annotation include the individual rating based on subjective assessment, the kinematic related incidents evaluate the measurements of the vehicle dynamics and compare those to predefined thresholds. Details on the different ways of detecting incidents can be found in [37] and in the Annex 2 of D6.4.

The effect of ACC+FCW on the number of FCW's was also investigated using 2197 warnings in total. During baseline, warnings were logged but not displayed to the driver. Results show a reduction in risk in treatment, which suggests a positive effect of the bundle.

We also investigated hypothesised negative side effects of ACC+FCW in terms of increased secondary task engagement and attention to forward roadway. Given that ACC+FCW was expected to lower drivers' workload, secondary task engagement might increase. However, results showed no such effects. Overall, drivers kept their focus on the road while using ACC+FCW. This was true both during normal driving and during crash relevant events, although the analysis with incidents was based on a comparatively small number of events, and hence is to be viewed more as a trend.

Transfer of safety results into accident reductions

In terms of projecting what the safety indicators changes would mean if ACC+FCW was widely deployed in the EU-27, it was concluded that ACC+FCW in passenger cars might have a positive effect on the overall crash population. In trucks, this conclusion could only be made for motorways. Hence, given the assumption that the safety-related indicators are good indicators for how the accident scenario would change if all vehicles were equipped, ACC+FCW cars could potentially affect up to 2.2-5.7% of the injury accidents on motorways, while ACC+FCW trucks could potentially affect up to 0.2-0.6% of these accidents, see Figure 30.

Further estimations based on the relevant rear-end target crash population accident data can be made for EU-27, e. g. regarding involved injured individuals. Note that these results are based on a set of assumptions. They are therefore to be used with caution, and need to be put into the perspective in the light of all the assumptions made within the analysis framework.

ACC+FCW cars	Motorway		Rural		Urban	
EU-27 target group	low	high	low	high	low	high
Fatally inj. car occ.	1.68%	4.25%	0.09%	0.13%	0.10%	0.10%
Injured car occ.	2.54%	6.42%	0.52%	0.73%	0.36%	0.36%
Fatalities (all)	1.16%	2.95%	0.06%	0.08%	0.02%	0.02%
Injuries (all)	2.24%	5.66%	0.42%	0.59%	0.16%	0.16%
Injury accidents (all)	2.24%	5.68%	0.47%	0.65%	0.14%	0.14%

ACC+FCW trucks	Motorway	
EU-27 target group	low	high
Fatalities (all)	0.33%	0.85%
Injuries (all)	0.18%	0.45%
Injury accidents (all)	0.21%	0.55%

Figure 30: Proportion of the total crash population that ACC+FCW might positively address

5.1.3 Traffic efficiency

The change in average speed was already discussed in section 5.1.1 and 5.1.2 since it gives insights on various aspects of the driving behaviour. Average speed is increased between 2% on motorways and urban roads and 4.5% on rural roads for passenger cars. Considering both phases with active ACC as well as those without ACC activation in the treatment period there is a small reduction (less than 0.4%) in average speed on motorways and urban roads for passenger cars, compared to the baseline. The effect on trucks is minimal (less than 0.1%). The differences between ACC+FCW being active and being not active within the treatment period can be seen in Figure 31 and Figure 32. As it was already discussed in section 5.1.1 the average speeds using ACC+FCW is higher than in driving without the system. The subjective results show no effect of ACC and FCW on mobility behaviour, route choice and choice of road type.

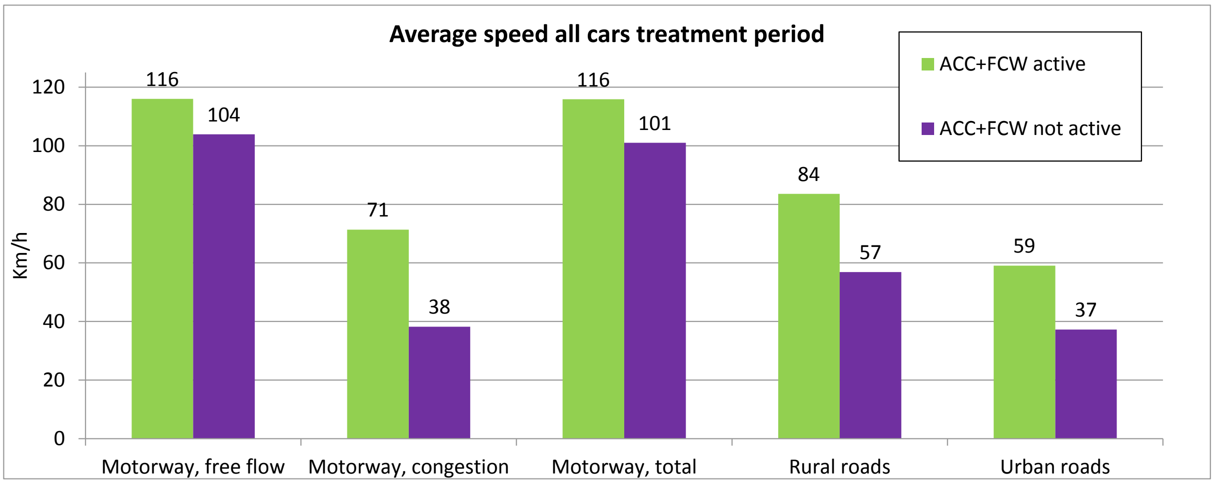

Figure 31: Average speeds per road type within treatment period (passenger cars)

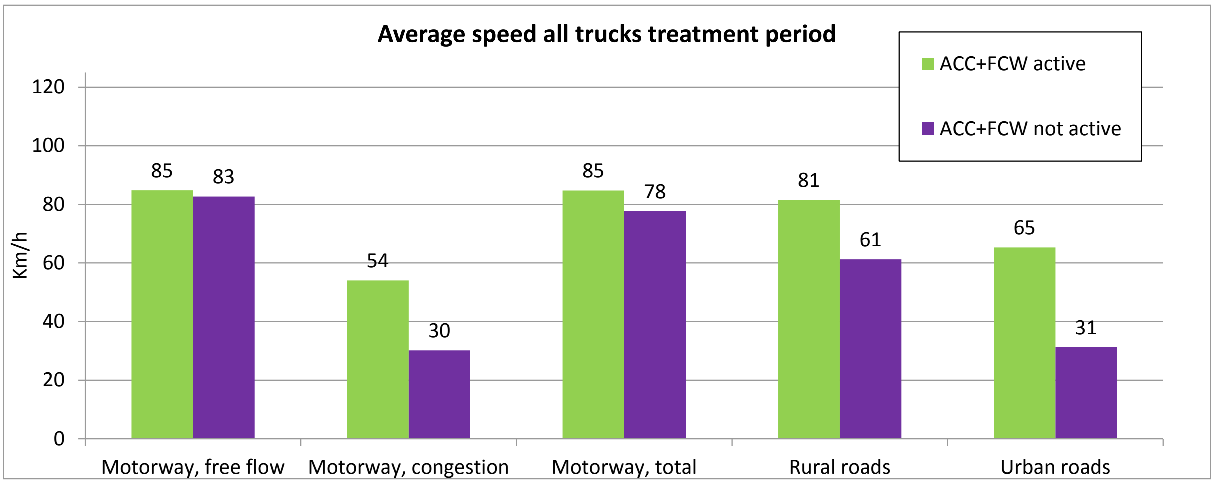

Figure 32: Average speeds per road type within treatment period (trucks)

In addition to the objective and subjective data gathered during the FOT traffic simulations were conducted to be able to vary ACC penetration rates within the analysed vehicle fleet. The simulations show that the effect on network speed is similar in size to the effect found in the FOT and generally linear when more vehicles are equipped. The simulations are based on the driving behaviour and system usage observed in the FOT. The effect scales linear with the penetration of equipped vehicles in most situations. Only in heavy traffic scenarios when less than 25% of the vehicles are equipped, the average network speed is reduced slightly stronger than the speed reduction measured for the individual FOT vehicles.

Beside the function related direct changes of traffic efficiency (e.g. less congestion because of more homogenous traffic flow) an indirect positive contribution to traffic efficiency is based on increased safety when using the ACC+FCW. Therefore, a reduction in terms of accidents (derived in the safety impact assessment, see D6.4) can be transferred into decrease of incidental delay that is measured in lost vehicle hours. Considering both fatal and injury accidents annual savings of more than three million of vehicle loss hours level could be reached on an EU-27. Details on the methodology of calculating the incidental delay can be found in D6.5.

5.1.4 Environment

As assumed in the hypothesis there is a significant reduction in fuel consumption while driving with ACC and FCW for both vehicle types (passenger cars and trucks). For passenger cars a decrease of 2.77% was found while the reduction for trucks is somewhat lower with 1.78%. Notice that the results consider only driving on motorways and are based on the gathered FOT data. In Figure 33 the system related changes in fuel consumption are combined with the usage rates that were found during the FOT and scaled up to the EU-27 level. The overall fuel saving potential for passenger cars is 1.37% and slightly below 1% for trucks. This accounts for 790 million litres of fuel every year and almost 2 million tons of CO_2 based on the average fuel consumption that was evaluated with the objective data. However, the results for trucks seem more reliable since the driving patterns that are compared in baseline and treatment are very similar because of the general traffic and driving situation (car following situations with little speed variation).

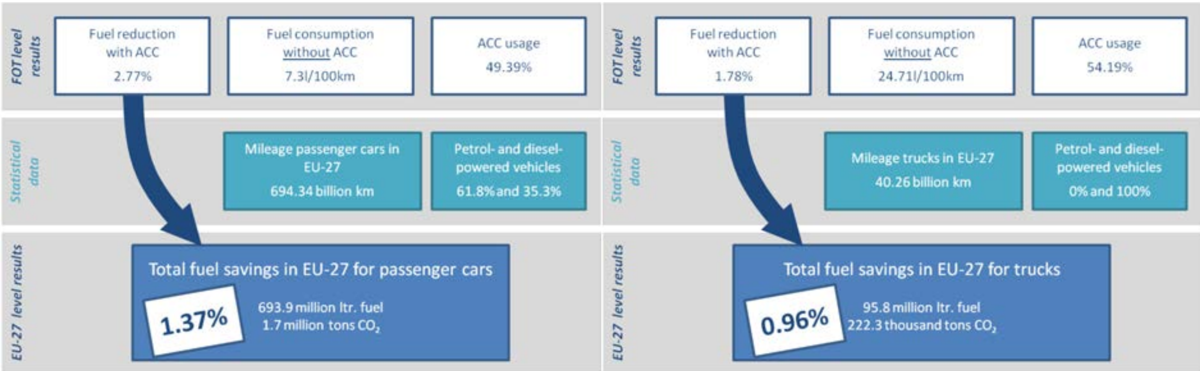

Figure 33: Fuel saving potential for passenger cars and trucks when using ACC

In addition, simulations were used to calculate also effects of the ACC use on regulated emissions (CO, NO_x, PM10, HC) as well as CO_2. As input to these simulations speed profiles gathered in the FOT were used. The data set included almost 100 hours of driving time from nine different drivers. Considering usage rates for the different road types the effects on CO_2, HC and PM show a very small increase of less than 1%. Only CO and NO_x emissions show increases higher than 2% on motorways.

5.1.5 Cost-Benefit Analysis

The result of the cost-benefit analysis is composed of the costs and benefits stemming from road safety improvement, more efficient and more environmental friendly traffic. The main findings can be summarised as follows:

1. The costs of equipping the passenger cars and heavy trucks with the combined system lead to annually approx. 1.6 billion € (passenger cars) and approx. 28 million € for heavy trucks (because of the smaller fleet). When only parts of the fleet will be equipped (e.g. 10% of the car fleet), the costs amount to 240 Mn EUR.
2. Annual benefits for cars add up to 0.8 to 1.2 Bn EUR (full penetration) respectively 126 to 175 Mn Euro (10% penetration rate), depending on the magnitude of safety impact. The result is dominated by the safety impact which accounts for approximately half of the benefits in the lower bound scenario and two thirds in the upper bound scenario. However, also traffic impacts and environmental effects provide substantial contributions to the benefits.
3. Annual benefits for trucks amount to approximately 108 and 146 million €. The same pattern of results as for cars appears also here. Safety is dominant in the upper bound scenario whereas traffic represents the biggest impact in lower bound scenario

4. For trucks, the ACC+FCW bundle is clearly profitable from society point of view. The benefit-cost ratio is between 3.9 and 5.2.
5. For cars, the attainable benefits are not sufficient to outweigh the costs. The benefit-cost ratio ranges between 0.5 and 0.7. The system is either too expensive or users on average drive too less km for pay off of the “investment”. It has to be kept in mind that the tested system ACC+FCW represents foremost a comfort system. These effects are however not subject of monetisation in a transport-focused cost-benefit analysis.
6. Sensitivity of the results was tested for the cars scenario. The overall result was that modifying input parameters (such as higher cost-unit rates for impact appraisal, considering potential underreporting of injury accidents) would bring the benefit-cost ratio close to or even above 1. Changing of the penetration rate and taking different levels of economies of scale into account provides a BCR above 1 for a scenario assuming large economies of scale and a penetration rate of at least 50%.
7. Former ex-ante impact assessment studies have indicated more favourable benefit-cost results (e.g. eIMPACT). The differences for euroFOT can be explained by making use of in-depth databases for modelling the accident target group, considering empirical evidence of usage rates and the estimation of system cost (expert estimations vs. market price based assessment).
8. For passing the profitability threshold it would require to widen the scope of the assessment by including also benefits from avoiding property damages. In this context, a first best estimate study on the basis of Allianz insurance databases with PDO claims (minor, TPL and MoD) using euroFOT results revealed, that in EU-27 each year approximately 500,000 PDO claims could be avoided or at least mitigated if all passenger cars would be equipped with ACC+FCW (generation 2008). This is particularly remarkable as for newer generations of ACC+FCW even higher accident avoidance is probable. Further benefits are expected if wider economic impacts in terms of growth and employment will be considered.

Table 13: Results of the cost-benefit analysis for ACC+FCW (passenger cars)

	lower bound (impacts)	upper bound (impacts)
Safety	459.900.000	805.170.000
Traffic efficiency	286.584.283	300.562.283
Environment	84.000.000	84.000.000
Benefit of ACC+FCW	830.484.283	1.189.732.283
System cost (FESTA rule) [€/new reg. fleet]	1.624.000.000	1.624.000.000
Benefit cost ratio	**0.51**	**0.74**

Table 14: Results of the cost-benefit analysis for ACC+FCW (trucks)

	lower bound (impacts)	upper bound (impacts)
Safety	55.400.000	141.990.000
Traffic efficiency	71.628.443	74.357.443
Environment	15.540.000	15.540.000
Benefit of ACC+FCW	142.568.443	231.887.443
System cost (FESTA rule) [€/new reg. fleet]	83.250.000	83.250.000
Benefit cost ratio	**1.713**	**2.785**

5.1.6 Conclusion

Overall, ACC seems to be a highly appreciated and used function, that both increases driver comfort and safety. Questionnaire data indicates that the expectations of drivers to the system were fulfilled. The positive experiences of the drivers can also be seen in the increased use over the treatment (31% in travel time and 53% in the activation frequency).

With regard to safety aspects the benefit of the function can be attributed to the increased safety margins that are caused by the ACC settings which prevent the driver from (intended and unintended) close approaching manoeuvres. This benefit was measured in terms of reduction of critical time-headway situations, harsh braking manoeuvres and incidents. All these indicators show significant high reductions between 30% and 80%.

Based on the positive influences of the ACC+FCW to the safety there are also positive (indirect) effects on traffic efficiency. Due to the reduction of accidents the annual incidental delay in terms of lost vehicle hours sums up to more than three million hours on an EU-27 level. Direct effects of the use of ACC+FCW however are hard to measure because of the fixed penetration rate within the considered vehicle fleet. The increase in average speed when driving with activated ACC cannot be assigned to function-related changes in the traffic flow but gives instead insights on the usage behaviour.

The environmental impact was measured with the help of the fuel consumption. Here, a reduction of about 3% was found evaluating the data for passenger cars and almost 2% for trucks. This effect can be directly related to changes in driving behaviour with active ACC and does not consider additional effects on fuel consumption that result from changes in traffic efficiency (e.g. the found reduction in incidental delay).

The results of the cost-benefit analysis show the profitability of ACC+FCW for heavy goods vehicles. An important factor for the profitability is the high annual mileage of heavy goods vehicles. The benefit-cost ratios for cars does not pass the profitability threshold of BCR>1 in the base case. Nevertheless, sensitivity tests have demonstrated that under modified assumptions (target population, impact assessment, impact appraisal and level of costs) profitability may be also reached for cars.

5.2 Lane departure Warning (LDW) & Impairment Warning (IW)

In euroFOT, the impact of Lane Departure Warning (LDW) was investigated. LDW was meant to be evaluated at two test sites namely German1, Swedish and Italian VMC. Objective data from the CAN-Bus and subjective data (questionnaires) were gathered at the German1 and Swedish VMC. Furthermore additional subjective data for LDW were gathered at the Italian VMC. (see

Ford	MAN	VOLVO	VCC	VW	CRF	IFSTTAR	BMW	DAG
		✓	✓		✓*			

Figure 34). However, in the end the German1 test site could not deliver data; hence the below results are based on data from the Swedish VMC only.

Ford	MAN	VOLVO	VCC	VW	CRF	IFSTTAR	BMW	DAG
		✓	✓		✓*			

Figure 34: Manufacturers that gathered LDW (+IW) data during the FOT (*only subjective data)

At the Swedish VMC in the passenger car fleet, LDW was analysed as a bundle with another function: the Impairment Warning (IW). Drivers experienced two conditions; 1) driving without the system for 4 months (baseline condition) and 2) driving with the system available for use during approximately 8 months (treatment condition). While subjective data (i.e. questionnaire responses) is available for the full duration, due to delays in the collection of objective data, the objective data analysis is based on 3 months of baseline and 3 months of treatment.

During this time 98 passenger car and over 100 truck drivers travelled more than 4 million kilometres of which more than 550.000 km were used for the statistical analysis. In addition to the collection of objective data subjective impacts of the drivers were assessed by four questionnaires that were filled in by the drivers during the progress of the FOT.

With the help of the gathered data it was possible to answer the research questions on user acceptance and driver related aspects, safety, traffic efficiency and environment defined at the beginning of the project, but not a full safety impact or CBA (reasons why are explained in section 3.2.5). Statistical analysis of the objective and subjective data was carried out to test the hypotheses derived based from the research questions. In Table 15 and Table 16 the list of hypotheses addressed is presented.

Table 15: Hypotheses tested with objective data

	Hypothesis
Safety	LDW+IW decreases the number of crashes, near crashes and incidents
	LDW + IW decreases drowsy driving
	LDW + IW issues warning when the driver is not looking at the road ahead
	Using LDW+IW, focus on primary task (time in which the driver looks straight ahead) is lower in crash relevant events (CRE)
	Using LDW+IW focus and level of engagement on secondary tasks will increase
	Using LDW+IW focus and level of engagement on secondary tasks will increase in crash relevant events (CRE)
	LDW+IW increases the use of turn indicators in lane change situations
	LDW+IW increases night driving

Table 16: Hypotheses tested with questionnaire data

	Hypothesis
Acceptance	LDW and IW* is well accepted by the driver
	Certain features of the LDW and IW functions, in terms of usefulness, influence acceptance
	Acceptance changes over time with function use
	Certain features of the functions, in terms of usability, influence acceptance.
Trust	Trust in function changes over time with system use
Workload	Driver workload decreases over time with function use.
User practices	User practices (heuristics, rules) will change over time during the FOT
Abuse/Misuse	Drivers will not abuse or misuse LDW and IW
LDW+IW	LDW+IW influence lateral driving performance.

* Note that drivers responded separately to LDW and IW questions, and hence questionnaire based answers are available on a per function level in D6.3. Since the tested hypotheses are the same however, they are grouped together in this table.

5.2.1 User acceptance and user related aspects

For the analysis of the user acceptance and user related aspects (based on subjective data) the evaluation for LDW and IW is conducted separately, because specific function related questions on LDW and IW have been asked in the questionnaires. In the following the results for LDW and IW are presented.

LDW

The user acceptance and user related aspects analysis was done based on two approaches with two different experimental setups on two different test sites. While in the Italian VMC only questionnaires in a higher frequency were used to evaluate user related aspects, the Swedish VMC additionally considered objective data from the CAN bus and video cameras. Because of the differences in the experimental setup the results for these two fleets are therefore presented separately.

5.2.1.1 Objective and subjective data

The results of the questionnaire responses regarding LDW can be summarised according to the following:

- Participants find the LDW system useful but the satisfaction with the system is low.
- There is no difference in the satisfaction rating before and after use, which indicates that the expectations were low already before drivers started to use LDW.
- Average trust in LDW changes significantly (decrease) over time, i.e. drivers expected more of the system than it could fulfil.
- Most drivers found LDW very easy to use but many also found the warning irritating and commented on the warning timing.

It is clear that while most participants find LDW useful, they also indicate low levels of satisfaction with the system (see Figure 35). This fact is also reflected among the additional items in the questionnaire where the participants perceive the system as effective and intuitive but not attractive to buy.

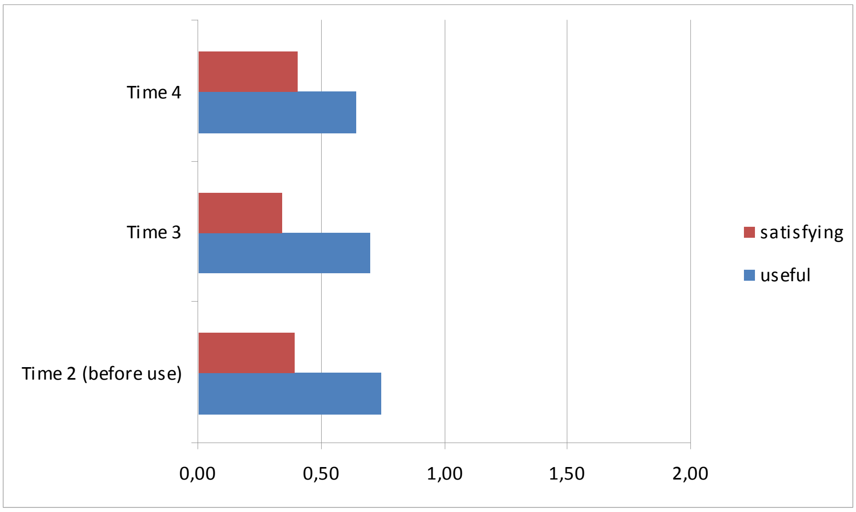

Figure 35: Acceptance rating in terms of usefulness and satisfaction (LDW)

This lack of satisfaction seems attributable to the weak coupling that exists between a warning from LDW and potential situational risk. For most drivers, drifting out of lane is only a problem given that there are objects and/or places nearby that warrant avoidance (oncoming vehicles, ditches, etc.). If no such objects are present or places near, then a leaving lane warning seems unnecessary, even though it is technically correct.

This highlights the need for future LDW systems to be capable of a more sophisticated traffic environment assessment in order to determine when a warning will be perceived as relevant by the driver. Of course, one could also envision a closer coupling to driver state assessments as well (e.g. drowsy drivers might appreciate a lane departure warning even if there is no apparent threat nearby). Regardless of which approach is selected, some method for bringing warning timing closer to perceived relevance in the driver is likely necessary.

5.2.1.2 Subjective data

The second fleet involved n=570 customers and followed them for nine months. Two groups of users were compared:

- LDW group: n=280 vehicles equipped with LDW (feedback through a torque applied through the steering wheel);
- Control group: n=290 vehicles not equipped with LDW.

A subjective field test was carried out in which drivers were asked to fill-in periodical questionnaires self-reporting their experience and perception about the LDW system (i.e. five periodical questionnaires, weekly and event registers). Vehicles were not equipped with Data Acquisition System and extra sensors. Operational Centres did not need to meet face-to-face the project participants and no installation or extra maintenance were required for vehicles.

The investigated research questions and hypotheses concern subjective users' related aspects of the function impact. The users' perceptions about these aspects have been assessed. Driver behaviour, users' workload, acceptance, usability and safety were the areas of impact of the function. In particular, users' perception of the impact of the LDW on safety (i.e. the perceived LDW impact on lateral incidents and chance of accidents and the impact on lateral driving performance) has been investigated.

According to the particular experimental design, descriptive statistics and derived aggregated values were considered very important for the interpretation of results according to research questions and hypotheses. Moreover inferential statistics were used to test trends and change of perception during test period.

Drivers involved in the Italian test site of the euroFOT project report positive perceptions of the main users' related aspects investigated in the test about LDW system impact. In particular users perceived a good impact of the system on the overall road safety and also on their behaviour of turn indicator activation. System seems to be well accepted as users recognised the LDW as useful and satisfying. LDW is useful in order to avoid dangerous situations and also driving in critical conditions such as when driver is tired or when there is a high risk to fall asleep at the wheel or at night. According to test results the system seems not to affect drivers' workload in driving conditions. That seems to fit the research expectations as the system provides warning just in specific situational conditions recognised as risky for driver safety. Drivers report some occurrences of system misuses and abuse, which seem however to be quite rare. They performed some tests in order to assess their confidence with system usage and intervention. Users report very rare occurrences in which the system is misused in order to improve performances of secondary tasks while driving with potential impact on safety.

Some of the most important results are briefly listed below.

Most of the sample (more than 90%) found the LDW system effective in increasing the driving and road safety and this perception is stable along time. In particular users perceived the system useful to avoid dangerous situations and helpful in case of falling asleep at the wheel (94% at Time 3a, 88% at Time 4). The overall driving performance seems not to be affected by the usage of LDW system, but it seems to impact on drivers' ability to keep within the lane. Also users' perception of the usage of turn indicators seem to be affected by LDW, as users recognize a positive effect of the system on this behaviour that increases over time [$F(2.230) = 5.24$, $p = 0.006$].

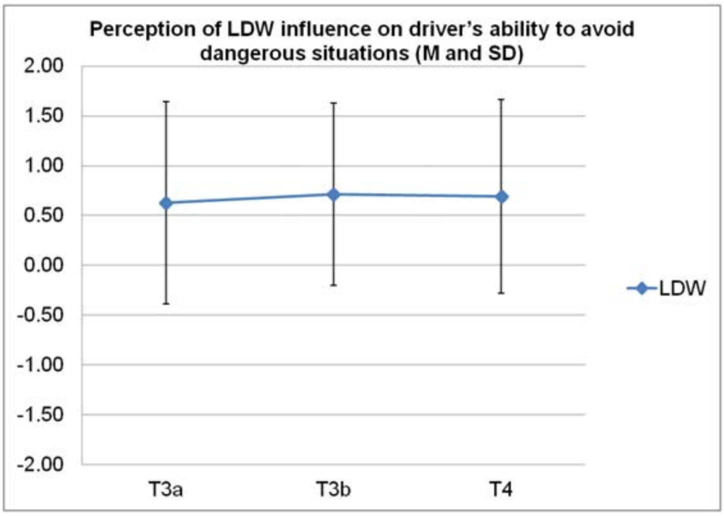

Figure 36: Subjective perception of LDW influence on driver's ability to avoid dangerous situations (M and SD)

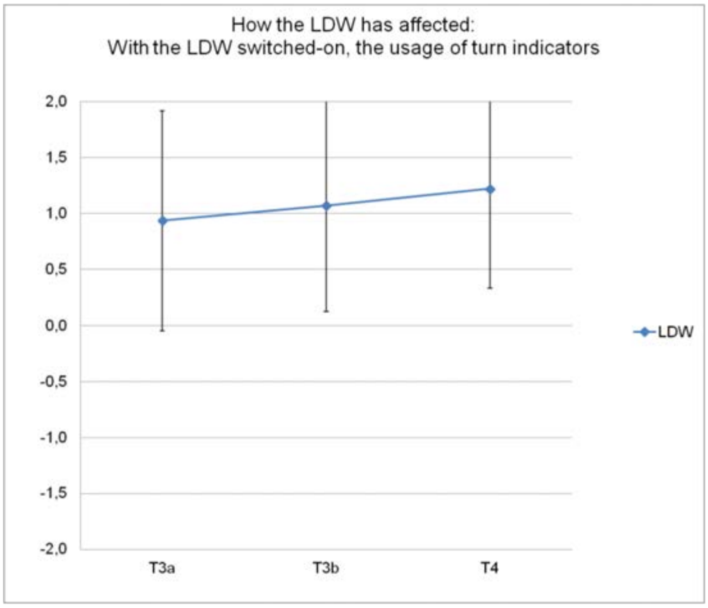

Figure 37: Responses to item "How the system has affected, with the LDW SWITCHED-ON, the usage of turn indicators?" (M and SD)

The acceptability of the LDW system is high for all the considered features. Drivers found the system very useful and satisfying with a prevalence of the former. The acceptance is stable over time. Users also report positive perception for other investigated aspects in relation to users' acceptance such as perceived quality or user-friendliness.

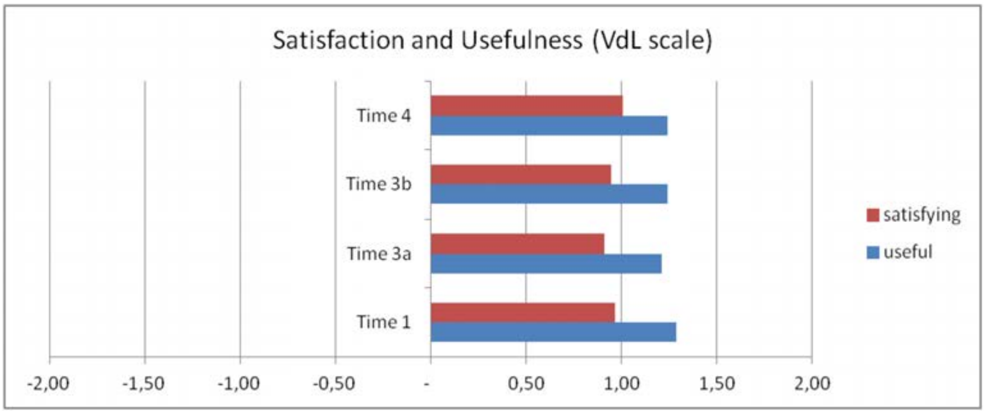

Figure 38: Acceptance rating in terms of usefulness and satisfaction (LDW)

Only a low percentage of the drivers experienced situations in which they did not trust the LDW and also misuses of the system seem to be quite rare, except for behaviours performed to test the system (i.e. make unnecessary lane changes).

Several aspects seem to impact on system usability as the ease of use, the ease of learning, the ease of remembering and the perceived system comfort. Three considered features, level of confidence, trustworthy and reliability of the LDW system are high and stable at all the time points.

About LDW distracting impact on other driving activities, the system is perceived as less impacting at the end of the test period [$F(2.236) = 3.70$, $p = 0.03$]

Users' workload is higher in some specific driving conditions as driving in poor weather conditions or driving while drowsy. No relevant different trends (for the LDW and Control groups) emerged during the test period.

The results of the questionnaire responses regarding IW can be summarised as follows:

- IW is rated very positive in terms of acceptance, satisfaction and usefulness.
- The IW ratings are stable, i.e. they do not change over time.
- Many respondents feel that IW increased safety.
- IW is perceived as most useful on motorway in normal traffic
- Trust in IW is overall high and does not change with time. This indicates that drivers agree with IW's assessment of their level of attention/drowsiness

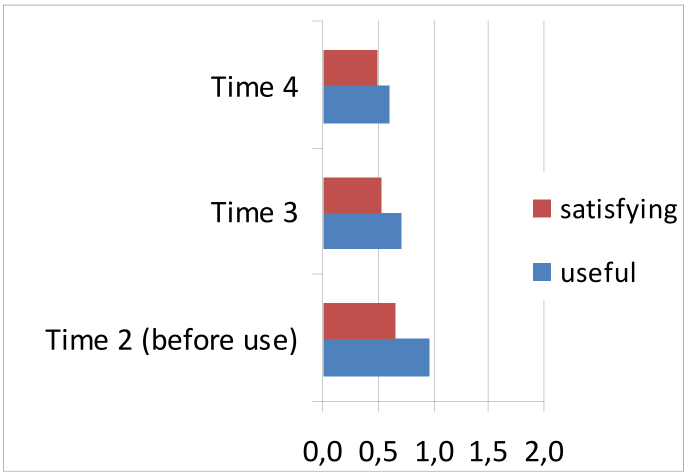

Figure 39: Acceptance rating in terms of usefulness and satisfaction (IW)

Impairment Warning scores very high on usability. This is perhaps not surprising, given the usage is extremely simple (i.e. turn the system on) and its intuitive warning interface (the coffee cup and the text "time for a break" light up when the driver's lane keeping performance indicates drowsiness). In other words, there is little interaction required, and interpreting system output is very easy. This is also reflected in the low workload score.

A highly interesting aspect of the questionnaire data is whether drivers agree with IW's assessment of their level of drowsiness/inattention. Here, drivers report a high level of trust in the system, and this rating does not change with time. Drivers thus seem to agree with the system's assessment of their level of attention/drowsiness, which is both good in terms of showing that the system is accurately tuned, and a necessary prerequisite for drivers to act on the information given.

Regarding the latter, many comments indicate that the real obstacle to efficient impairment warning may not be the detection of impairment per se, but rather finding ways and means for the driver to do something about it. The number of places to stop and take a break on the motorway is limited, and other factors such as a desire to get home (also referred to as extra motives, by e.g. [39] and [40]) show a type of social pressure that forces the driver to disregard the impairment warning.

As for potential misuse of the system, only one driver reported driving while drowsy and relying on IW to indicate when to take a break. On the other hand, as the indication IW gives seems to match the drivers' own state assessment, it is unclear whether this is to be viewed as misuse or simply efficient use of the system.

5.2.2 Safety

The LDW+IW function was expected to support the driver in avoiding unintended lane departures, either due to distraction (LDW) or drowsiness (IW). Overall, the bundle was expected to have a positive effect on both comfort and safety.

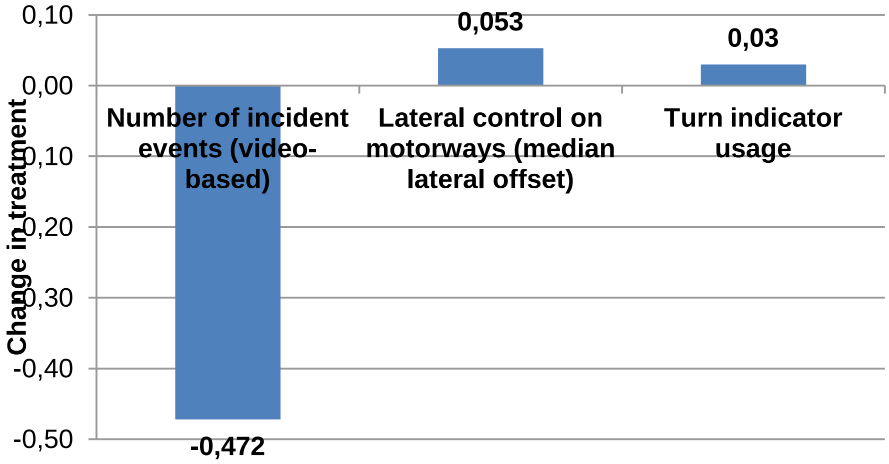

Figure 40: Overview safety indicators for LDW (passenger cars)

Based on the overall results for LDW+IW in passenger cars, it can be concluded that some of the indicators point toward an increase in safety when drivers use LDW+IW. The mean steering wheel angle was somewhat reduced and use of turn indicators increased, both of which indicate improved lateral control. Also, results show that LDW issues warnings mainly when drivers are not looking at the road ahead. Hence, it mainly addresses potentially unsafe situations. The likelihood of experiencing a lateral crash relevant event also decreased when drivers used LDW+IW. However, that decrease was not statistically significant, mainly because the number of annotated events in the end judged relevant for LDW+IW was small. Crash relevant events and near crashes are truly rare events. Reviewing over 1200 potential conflicts based on video and kinematic data only 133 were judged to be truly relevant events for this dataset, and hence retained for the analysis.

Results for trucks were very similar. Some of the indicators point toward increased safety when truck drivers used LDW. The mean lateral offset (i.e. vehicle distance from road edge) was somewhat increased and also the use of turn indicators to indicate lane changes was increased, both of which suggest improved lateral control. The likelihood of experiencing a lateral crash relevant event also decreased when drivers used LDW. However, that decrease was not statistically significant, again mainly because the number of annotated events in the end judged relevant for LDW was small. From over 1000 potential conflicts, only 19 were judged to be truly relevant lateral conflict events, and hence retained for the analysis.

We also investigated possible negative side effects of LDW+IW in terms of secondary task engagement, attention to forward roadway and drowsy trip frequency. Results showed some interesting effect of LDW+IW in the first two measures for passenger car drivers. First, during normal driving, the likelihood of a passenger car driver using a nomadic device almost tripled when drivers were using LDW+IW. However, during crash relevant events, no such difference was found, which indicates that drivers seem capable of adjusting nomadic device usage to situations where safety is not compromised. This line of reasoning is supported by the fact that there was no difference in visual attention to the forward roadway during critical events in baseline and treatment. For truck drivers on the other hand, the data did not show any difference between baseline and treatment, i.e. truck drivers had similar focus on the forward roadway and did not engage more in secondary tasks when using LDW compared to baseline.

5.2.3 Traffic efficiency

LDW and IW were not expected to have direct traffic efficiency effects. Furthermore, as described in deliverable D6.4, the safety impact results for LDW+IW did not warrant an up-scaling to the EU-27 level. Hence indirect traffic efficiency effects could not be assessed either.

5.2.4 Environment

LDW and IW were not expected to have significant environmental effects.

5.2.5 CBA

Carrying out a cost-benefit analysis was considered not being feasible for this function because of performance restrictions resulting from the Field Operational Test (non-applicable and / or insignificant impacts, insufficient knowledge on EU-wide driver behaviour and network characteristics required for up-scaling).

5.2.6 Conclusion

Overall, drivers indicate that LDW is a useful function. However, it is also clear that drivers perceive the coupling between a warning and the risk of a crash as weaker for LDW than for ACC+FCW. Hence many warnings are perceived as nuisance warnings.

It follows that in order to increase user satisfaction for future LDW systems, a more sophisticated traffic environment assessment is probably required, to better determine when a warning will be perceived as relevant by the driver. For example, if the driver drifts to the left on a non-divided rural road, the warning could be made dependent on whether there is oncoming traffic. Of course, one could also envision a closer coupling to driver state assessments as well (e.g. drowsy drivers might appreciate a lane departure warning even if there is no apparent threat nearby). Regardless of which approach is selected, some method for bringing warning issuing and timing closer to perceived relevance for the driver is likely necessary.

For IW, user ratings were highly positive and stable over time, which indicates that the function does what the driver expects it to do in a reliable way. Moreover, drivers seem to agree with IW’s assessment of their level of drowsiness/inattentiveness, which is good. The challenge rather seems to be to find a way in which the driver safely and meaningfully can act on the impairment warning.

With regard to safety aspects, some indicators pointed toward LDW+IW having a beneficial safety impact, particularly the relative frequency of lateral incidents. However, within the time and resource frame available for the analysis, it was not possible to review a sufficient number of such incidents to test whether that trend also was statistically significant. Since it was not possible to identify a significant starting point from which a national / EU-27 level safety impact could be calculated, there was no up-scaling or CBA for LDW+IW. In addition, traffic efficiency and environmental impact was not tested for LDW+IW.

5.3 Navigation Systems

The navigation system was evaluated at the German2 VMC. In the FOT navigation systems with two levels of integration - mobile device and built-in navigation system - were compared. This results in three experimental conditions: driving with a built-in device, driving with a mobile device and having no navigation system available (baseline condition). Each driver participated in each experimental condition for about one month. The order of the conditions was balanced between drivers.

Data for evaluating navigation systems was collected by two car manufacturers and that are BMW and Daimler. In total 110 drivers participated in the FOT. Of those, objective data for 99 drivers can be analysed. Objective driving data used for the analysis adds up to nearly 1 million kilometres or more than 13.400 hours of driving time. In addition to the collection of objective data, subjective evaluations of the drivers were assessed by questionnaires filled in at the beginning, during and at the end of each experimental condition.

Ford	MAN	VOLVO	VCC	VW	CRF	IFSTTAR	BMW	DAG
							✓	✓

Figure 41: Manufacturers that gathered data for navigation systems

With the help of the gathered data it was possible to answer the research questions on user acceptance and driver related aspects, safety, traffic efficiency and environment. To that end the statistical analysis of the objective and subjective data aimed to test the hypotheses that were derived based on the research questions. For navigation systems, no up-scaling and no cost-benefit analysis was conducted.

Table 17: Overview of used data

	Baseline	Built-in	Mobile	Overall
Mileage overall [km]	295.515	332.825	290.822	919.162
Mileage motorway [km]	200.329	229.329	190.128	619.787
Mileage rural [km]	45.531	51.533	49.268	146.332
Mileage urban [km]	49.373	51.670	51.108	152.152
Duration overall [h]	4.359	4.769	4.325	13.453
Proportion system active overall [%]	0.00%	47.14%	32.95%	40.39%

In Table 18 the list of hypotheses to be answered for the navigation system is presented.

Table 18: List of hypotheses for navigation systems (SafeHMI)

	Hypothesis	Performance indicator
Safety	SafeHMI decreases incidents while approaching decision points	N incidents / intersection
Safety	SafeHMI influences share of critical TLC	% time of TLC < 1.0 sec
Safety	SafeHMI influences share of critical THW	% time of THW < 0.5 sec
Safety	SafeHMI influences share of critical TTC	% time of TTC < 1.75 sec
Safety	SafeHMI influences the number of unintended line crossing	N line crossings / hour

	SafeHMI influences the frequency of hard braking	N hard brakings / hour
Traffic efficiency	SafeHMI increases journey efficiency, based upon surrogate measures.	change of relative travel time change of relative travel distance % time spent in congestion
Traffic efficiency	SafeHMI influences mean speed	m(v)
Traffic efficiency	SafeHMI influences sd speed	sd(v)
Traffic efficiency	SafeHMI influences the number of trips	N trips / day
Traffic efficiency	SafeHMI influences the number of kilometres travelled per road category	% time spent on road category
Environment	SafeHMI influences the fuel consumption per kilometre	average fuel consumption
Driver behaviour, user acceptance and driver related aspects	SafeHMI influences the frequency of hard acceleration	N hard accelerations / hour
Driver behaviour, user acceptance and driver related aspects	SafeHMI increases the time spent on secondary tasks.	% time spent on measurable secondary tasks
Driver behaviour, user acceptance and driver related aspects	SafeHMI increases compliance with traffic rules.	% time speeding % left / right turns with using turn indicator subjective rating
Driver behaviour, user acceptance and driver related aspects	SafeHMI handling occurs mainly in low demanding situations	% of system handling in specific situation divided through proportion of time spent in situation
Driver behaviour, user acceptance and driver related aspects	Handling of SafeHMI increases active compensation by the driver.	change of speed, THW and sdlp in periods of system handling
Driver behaviour, user acceptance and driver related aspects	Handling of SafeHMI does not decrease safety, based upon surrogate measures.	change of % time of THW < 0.5 sec, % time of TTC < 1.75 sec, % time of TLC < 1.0 sec and N line crossings / hour in periods of system handling
Driver behaviour, user acceptance and driver related aspects	SafeHMI decreases driver load at decision points.	subjective rating m(v) on intersection % time spent driving very slowly before intersection
Driver behaviour, user acceptance and driver related aspects	SafeHMI increases perceived driving comfort.	subjective rating
Driver behaviour, user acceptance and driver related aspects	Acceptance and trust of SafeHMI will increase with experience.	subjective rating
Driver behaviour, user acceptance and driver related aspects	System usage will increase over time	% time navigation system active
Driver behaviour, user acceptance and driver related aspects	Driver workload decreases over time with system use.	subjective rating
Driver behaviour, user acceptance and driver related aspects	User practices (heuristics, rules) will change over time during the FOT	subjective rating
Driver behaviour, user acceptance and driver related aspects	Drivers will not abuse or misuse navigation systems	subjective rating
Driver behaviour, user acceptance and driver related aspects	Type of SafeHMI affects SafeHMI interaction.	subjective rating % time navigation system active

5.3.1 User acceptance and user related aspects

The analysis of subjective ratings of the two tested HMI-solutions shows a strong and overall preference of the built-in navigation system compared to the mobile device. The analysis especially of open questions indicates that this difference is not only based on the HMI-solution but at least also partly on other characteristics of the systems (e.g. selection of routes).

The detailed analysis of user acceptance over time reveals that the expectation on both HMI solutions is positive but for the mobile device that expectation is not fulfilled. The subjective evaluation of the two systems is reflected in objectively measured usage: the mobile device is used less often than the built-in system and its usage decreases over time. Overall, compared to other function (e.g. ACC) navigation systems are highly used systems. Even the less popular mobile device is used between 30% and 40% of total driving time.

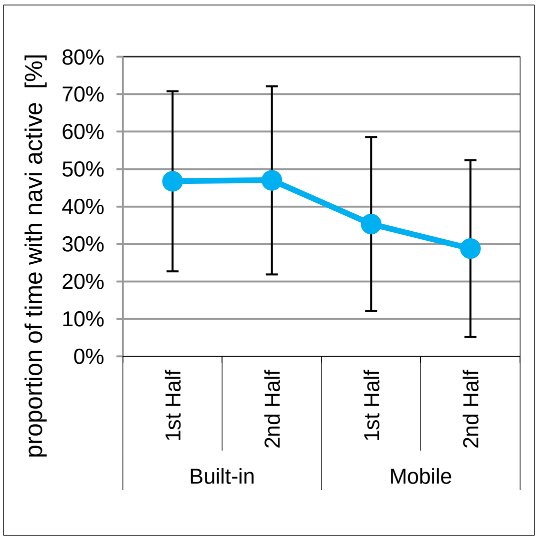

Figure 42: Change of objective system usage over time for the built-in and the mobile device

A further analysis of system usage shows that it depends on the familiarity of a trip and the length of a trip. The navigation system is activated more often on long and on unfamiliar trips. The mobile device is used less often than the built-in device especially in situations where overall system usage is less likely (short trips, unfamiliar trips).

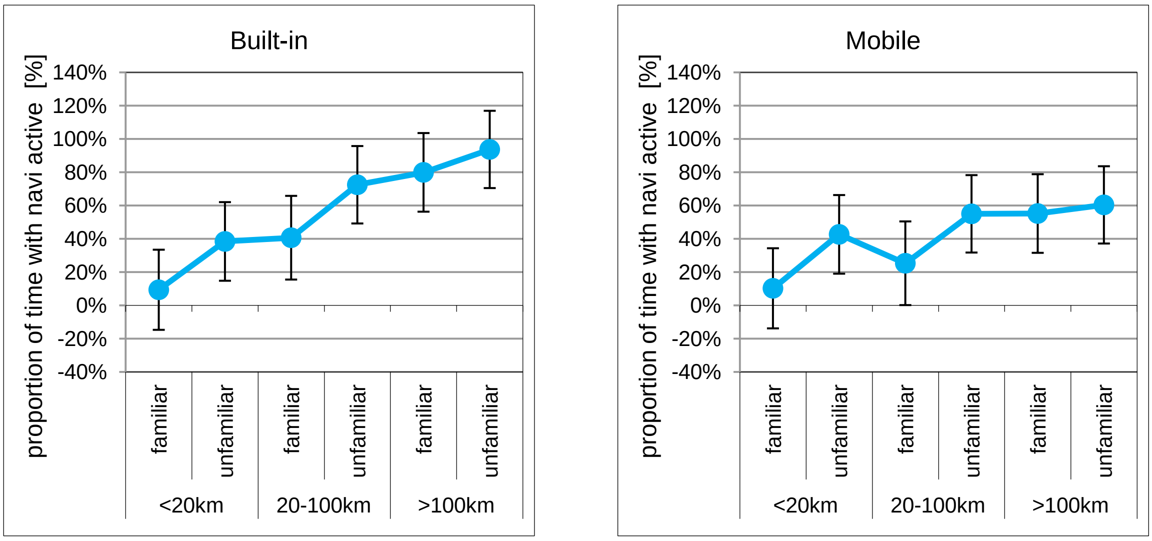

Figure 43: Proportion of trips with active navigation system separate for familiarity of route and trip length

Regarding driver behaviour one area of interest specific for navigation systems was analysed. Results indicate that drivers prefer low demanding situations like standstill or very low speeds for handling of the navigation system. In case they make system inputs while driving, they adapt their driving behaviour directly prior and during the system input in order to compensate for the extra load during system handling. As can be seen in Table 19, there are significant adaptations in speed and distance for both HMI-solutions in urban and rural areas. Furthermore, results indicate that the handling of the mobile device is more complex since lane keeping performance decreases during system inputs on highways and urban areas. An analysis of indicators related to driving safety shows for both HMI-solutions no increase in safety critical driving behaviour related to system inputs.

Table 19: Summary of results regarding the impact of system handling on driving

	Built-in			Mobile		
	Motorway	Rural	Urban	Motorway	Rural	Urban
Average speed		<<	<<		<<	
Distance to lead vehicle		>	>>		>	>>
Lane keeping performance				<<		<<

>> indicates a significant increase (p<0.05), > a tendency for an increase (p<0.1), << indicates a significant decrease (p<0.05), < a tendency for a decrease (p<1). Empty cells show that there was no statistically significant effect

5.3.2 Safety

Because of the functionality of a navigation system, a potential safety benefit is expected to be most likely in urban and rural areas and here especially prior to or on intersections. Several indicators relating to driving safety were analysed for urban and rural areas. For the built-in device most of the indicators show that driving safety is significantly enhanced if the navigation system is active. For the mobile device, a significant safety benefit can be found as well but it is less global than the effect for the built-in navigation system. Table 20 gives an overview of the safety effects of navigation systems.

Table 20: Overall safety benefit of the navigation system

PI	Change			
	Rural		Urban	
	Built	Mob	Built	Mob
Incidents*	0	0	-5.4%	-21.5%
Critical THW			<<	<
Critical TTC	<<	<<	<<	<<
Hard brakings			<<	
Critical TLC			<<	
Lane exceedances	<<		<<	

* For incident frequency the change in percent of baseline condition is calculated. For all other indicators << indicates a significant reduction (p<0.05), < a tendency for a reduction (p<0.1). Empty cells indicate that the result was statistically not significant.

Looking at the differing results for the two HMI-solutions, the results for the mobile device seem less reliable than the results for the built-in navigation systems. This is not only because the effect is less global (only based on distance events) but it is also because drivers rated the mobile device as less favourable and as distracting. As a consequence, usage is lower. Therefore, the safety results are based on a smaller set of data with system active. The comparably high change in incident frequency for the mobile device is caused by a higher number of drivers with zero incidents while driving with the mobile device. Assuming that the likelihood of having at least one incident increases with travel time (exposure), probability of having no incident at all in one condition rises as time using the system gets smaller. Furthermore, it is also possible that drivers try to compensate for the expected errors of the mobile device by driving overcautiously.

Overall, the results indicate that driving with a navigation system activated relates to an enhancement of driving safety. This is a potential benefit of a navigation system that has not been in the focus of analysis up to now. Therefore, no experimental studies are known that explore the mechanisms behind a safety benefit of navigation systems in more detail. This makes an interpretation of the FOT-results regarding the impact of navigation systems on driving safety difficult.

5.3.3 Traffic efficiency

The main goal of a navigation system is to improve travel efficiency especially on unfamiliar routes. To evaluate that function, the measured travel times and distances were compared to travel times and distances that were estimated with a reference route planner.

Results show that for both navigation systems, relative travel time is shorter compared to driving in the baseline condition, where no navigation system was available. Furthermore, for the built-in navigation system, also relative travel distance is significantly reduced. Besides the routing function, all studied navigation systems also offer dynamic rerouting to help the driver to avoid traffic congestion. To analyse the effectiveness of that function, the proportion of time spent in congestion is analysed. In the available data, no proof can be found that the dynamic rerouting does successfully support the driver to avoid traffic jams. The proportion of time spent in congestion does not significantly change if the navigation system is active. For the analysed indicators, Figure 44 shows the mean change in percent of baseline values.

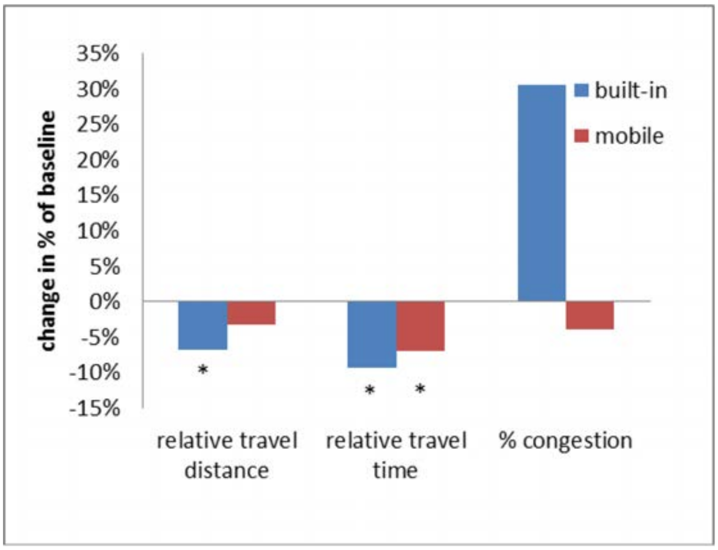

Figure 44: Change in percent of baseline for analysed indicators for travel efficiency

(* indicates that the change is significant)

To keep the presentation of results harmonized over the document, Figure 44 shows means. The larger mean increase of time spent in congestion is caused by a few outliers. For the whole sample, there is no significant change of the proportion of time spent in congestion while using a navigation system.

5.3.4 Environment

To evaluate the impact of navigation systems on the environment, average fuel consumption is analysed. For the built-in navigation system, it decreases significantly on urban and rural roads if the system is active. For the mobile device the decrease is not significant.

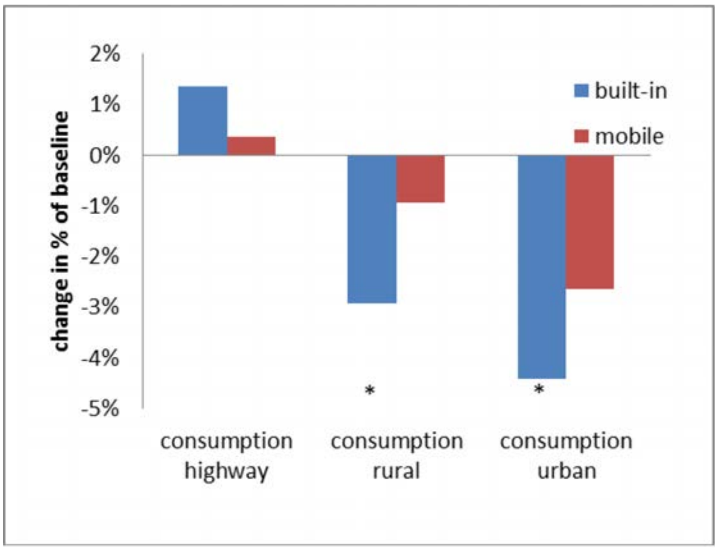

Figure 45: Change of average fuel consumption in percent of baseline

(* indicates that the change is significant)

Further analyses show that the routing algorithms used by the two types of navigation systems differ. The tested built-in systems prefer staying on main roads on which average speed is higher. Compared to that, the mobile device more often chooses also smaller routes (e.g. residential routes). As shown in the section on traffic efficiency, both approaches proofed equally efficient regarding the reduction of relative travel time. Nevertheless, it is assumed that the difference between the two systems regarding their effect on fuel consumption is related to the differences in route choice. This means, that depending on the routing algorithm used by a navigation system, the system has the potential not only to shorten travel times but also to reduce fuel consumption on the way driven.

5.3.5 CBA

Carrying out a cost-benefit analysis was considered not being feasible for this function because of performance restrictions resulting from the Field Operational Test (non-applicable and / or insignificant impacts, insufficient knowledge on EU-wide driver behaviour and network characteristics required for up-scaling).

5.3.6 Conclusion

The analysis conducted in euroFOT shows that navigation systems are highly accepted and also widely used driver assistant systems. Based on the collected data, several positive effects of driving with a navigation system active were found:

Navigation systems reach their main goal that is to support the driver to choose an efficient route. Results indicate that route choice while driving with a navigation system active is more time efficient than in baseline condition where no navigation system was available. Furthermore, depending on the routing algorithm used by the navigation system, navigation systems seem to be able to support a fuel efficient route choice, too.

Especially in urban areas, driving is safer if a navigation system is active. The main problem by interpreting the FOT-results on driving safety is that compared to other systems navigation systems are less widely studied in the literature. Especially a potential safety benefit of navigation systems has not been in the focus of research up to now. This makes the interpretation and up-scaling of the FOT results very difficult.

Although a potential safety benefit has not been studied in the literature, potential negative safety effects of handling of navigation systems (e.g. entering a destination) are widely investigated. Results from euroFOT imply that safety decreases through system handling reported in the literature might be overestimated. The FOT data indicates that drivers first of all prefer making system inputs in low demanding driving situations (e.g. standstill) in which a potential negative impact on driving safety is very unlikely or - like in standstill – even impossible.. This choosing of appropriate situations for interacting with the navigation system is normally not possible in experimental setups. Here, drivers are instructed when and where to interact with the system. Second, in case system inputs occur while driving, drivers compensate the distraction by adapting their driving behaviour accordingly. This way of compensation is ex-pected from experimental results. In the literature, it is normally concluded that although compensatory behaviour occurs on the level of driving, this is not sufficient to compensate for the change in reaction time. To study this possibility, parameters directly related to driving safety (e.g. critical distances) have been evaluated. No increase of critical events during sys-tem inputs can be found. Instead, if significant results occur they even indicate an increase of driving safety during system inputs. Whether these changes during system inputs reflect a real increase in driving safety or are caused by a change towards a more defensive driving style cannot be decided from the FOT-data. Overall, the results indicate that the impact of system inputs on driving safety reported in the literature might be overestimated because drivers do not only compensate on the level of adjusting certain aspects of driving (e.g. dis-tance) but also by choosing appropriate driving situations (like standstill). Since this result is not in-line with the literature, the impact of other ways of compensation should be investigat-ing this in more detail in the future. One possibility is to use the results from real driving data obtained in the FOT to adapt experimental setups. Through this it might be possible to study the effects of handling a navigation system on driving in an experimental setup that is closer to real driving than the setups used at the moment.

5.4 Speed Regulation System (SRS)

The Speed Regulation System (SRS) included two functions: speed limiter and cruise control. It was evaluated by the French VMC at a test site near Paris. Data from two passenger cars (Renault Clio and Renault Laguna) was gathered during the FOT (12 months) which lasted at least three months for baseline period and six months of treatment. During this time 35 passenger cars travelled more than 500.000 kilometres.

Ford	MAN	VOLVO	VCC	VW	CRF	IFSTTAR	BMW	DAG
						✓		

Figure 46: Manufacturers that gathered SRS data during the FOT

In the table below the used data for the evaluation of the SRS is presented. In total more than 500.000 kilometres are processed and considered for the analysis.

Table 21: Overview of used data

Mileage [km]	545.340
Hours driven [h]	12.590
Number of drivers	35

In addition to the collection of objective data subjective impacts of the drivers were assessed by three questionnaires that were filled in by the drivers during the progress of the FOT.

Table 22: SL/CC hypotheses list for objective data

	Hypothesis	Performance indicator
Safety	Using SRS, the number of incidents will decrease	Number of incidents per 100 driven km
	SRC decreases the number of critical time gaps to the leading vehicle	Presence of a critical time gap event in the chunk
	SL/CC use increases over time	Average per month of the driver's % of SL or CC usage per day
	Using SL/CC, the number of harsh braking will decrease.	Presence of a hard braking event in the chunk
	Using SL/CC, the number of strong jerk will decrease.	Presence of a strong jerk event in the chunk
	Using SL/CC, reduces speeding occurrences	Presence of an over speeding event in the chunk
Safety	SL/CC decreases average speed.	Average speed per chunk
	The number of trips made will increase	Number of trips per month
	The number of vehicle km travelled will increase	Number of km travelled per month
Environment	SL/CC reduces the average fuel consumption	Average fuel consumption per chunk
Driver Behaviour	CC will be used more on roads with few curves or intersection	Presence of a CC-active event in the chunk, and relation with intersection or curve density
	SL will be used more on roads with few curves or intersection	Presence of a SL-active event in the chunk, and relation with intersection or curve density
	The SL/CC selected speed will be below legal speed for non-sensation seekers drivers and will be above legal speed for sensation seekers drivers	Differences between SL or CC selected speed and legal speed limit

Table 23: SL/CC hypotheses list for subjective data

	Hypothesis	Performance indicator
Acceptance	SL/CC increases driving perceived pleasure to drive and comfort.	Subjective rating
	Certain features of the functions, in terms of usability, influence acceptance.	Subjective rating
	Certain features of the functions, in terms of usefulness, influence user acceptance?	Subjective rating
	The level of SL/CC acceptance will increase with SL/CC experience	Subjective rating
Trust	The level of SL/CC trust will increase with SL/CC experience.	Subjective rating
Workload	Driver workload decreases over time with function use.	Subjective rating
User practices	User practices (heuristics, rules) will change over time during the FOT	Subjective rating
Abuse/Misuse	Drivers will not abuse or misuse SL/CC	Subjective rating

5.4.1 User acceptance and user related aspects

The SRS system is used on demand, according to the driving situations chosen by drivers. Moreover, drivers tend to use more one of the two systems than the other. The cruise control is often used on motorways or freeways (>= 110 km/h) and speed limiter is often used on all roads except motorways (130 km/h).

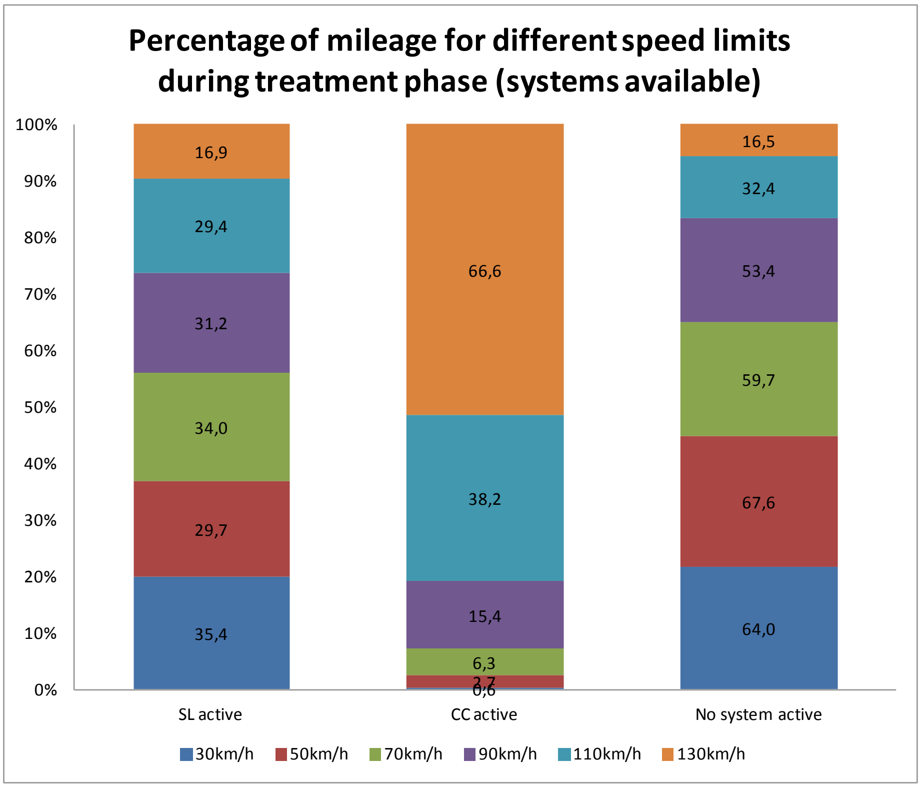

Figure 47: Percentage of mileage for different speed limits during treatment

With the SL system drivers state that they engage in misuse behaviours only to select a top speed above the speed limit and to use buttons to adjust SL speed instead accelerator pedal. With the CC system drivers state that they engage in misuse behaviours only to select a cruise speed above speed limit and to use CC to overtake a vehicle. For both systems there is no systematic change of workload over the period of system usage.

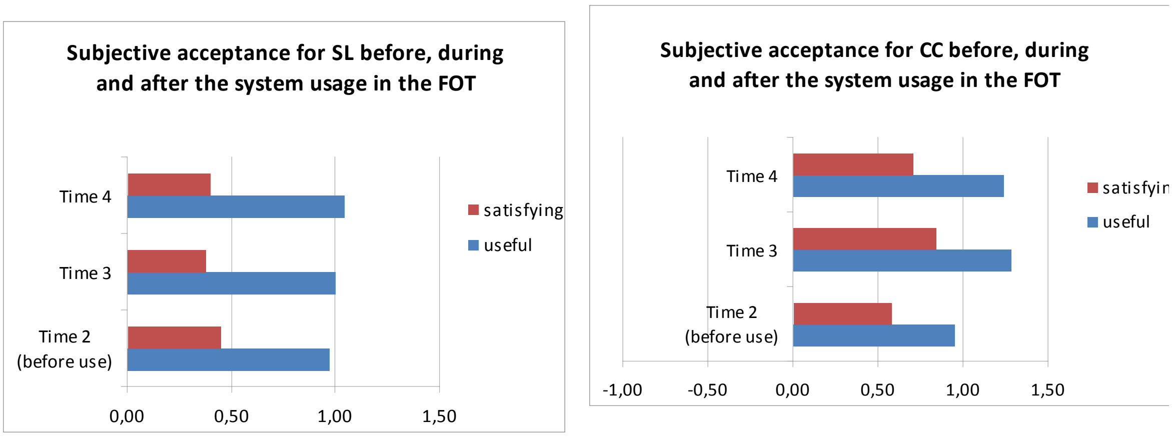

Figure 48: Subjective acceptance for the SL (left) and CC (right) system during the FOT

For the two systems, drivers have positive expectations at the beginning of the FOT and expectations are confirmed and overall there is no significantly change of acceptance and trust over time. The system SL was judged as necessary, good, assisting the driving and useful. The CC system was judged as necessary, good, assisting the driving desirable, pleasant and useful.

The SL is perceived as increasing the driver comfort for 46% of the drivers and the CC for 80%. For the speed limiter system, this increasing is not statistically significant (< 50%) but for the CC, this increasing is significant. The SL is perceived as increasing the pleasure to drive for 35% of the drivers and the CC for 63%. For the speed limiter system, this increasing is not statistically significant (< 50%) but for the CC, this increasing is not significant due to the low number of driver.

5.4.2 Safety

SRS system refer to two different systems with different purposes: SL is used to limit the vehicle speed on a voluntary basis, while CC is used to maintain a constant speed when the driving conditions allow it. These two systems cannot be active at the same time, and the choice to activate it or to change from system to another one largely depends on various parameters of driving conditions that are difficult to control. Intersections, curves, or rush hours reduce the likelihood of using these systems. Despite SL is used on almost all the speed limits, CC are used much more on high speed roads (110 or 130 km/h roads). Moreover, it has been shown that when SL is active the likelihood of observing an over-speeding event is divided by 2 and hard braking occurrences by 30%, but very small or insignificant effects are found for other events (strong jerk or critical time gap). The CC effect on over-speeding is opposite showing a strong increase, but strong jerk, critical time gap, and hard braking occurrences are divided by 3 (not surprising if CC is used under free flow conditions)

These findings highlight the relationships between systems usage and driving conditions, showing that level of traffic is likely to be an important parameter. Although the congestion level cannot be estimated for the euroFOT data, there are sufficient clues to make the following hypotheses:

SL is used on all speed limits, but mainly when the likelihood of being caught by a speed enforcement camera is high (road with low speed limits or low congestion level).

CC is used in high speed roads under free flow conditions that allow to speed without deteriorating safety too much (drivers use the system to speed, but when it is possible to do so).

It is not possible to check such assumptions with the euroFOT data sets because of a lack of information about congestion levels. The fact that SL or CC usage is closely related to driving conditions leads to some difficulties in up-scaling the safety effects for the following reasons:

Speed-accident relationships are in the form of single formulae statistically estimated using data merging all driving conditions.The increasing speed observed both when SL or CC active would lead to a negative impact on safety, but this do not take into account the fact that SRS usage is stronger for free flow conditions associated with an high safety level.This kind of behaviour is quite frequent for longitudinal systems: Due to a higher safety level for the use cases of cruise control (or adaptive cruise control), drivers tend to drive faster to maintain constant their own acceptable risk level. It is likely that higher risks due to increased average speed may be compensated by the absence of congestion, and the ability of the driver to concentrate on other driving tasks.

The SL part of the SRS is more suitable for a classical analysis based on speeds, but as the two systems cannot be used at the same time, no one can know what would have been the behaviour of the euroFOT drivers if they get only the SL in the car.

The SRS system is made for comfort purposes, and not designed for safety issues, while interactions can happen depending on driver's usage. euroFOT data helped showing that no critical or unexpected safety effect occurs when using that system, despite the fact that using it to speed will lead to a negative impact. SRS impact therefore depends largely on the way drivers use it. Up-scaling the behaviour of 35 French drivers to the European level is not realistic without any clear picture of differences of usage among European countries. This would have been feasible if a strong evidence of a safety impact would have been found. However, the emerging picture of the system is not clear enough (small effects) to be representative.

5.4.3 Traffic efficiency

The SRS slightly increases average speed. The traffic efficiency effect is determined by the change in average speed, since from the analyses it resulted that SRS does not affect mobility behaviour, route choice or choice of road type.

The traffic efficiency effect of the sub functions Cruise Control and Speed Limiter cannot be determined because of the experimental set-up. Since the analyses show that CC is used at higher speeds, the average speed when CC is active is higher than the average speed. When SL is active, speeds are about the same as the average speed. When SL and CC are off speeds are lower than the average speed. These results are caused by the fact that CC and SL are active under certain conditions which are not representative for the whole data set.

In Table 24 a summary of effects of SRS on the FOT vehicles can be found. This is the effect of having the system in the vehicle (treatment period), versus not having it (baseline period). It is not the effect of an active system versus an inactive one. The results show that, on average, SRS-equipped vehicles record higher speeds than unequipped vehicles on all road types. There is no effect on the average length of a trip, which suggests that mobility patterns (trip choice, modal choice and route choice) are not affected by the function. These effects on speed and mobility translate into lower average travel times and less delay. Note that the term delay is not used here to indicate extra travel time as a result of a driver not being able to travel at his desired speed (up to the legal limit) because of an external influ-ence, e.g. weather or congestion.

Table 24: Summary of effects on FOT level (only passenger cars) of Speed Limiter and Cruise Control

Change in average travel time per trip	Urban	-1.4%
	Rural	-0.8%
	Motorway	-2.4%
Change in mobility (kilometres driven per trip)		0%
Change in average speed	Urban	+1.4%
	Rural	+0.8%
	Motorway	+2.4%
Change in delay (range, depends on road type)	Incidental delay	not significant
	Recurrent delay	-5.5% to -3.5%

Table 24 shows the results that were obtained directly from the FOT data. Additionally, microscopic traffic simulations have been performed for SRS to determine the interaction effects with higher penetration rates of SRS equipped vehicles.

Additionally to the results given above, the main results from the simulations are:

- The use of SRS makes the speed distribution narrower, so the variation in speeds is reduced. This is even more so for the CC function.
- Most of the speed increase takes place on rural roads and urban roads.
- Cruise control is mainly used on motorways in free flow situations. Speed limiter is used on all road types, but most on urban roads and rural roads (around half of the kilometres driven). CC is used on longer trips.

The simulation results show that the average speed increases linearly with the penetration of equipped vehicles. There are no interaction effects on traffic efficiency indicators average network speed and average delay. Interaction effects are visible for safety indicators, such as small time headways and short times to collision. For the scaling up of the traffic efficiency effects, the effects found in the FOT are used and the effect is considered linear for higher penetrations.

5.4.4 Environment

The objective data shows a significant influence of SL usage to the fuel consumption on all road types. The reduction varies between 1.55% on motorways, 3.75% on rural roads, and 5.19% on urban roads. These results might be influenced by the choice of the driver when to use the system. The higher influence of the SL at lower vehicle speeds is due to thermal engine fuel consumption which is lower for high speeds until 90 km/h. For higher speeds the fuel consumption increases again. The reduction can be attributed to a more constant speed while using SL.

The SL environmental effects are projected for driving on motorways. It is assumed that the usage rate derived from the FOT data can be transferred to the whole EU-27. Combining the usage rate with the reduction in fuel consumption a fuel saving of 0.26% could be achieved in the European passenger vehicle fleet.

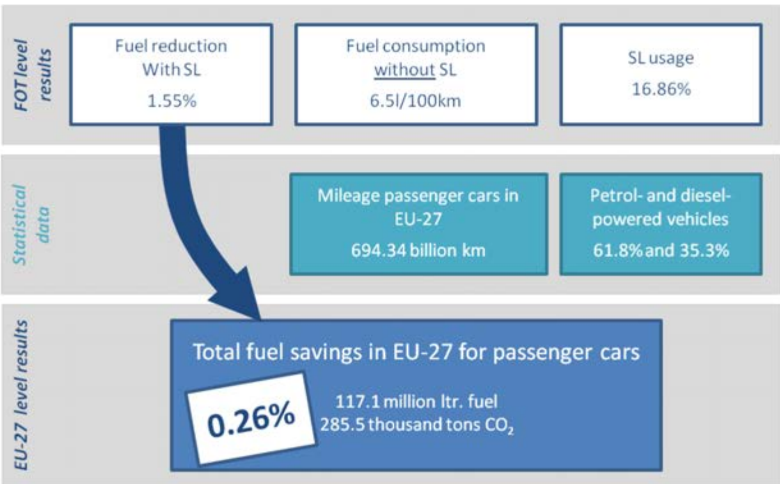

Figure 49: Potential in fuel saving with SL equipped passenger cars for EU-27

For all road types a significant influence of the use of CC on the fuel consumption was found during the analysis. The reductions vary between 1% for motorways, 13.2% on rural roads, and 36.3% for urban roads. However, the high influence on urban roads might be caused by the selection of driving situations when the system is used. This inherent bias in the FOT data on CC leads to precaution in the interpretation of the results as they do overestimate the benefits. The low usage rate of 2.69% is an additional indicator that the system is only used under certain driving situations whose driving pattern might be different from the rest of urban driving.

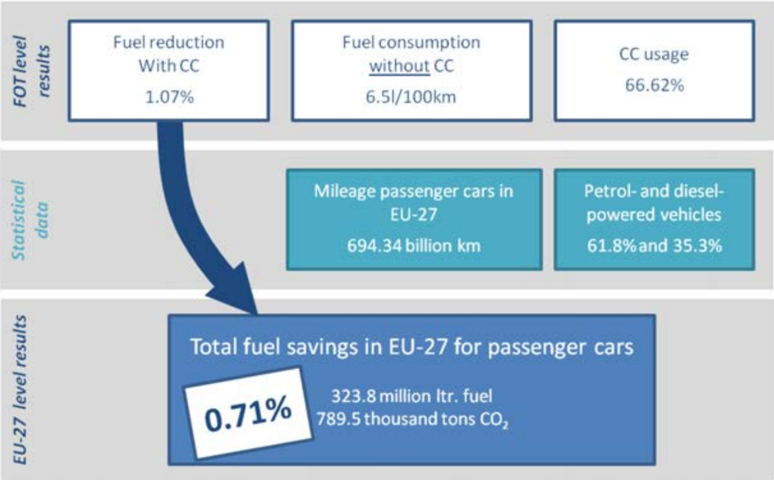

Figure 50: Potential in fuel saving with CC equipped passenger cars for EU-27

A surprising result from the simulation is that the homogeneity of the traffic flow is independent of the penetration of SRS equipped vehicles. This means that the traffic efficiency and environmental effects can be considered linear with the penetration.

5.4.5 CBA

Carrying out a cost-benefit analysis was considered not being feasible for this function because of performance restrictions resulting from the Field Operational Test (non-applicable and / or insignificant impacts, insufficient knowledge on EU-wide driver behaviour and network characteristics required for up-scaling).

5.4.6 Conclusion

The French VMC allowed a study the behaviour of 35 drivers using a system of speed regulation and limitation during a 9 months period. As expected, the drivers use the cruise control (CC) a lot during free flow conditions and on motorways (40% and 66% of the crossed distances). On the other hand the speed limiter (SL) is used a third of the crossed distances on every type of ways (except on motorway: 16%).

Both systems are less used in curves and the CC is less used where there are intersections. The use of both systems did not change during the experiment; Most of the drivers consider that these systems allow increasing the safety, although both systems do not change the workload, or the number of incidents.

This experiment allowed us to estimate the acceptance of these systems. The a priori acceptance was already very positive and the use of the systems confirmed this tendency. In terms of usefulness, the opinion of the drivers slightly increased for the SL and strongly for the CC whereas the satisfaction slightly decreased for the SL and increased for the CC.

Also, the trust in the system was already positive before its use and during its use, it slightly strengthened for the SL and very strongly for the CC especially in terms of reliability and trust in the system. Both systems were considered very usable by two thirds of the drivers as well in terms of access to visual information, manipulation of the commands and management of the interactions. The only function collecting only half of the positive opinions is the access to on/off button of the system, which is indeed not easily accessible in both vehicle types used for the experiment.

In term of comfort, a third of drivers did not report change for the SL, half reported an increase and the others a decrease. In terms of pleasure to drive, the proportion of positive change is lower (35%) and that of negatives is more important (37%). It is true that the SL does not lead a lot of change in the driving because it intervenes only to prevent from exceeding a given speed.

This can lead to inconveniences in certain situations (for example, during insertion on a motorway if the driver forgets to change the speed of the system). The fact that the driver continues to manage the longitudinal control can explain that the influence of the system in term of critical headway, of hard braking and strong jerks is low. On the other hand it allows reducing considerably the over-speeding on freeways and the consumption of fuel, although the average speed slightly increases by 2 km/h.

For the CC, opinions of the drivers are much more positive: 80% report improved comfort, and 62% report improved pleasure to drive. This can be explained by the fact that the system allows to unload the driver of a part of the driving task. This automation also has effects in terms of reduction of critical time gaps, strong jerks, and hard braking as well as consumption of fuel although the average speed increases by more than 10 km/h.

These two systems cannot be active at the same time, and the choice to activate it or to change from one system to another largely depends on various parameters of driving conditions that are difficult to control. The previous findings highlight the relationships between systems usage and driving conditions, showing that level of traffic is likely to be an important parameter. Although the congestion level cannot be precisely estimated for the euroFOT data, there are sufficient clues to make the following hypotheses:

- SL is used on all speed limits, but mainly when the likelihood of being caught by a speed enforcement camera is high (road with low speed limits or low congestion level).
- CC is comfortable to use in conditions when constant speeds can be maintained. People use the CC when the road and traffic conditions allow for fast driving.

The increasing speed observed both when SL or CC are active would lead to a negative impact on safety, but this does not take into account the fact that SRS usage is stronger for free flow conditions associated with an high safety level.

This kind of behaviour is quite frequent for longitudinal assistance systems: Due to a higher safety level for the use cases of cruise control (or adaptive cruise control), drivers tend to drive faster to maintain constant their own acceptable risk level. It is likely that higher risks due to increased average speed may be compensated by the absence of congestion, and the ability of the driver to concentrate on other driving tasks.

The low usage rate of 2.69% is an additional indicator that the system is only used under certain driving situations whose driving pattern might be different from the rest of urban driving. Further investigations in a follow up project are needed to improve baseline selection in order to eliminate bias due to driver’s system usage related to specific driven conditions.

5.5 Blind Spot Information System (BLIS)

BLIS was evaluated at the Swedish VMC in the passenger car fleet. Drivers experienced two conditions; 1) driving without the system for about 4 months (baseline condition) and 2) driving with the system available for use during approximately 8 months (treatment condition). While subjective data (i.e. questionnaire responses) is available for the full duration, due to delays in the collection of objective data, the objective data analysis is based on 3 months of baseline and 3 months of treatment.

Ford	MAN	VOLVO	VCC	VW	CRF	IFSTTAR	BMW	DAG
			✓					

Figure 51: Manufacturers that gathered BLIS data during the FOT

In total 98 drivers participated in the evaluation of BLIS. Table 25 shows the number of drivers for which complete sets of objective and subjective data was available for the analysis of BLIS. Note that for different hypothesis, the number of complete data sets varied, with the lowest being 58 and the highest 73.

Table 25: Number of drivers with complete data sets available for the data analysis

	Number of drivers
Questionnaire data	58 - 73
Objective data (6 months)	58 - 73

Table 26 gives an overview over the number of kilometres and hours of driving on which the analysis for BLIS is based. The number of respondents varied between the questions.

Table 26: Description of objective data used for the analysis of BLIS

	Baseline	**Treatment**
Mileage overall [km]	347.774	492.976
Mileage motorway [km]	89.028	119.003
Mileage rural [km]	87.157	111.496
Mileage urban [km]	116.555	160.420

With the help of the gathered data it was possible to answer the research questions on user acceptance and driver related aspects, safety, traffic efficiency and environment defined at the beginning of the project, but not a full safety impact and CBA (reasons are explained in section 3.5.5). Statistical analysis of the objective and subjective data was carried out to test the hypotheses derived based from the research questions. In Table 27 and Table 28 the list of hypotheses addressed is presented.

Table 27: Hypothesis tested with objective data

	Hypothesis
Safety	BLIS decreases the number of crashes, near crashes and incidents
	BLIS decreases the use of turn indicators

Table 28: Hypothesis tested with questionnaire data.

Note that drivers responded separately to LDW and IW questions, and hence questionnaire based answers are available on a per function level in D6.3. Since they hypothesis tested are the same however, they are grouped together in this table.

	Hypothesis
Acceptance	Certain features of the LDW and IW functions, in terms of usefulness, influence acceptance
	Acceptance changes over time with function use
	Certain features of the functions, in terms of usability, influence acceptance.
Trust	Trust in function changes over time with system use
Workload	Driver workload decreases over time with function use.
User practices	User practices (heuristics, rules) will change over time during the FOT
Abuse/Misuse	Drivers will not abuse or misuse LDW or IW

5.5.1 User acceptance and user related aspects

In euroFOT, the impact of Blind Spot Information System (BLIS) on driving was investigated. The results of the FOT on BLIS can be summarised as follows:

- Overall usability and acceptance scores for BLIS are very high (over 90% positive ratings), and this rating does not change over time.
- Approximately 80% of drivers feel that BLIS increases safety.
- BLIS is perceived as most useful on urban roads in heavy traffic.
- BLIS does not increase workload.

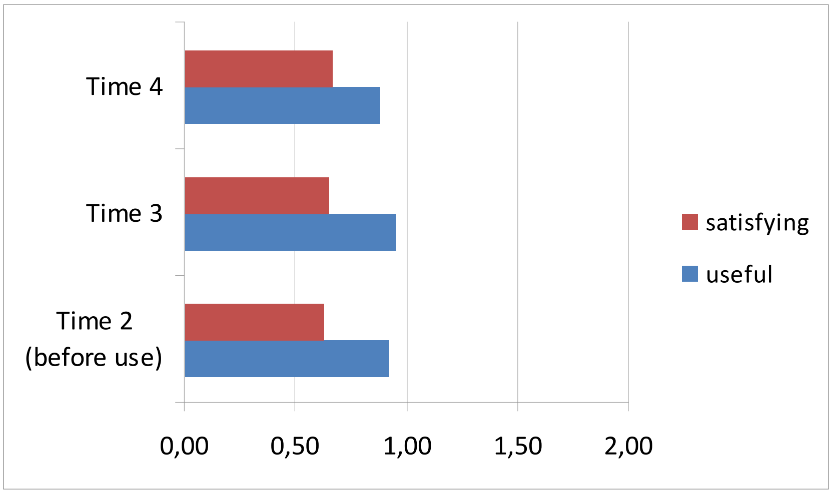

Figure 52: Acceptance rating in terms of usefulness and satisfaction (BLIS)

BLIS gets very good feedback from the drivers who have evaluated the system. A majority of drivers only gave positive responses on all the questionnaire items related to BLIS. Furthermore, acceptance does not change over time, i.e. it is stays high with continues use of the system, which indicates that drivers continue to perceive it as useful as they experience interacting with it over an increasingly large variety of conditions.

BLIS does not impair workload, and workload scores remain low throughout treatment. This is not very surprising, as the interface is highly intuitive and as simple as it can be (light indicators are situated on A-pillars; light on indicates vehicle in blind spot).

From the free text comments, it is clear that many view BLIS as an important complement to visual checks, rather than as a primary source of information. This probably relates to the fact that the system at times gives false positives, i.e. it lights up its indicator even when there is no vehicle in the blind spot.

Drivers seem well aware of this performance limitation in the system, i.e. they score the system low on items related to reliability and trust. Furthermore, while a few drivers stated that they substitute visual checks with BLIS information, a closer reading shows that this is not unconditionally done. Rather they over time seem to learn where and when the BLIS indications are highly accurate, and that's where they feel comfortable relying on BLIS only.

Thus, interestingly, drivers are overall very pleased with a system that gives false positives on a more or less regular basis. This goes to show that as long as the core functionality of the system is perceived as useful, drivers seem willing to learn where the limitations are and/or "forgive" the systems for this type of imperfections

5.5.2 Safety

The BLIS function was expected to support the driver in avoiding initiating lane changes when there is a vehicle in the adjacent lane, in particular when that vehicle cannot be seen through the regular rear view mirrors. However, the frequency of lateral incidents studied for BLIS showed no difference between baseline and treatment. The one indicator that did show a significant difference was the use of turn indicators, which decreased by approximately 10% when BLIS was in use. This is an interesting contrast to the LDW+IW findings, which indicate the opposite, i.e. a 10% increase in turn indicator use when LDW+IW is in use.

These results are not contradicting however, since the BLIS data studied was selected from the portions of driving when LDW+IW were switched off. Rather, they seem to reflect a clear case of driver adaptation. When LDW+IW is active, drivers use the indicators more to avoid the warning sound that they otherwise get if they change lanes without signalling. When BLIS is active, the questionnaire data confirms that drivers trust the system not to give false negatives (i.e. not warn even though there is a vehicle in the blind spot). Hence the need to use the turn indicator is reduced, because drivers perceive that the really know that there is no other vehicle in the lane they are changing into.

5.5.3 Traffic efficiency

BLIS was not expected to have direct traffic efficiency effects. Furthermore, as described in D6.4, the safety impact results for BLIS did not warrant an up-scaling to the EU-27 level. Hence indirect traffic efficiency effects could not be assessed either.

5.5.4 Environment

BLIS was not expected to have any environmental impacts, and hence this was not tested.

5.5.5 CBA

Carrying out a cost-benefit analysis was considered not being feasible for this function because of performance restrictions resulting from the Field Operational Test (non-applicable and / or insignificant impacts, insufficient knowledge on EU-wide driver behaviour and network characteristics required for up-scaling).

5.5.6 Conclusion

Overall, drivers indicate that BLIS is highly appreciated, despite the fact that it sometimes warns even though there is no vehicle present in the blind spot. Thus, interestingly, drivers were very pleased with a system that gives false positives on a more or less regular basis. This indicates that as long as the core functionality is perceived as useful and the limitations are comprehensible, drivers seem willing to "forgive" the function for its imperfections.

With regard to safety aspects, some indicators pointed toward BLIS having a beneficial safety impact, particularly the relative frequency of lateral incidents. However, as for LDW+IW, it was not possible to review a sufficient number of such incidents to test whether that trend also was statistically significant given the time and resources available for analysis. Consequently, as it was not possible to identify a significant starting point from which a national / EU-27 level safety impact could be calculated, there was no up-scaling or CBA for BLIS. In addition, traffic efficiency and environmental impact was not tested for LDW+IW.

5.6 Curve Speed Warning (CSW)

The Curve Speed Warning (CSW) was not bundled; only one function per vehicle was evaluated. Data from two FORD passenger car manufacturers was obtained during the FOT (see Figure 53). The duration of this experiment is limited to a few days where thirty participants drove with the system. Drivers only filled in Time 1 and Time 4 questionnaires and it is necessary to point out that the time between both questionnaires was limited by the duration of the test.

Ford	MAN	VOLVO	VCC	VW	CRF	IFSTTAR	BMW	DAG
✓								

Figure 53: Manufacturers that gathered CSW data during the FOT

Due to the experiment extent and the number of passenger cars available with CSW systems a hypothesis testing based on the comparison of baseline against treatment phase is not applicable. Consequently the analysis is based on the distributed questionnaires (Time 1 and Time 4).

With the data gathered from the questionnaires, research questions on user acceptance, driver related aspects and cost-benefit could be answered. Statistical analysis was performed in subjective data to test the hypotheses derived on the research questions (see Table 29).

Table 29: CSW hypotheses list for subjective data

	Hypothesis	Performance indicator
Acceptance	Certain features of the functions, in terms of usability, influence acceptance.	Subjective rating
	Certain features of the functions, in terms of usefulness, influence user acceptance?	Subjective rating
	Acceptance changes over time with function use.	Subjective rating
Trust	Trust in function changes over time with function use.	Subjective rating
User practices	User practices (heuristics, rules) will change over time during the FOT	Subjective rating

The main outcomes of the analyses are presented in the next paragraphs. The results of the hypothesis testing can be found in the Annex of deliverable D6.5. For more detailed analyses the reader is referred to the D6.3 for user acceptance and driver related aspects and D6.7 for the cost-benefit analysis.

5.6.1 User acceptance and user related aspects

Data obtained from questionnaire items indicated that scores on satisfaction and usefulness are positive. User acceptance was evaluated according Van der Laan scale (items were evaluated in a scale from -2 to +2). As it is observed in Figure 54 the values for satisfaction and useful categories increase from Time 2 to Time 4 (the differences were statistically significant).

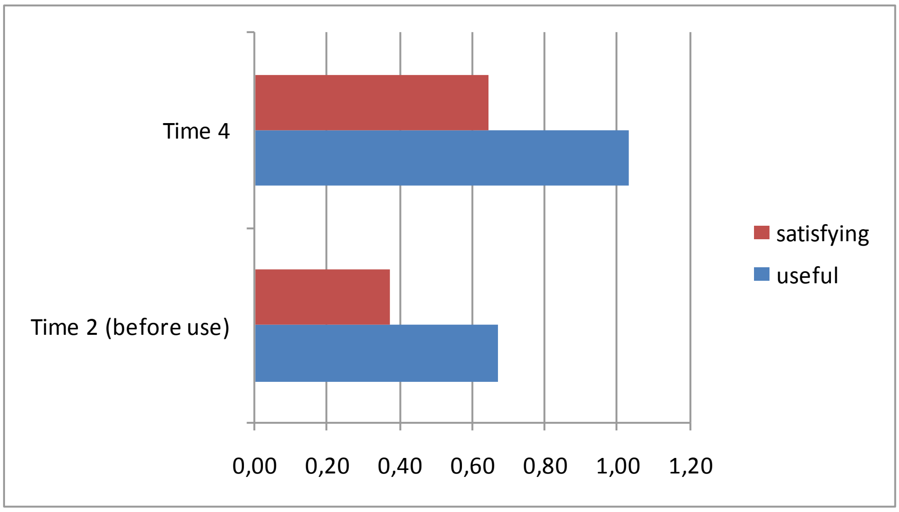

Figure 54: Acceptance rating in terms of usefulness and satisfaction

Considering usage, around a quarter of the drivers who drove with CSW answered they have changed their usage of the system. Some participants stated they used CSW as indicator or for practising a more defensive driving. Moreover, participants trusted more in the system after CSW usage in the FOT (all the trust scores were positive). For trustworthy and reliable scores differences were statistical significant between Time 2 and Time 4. In addition, three quarters of drivers felt that the system increased driving safety. Mainly they felt CSW is most useful in rural road driving (mainly in unfamiliar rural roads and normal traffic conditions).

Further, in general, participants found the visual information was easy to read, easy to understand and grabber their attention.

5.6.2 Conclusion

Altogether, CSW is a system where driver's expectations were fulfilled: usability and acceptance scores for the system are high, three quarters of the sample felt that CSW increased safety, and moreover, it is perceived as most useful on rural roads. Drivers also trust on it and they felt it was easy to understand how it worked. Thus, Curve Speed Warning system is appreciated by drivers and its use could be used frequently to avoid accidents caused by over-speeding in curves.

5.7 Fuel efficiency Advisor (FEA)

The Fuel Efficiency Advisor (FEA) is a system that provides in real time the current location of the vehicle, its fuel consumption, messages, driver times, service intervals and much more to support fuel-efficient driving, or eco-driving. It was tested at the Swedish VMC in 50 trucks that drove more than 3.6 million kilometres equally distributed over baseline and treatment.

It is the only system in the euroFOT project that directly aims to reduce fuel consumption. Since the system only gives advice and does not intervene, the effectiveness is depending on the willingness of the driver to follow the system instructions. Hence, it is difficult to distinguish between the fuel saving potential of the system and the influences that originate from the way of driving of the system user.

Only the environmental effects of the FEA are assessed because only a limited amount of information on the data gathered during the FOT is given and no effects on safety were hypothesised.

5.7.1 Environment

Unlike the other functions tested within euroFOT the FEA cannot be deactivated by the driver which implies a usage rate of 100% within the treatment phase. It cannot be evaluated in which periods the driver acts according to the system instructions. Therefore, the evaluation is not focused on the (technical) potential of the system to save fuel, but rather on the overall system design which includes the HMI and its benefit for decreasing fuel consumption. Limited information about the conditions under which the data was collected makes the results of the analysis difficult to interpret. The treatment phase showed a reduction in fuel consumption of 1.89%, but this effect is not significant. See Table 30 for the results.

Table 30: Results for change in fuel consumption for the Fuel Efficiency Advisor

Average fuel consumption (l/100 km)		**P-value**	**Absolute difference (l/100 km)**	**Relative difference (% of baseline)**
Baseline	Treatment			
37.09	36.39	> 0.05	-0.7	-1.89

Due to the very limited information on driving conditions that result in the evaluated fuel reduction and the related uncertainty ($p > 0.05$) scaling up the effect to EU-27 level is not applicable.

5.7.2 Conclusion

The limited information on the driving conditions under which the FOT data for the FEA was collected makes a reliable interpretation of the analysis very difficult. Influences from shifts in conditions (e.g. higher mileage on urban roads in baseline) or weather effects possibly overweigh the effects of the FEA use. The high mileage (more than 3.6 million kilometres) and the equal distribution over the two phases make it reasonable that the distribution over the different road types is equally spread too. However, the limited information on the data (no factors such as load could be included in the statistical analysis) and the fact that the trucks were used by different drivers in baseline and treatment show that further interpretation of the data is not possible with the applied experimental set up. Therefore, more research is needed in order to fully assess the potential benefits of FEA.

6 Management

euroFOT is amongst the first projects to realize a European scale Field Operational Test. The task of setting up and coordinating workflow and communication processes was an important and challenging effort. In the first project phase, euroFOT management established procedures for efficiently bringing together partners from different organisations, namely OEMs, suppliers, research institutes, and others to work together and openly share information and resources across company borders.

The cooperation across company and organisation borders was agreed upon and fixed in the Consortium Agreement (CA). During the initial project phase, euroFOT IP management team negotiated and completed the CA. With the project start, communication channels and documentation procedures were defined. This involved for example the provision and set up of a document sharing platform and several mailing lists.

For dissemination activities, euroFOT IP management team established communication rules and standards. For public communication a Corporate Identity (CI) guide was put together defining a sound and uniquely recognisable appearance. The branding including the use of colour scheme, logo and general appearance was defined. Various templates were created and distributed in the consortium to ease work and ensure conformity with the CI.

The euroFOT website was set up to increase visibility and to provide general as well as background information on the project itself as well as its progress. The website was constantly updated during the whole runtime of the euroFOT project.

The cooperation amongst partners led to a thorough test plan confirmed and agreed by external reviewers which was executed in the fourth quarter. The good collaboration and communication between the partners during the whole project made the successful conduction possible. The collaboration and the project progress as well as the consumption of resources were continuously tracked. Making use of an electronic tool allowing web-based data entry, the partner efforts and the time needed for reporting could be reduced. This way, the progress was tracked and all relevant data was communicated to the respective stakeholders.

During the course of euroFOT, management had to react to different challenges. The planning and schedule of euroFOT had to be revised several times due to changes in the consortium, technical problems, procedural problems and the economical situation. All these factors were not foreseeable for euroFOT.

More specifically, during the course of the project the following major problems surfaced. For each problem a satisfactory corrective actions or fall-back could be found.

The work performed in the first periode of euroFOT had shown the need of a comprehensive and detailed piloting phase as also stated by the EC reviewers. Thus, a task force had been established updating the current work plan and elaborating a new schedule. Due to the need of a very detailed and comprehensive pilot phase the SP2 needed more time and special planning for each VMC. The additional activities for the piloting phase made it necessary to also negotiate transfer of resources from other WPs to WP2000. The coordination amongst the SPs needed to be intensified.

Difficulties arose as well from practical implementation of the large body of preparation work done by FESTA, for instance the collection of an immense number of hypotheses which given the timeframe of euroFOT never all could be answered in full detail. As a consequence, the hypotheses needed to be prioritized and even then, the collection of measures proved to be extremely complex.

Special technical difficulties not foreseen before project start arose from the data loggers and the incorporation onto the vehicles itself. As for the field test, VMCs could not rely on easily manageable special vehicles driven by experienced test drivers. Instead, the cars came from common stock driven by customers requiring a much higher level of non-intrusiveness and reliability against malfunctions due to installation issues. Also, the installation process of additional technical components was complicated because the electronic body of the vehicle which had to be accessed was intricately hidden. Here, also solutions had to be discussed, found and followed for every OEM and suppliers alike.

Economic crisis in Europe struck the customers showing significant decrease in purchases of advanced assistance systems in cars making it difficult for euroFOT partners to find subjects enabled and willing to take part in the FOT. Partners struggled hard to reach planned numbers of vehicles and to find ways of inviting new participants. Partners consumed higher resources even from their internal reserves to continue efforts. Incentive schemes were elaborated and efforts taken to include company cars.

The economic crisis also resulted in lower sales volumes of upper middle class cars carrying advanced driver systems and also hindered the desired increase of awareness of advanced driver systems, one of the projects goals. This problem was closely monitored. Dissemination efforts were adapted and clear results delivered at the end of the project supporting the systems benefits.

The above challenges and solutions were not only discussed and executed in the consortium but were also documented and negotiated with the EC in amendments. During the course of euroFOT, four amendments were prepared. They reflect the countermeasures already described but also major issues in the consortium. Amendment 1 was negotiated due to changes in the consortium. “GIE Recherches et Etudes PSA Peugeot Citroen – Renault” left the consortium and was substituted by “Institut national de recherché sur les transports et leur securite”. As a consequence thereof, the resources had to be modified. In Amendment 2, again a change in consortium was indicated. “Irion Management Consulting GmbH” left the consortium and “Harman Becker Automotive Systems GmbH” entered the consortium, again with the consequence of a redistribution of work and resources.

Amendment 3 was necessary due to a change in the cost model of partner “MAN Nutzfahrzeuge AG” and the termination of partner “Delphi Delco Electronics Europe GmbH” replaced by “Delphi Deutschland GmbH”. “Institut national de recherché sur les transports et leur securite” was also terminated and replaced by “Institut Francais des Sciences et Technologies des Transports, de l’Amenagement et des Réseaux”. The remaining resources were reorganized and assigned to the new partners.

In Amendment 4 the reaction on the delays and needed countermeasures described above were the main concern. It was negotiated with partners from SP6 to postpone data evaluation. The project extension until month 50 and major redistributions of resources together with a reorganisation and assignment of work packages demanded a great effort both from management and partners.

Regarding contractual aspects, a relevant activity was an amendment to the Consortium Agreement regarding aspects of data sharing and the possibilities of how to store data, exchange data and make data available to partners.

6.1 Management structure

To ensure the success of the IP, euroFOT uses a management structure comprising the technical management on WP level carried out by the work package leaders, the overall technical and operational project management carried out by the Management Team, as well as the coordination with the project sponsors and strategic project lead carried out by the Steering Committee and the coordinator (as shown in the figure below).

In addition to the three management levels, representatives from all participating partners form the General Assembly. This body has the right to propose actions and decisions to the project's Coordinator and Steering Committee.

The project intern management structure is completed by the external partners, namely the European Commission and the EC reviewers.

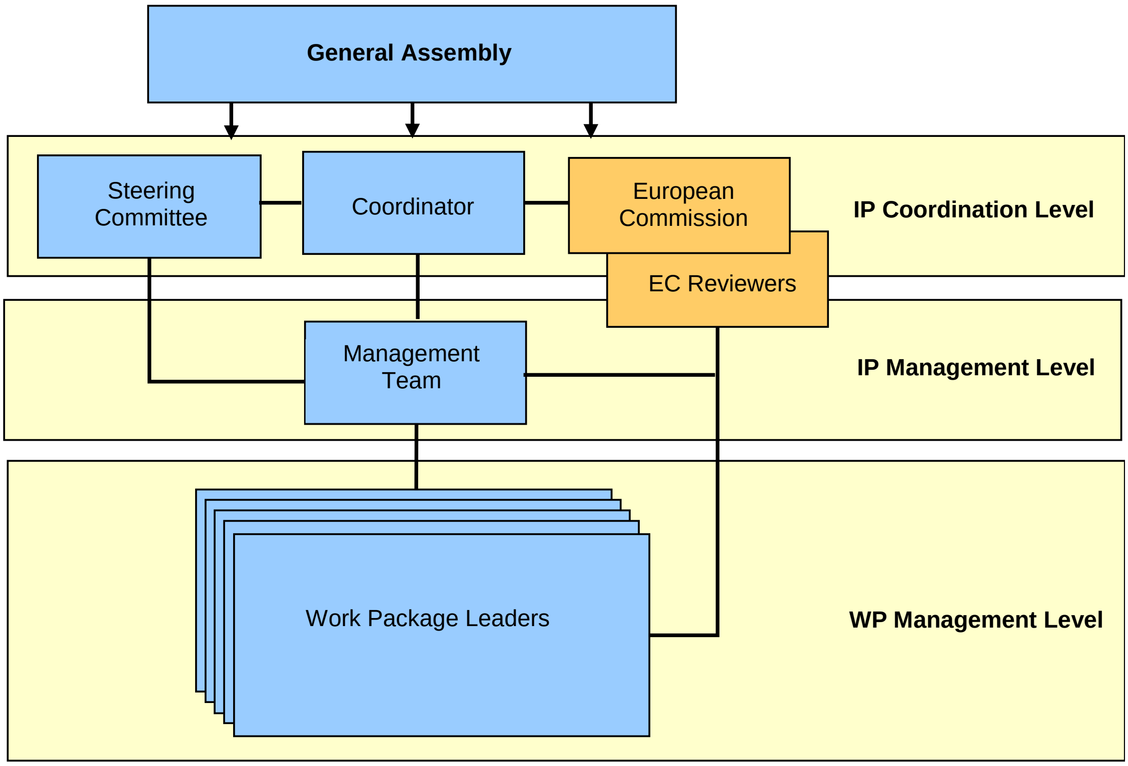

Figure 55: Integrated project management structure of euroFOT

The EC and the project mainly interact through the Coordinator. The Coordinator informs the EC about the project progress through reports, which were submitted on a quarterly basis, and through annual progress reports (periodic reports) including cost statements from all partners.

In addition, EC reviewers' interaction comprises the operational management level with regards to organizational issues of the annual review, mainly processing of deliverables and reports and the incorporation of requests and recommendations.

All bodies met on a regular basis to discuss the current status, decide the next steps and initiate any corrective measures.

6.1.1 Sub project structure

The euroFOT project was structured in six sub projects reflecting the different tasks to be carried out during the lifetime of the project (see Figure 56).

WP1000, IP Management & Dissemination, comprised all overall project tasks making sure that high quality project results can be achieved. Tasks included the overall IP coordination, the operational project management, handling of financial issues, reporting and project communication including documentation, strategic planning, and production of public project dissemination.

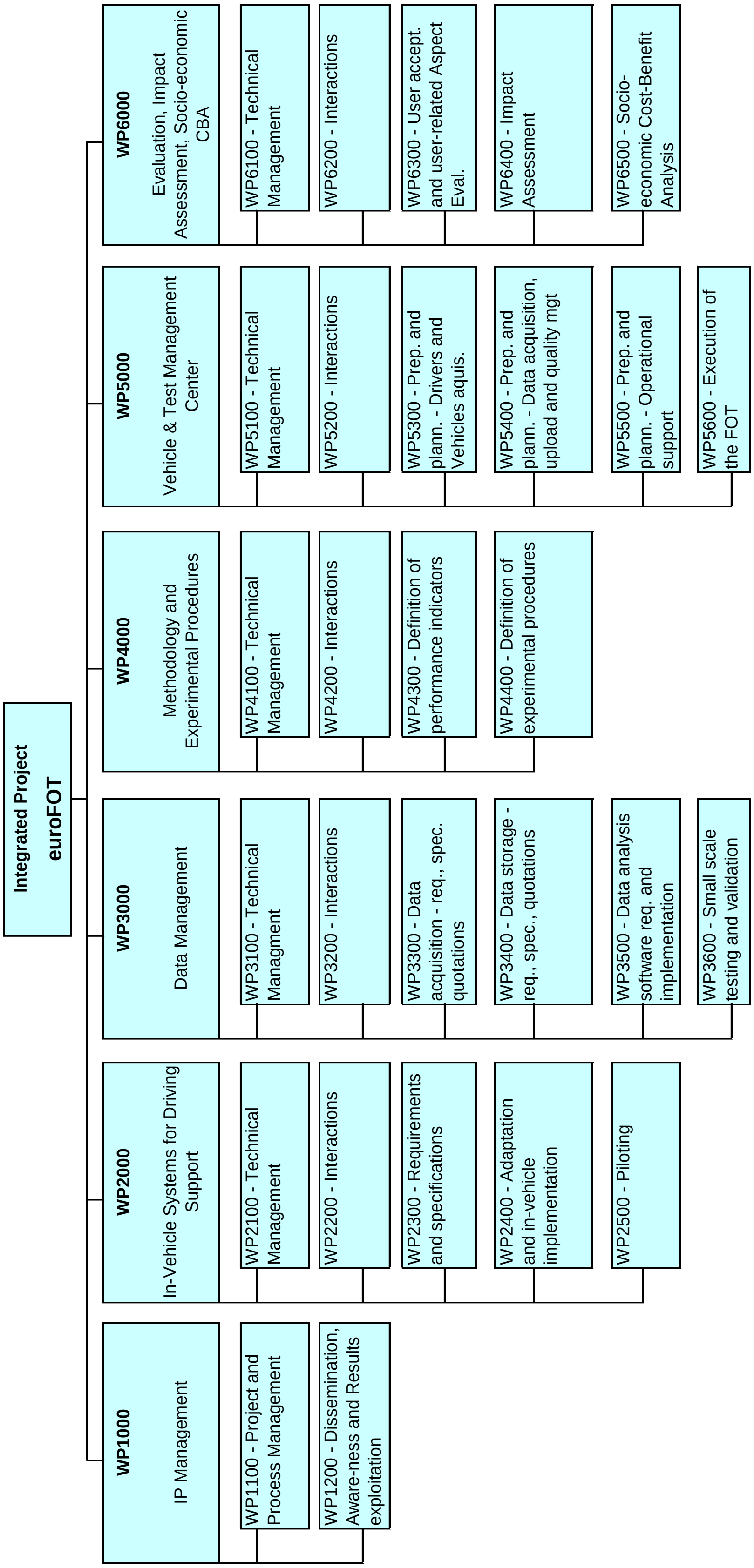

Figure 56: Project structure with subprojects (e.g. SP1 or WP1000) and workpackages (WP).

WP2000, In-Vehicle Systems for Driving Support, defined the objectives for in-vehicle systems and to specify which and how systems were to be tested in real environments.

WP3000, Data Management, specified the data acquisition and data storage systems and requirements. This WP developed analysis tools based on input and evaluation methodology.

WP4000, Methodology and Experimental Procedures, identified and selected performance indicators and specified the experimental procedures applied.

WP5000, Vehicle & Test Management Center, coordinated the vehicle management centres to ensure smooth test operation including provision of equipment, organizing driver interaction and guiding, and ensuring the defined experiment procedures were applied.

WP6000, Evaluation, Impact Assessment, Socio-economic CBA, analysed, evaluated, and studied the data gathered prior in the FOT and identified the impacts.

6.1.2 Partners

The euroFOT consortium was formed by a large number of partners from different organisations such as manufacturers, suppliers, research institutes, and management and consulting firms working together.

Manufacturers:

- Ford Forschungszentrum Aachen
- BMW Forschung und Technik GmbH
- Centro Ricerche Fiat S.C.p.A.
- Daimler AG
- MAN Truck & Bus AG
- Volvo Car Corporation
- Volvo Technology Corporation
- Volkswagen AG

Suppliers:

- Robert Bosch GmbH
- ADC Automotive Distance Control Systems GmbH
- Delphi Deutschland GmbH
- Harman Becker Automotive Systems GmbH

Research Institutes and public organizations:

- Chalmers Tekniska Hoegskola Aktiebolag
- Fundación para la Promoción de la Innovación, Investigación y Desarrollo Tecnológico de la Industria de Automoción de Galicia
- Institute of Communication and Computer Systems
- RWTH Aachen, Institut für Kraftfahrwesen Aachen
- Interdisziplinäres Zentrum für Verkehrswissenschaften an der Universität Würzburg
- University of Leeds
- Nederlandse Organisatie voor toegepast-natuurwetenschappelijk onderzoek TNO
- AZT Automotive GmbH - Allianz Center for Technology
- Bundesanstalt fur Straßenwesen
- Centre Européen d'Etudes de Sécurité et d'Analyse des Risques
- Politecnico di Torino

Management and Consulting Agencies:

- European Road Transport Telematics Implementation Coordination scrl
- ALCOR di Giancarlo Alessandretti
- European Center for Information and Communication Technologies GmbH
- Hagleitner Walter
- Institut Français des Sciences et Technologies des Transports, de l'Amenagement et des Réseaux

6.2 Budget

To achieve the euroFOT objectives, a project volume of over 21 Mio EUR was invested with a requested EU funding of 13.9 Mio EUR.

The distribution of resources to cost categories is shown in Figure 57. The vast majority of the budget is spent on personnel cost (about 16.8 Mio EUR corresponding to 80% of the total budget). For Computing and Travel only 2% of the budget had to be spent so that this position was no significant cost factor. The same holds for Consumables (1% of budget). For Subcontracting 4% of the budget was spent, showing the commitment of the partners for the project.

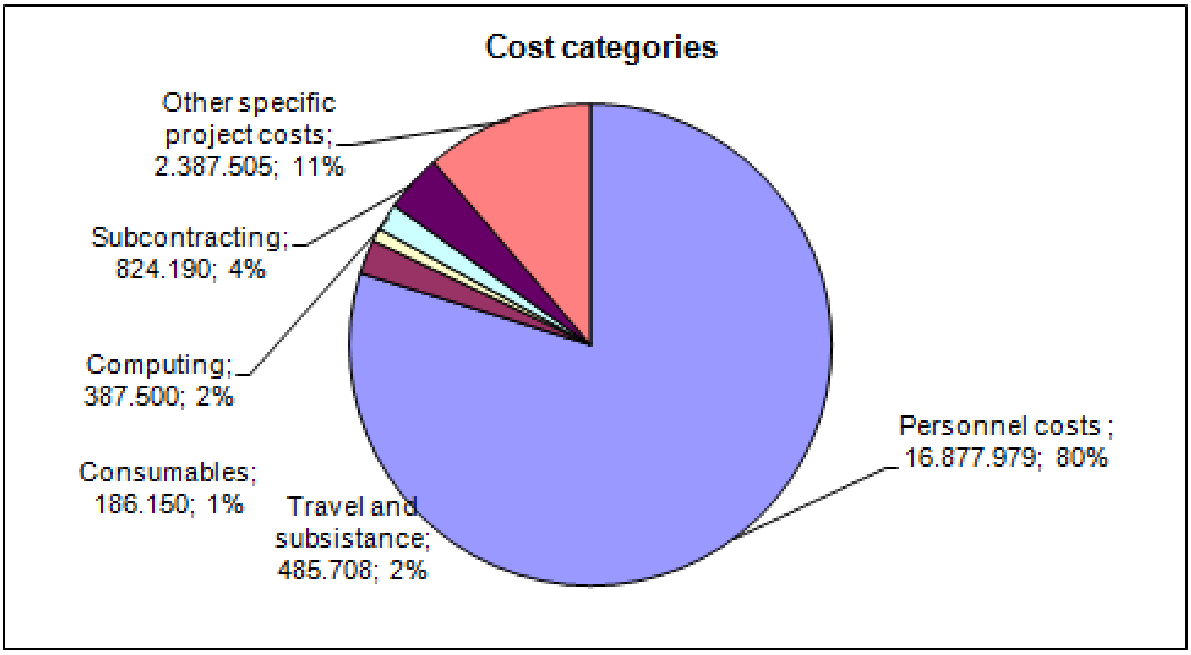

Figure 57: Distribution of costs to cost categories from the Description of Work.

A significant cost category, making up 11% of the total budget, was the item “Other Specific Project Costs”. The major cost factors in this category were equipment needed for data logging and incentives for the test drivers.

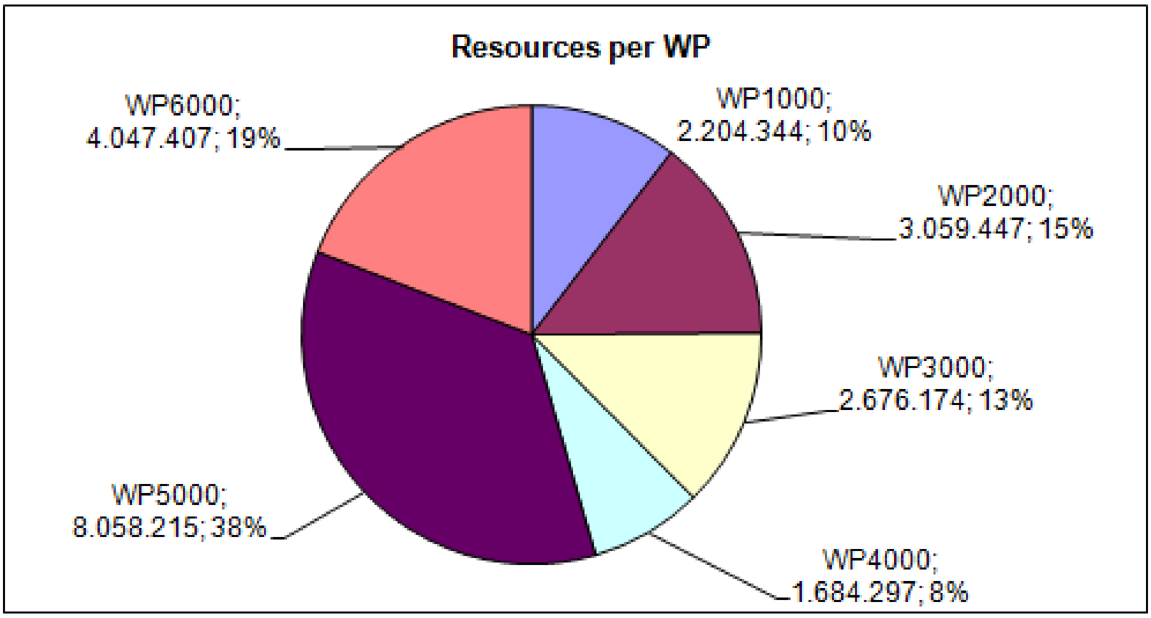

Figure 58: Resources per main workpackage (eg. WP1000 or SP1) from the Description of Work

The budget assigned to the different work packages was as follows:

- WP2000 defining the test requirements and providing vehicles was assigned 15% of the resources.
- WP 3000 that performed data management and consumed 13% of the resources.
- WP4000 elaborated the FOT methodology and required about 8% of the resources.
- WP5000 provided and coordinated the vehicle management centers, provided logging, sensing, and the operational platform and therefore consumed the most resources with 38% of the total budget.
- WP6000 summarized the results of the FOT by collecting and evaluating data, concluded impacts, and performed cost benefit analysis using 19% of the resources.
- The IP management in WP1000 claimed about 10% of the resources, though project and process management activities only required 6.1% of the overall budget. About 3% of the resources were set aside for awareness measures in WP1200 showing the importance of the dissemination part being within euroFOT.

The distribution of funds clearly reflects the focus on the respective parts of the project. IP Management made sure that the work plan was followed as closely as possible, and implemented any changes necessary to reach the project goals.

6.2.1 Resource consumption

The overall resource consumption in euroFOT is expected to be around 2000 person months (PM) at project's end, considerably more than the planned 1340 PM. These resources have been delivered without extending the project budget, either by shifting between cost categories or as contribution from the partner to the project. This is significant, considering the fact that the project runtime was extended, running 10 months longer with no additional funding.

6.3 Legal aspects

Many legal topics are part of project management and are not different from other projects. Two actions of an FOT, however, concern legal issues, namely recording and analysing of data. The issues are privacy, intellectual property and benchmarking.

6.3.1 Privacy

The monitoring equipment of several VMCs continually records the face of the driver. The routes of hundreds of drivers are recorded via GPS-tracks. Recordings are stored and then analysed, they represent personal data, and are subject to European Directives (e.g. D95/46/EC {ref]) as well as national laws. There are ways to anonymize the data – but they are very costly and/or reduce the information content of the data. To ensure privacy of personal data the following conditions are met in the project:

- Driver identification is realized with a number which identifies the subject with a tag (e.g. V_2_5_M_041).
- The name of a participant is kept separate from both the questionnaires and collected raw data.
- Access to the data server is password protected.
- Use of data is restricted to research purposes.
- Publication of videos needs the consent of the participant.

6.3.2 Intellectual property & Benchmarking

The manufacturers of the functions analyzed in an FOT need to protect their knowledge. The analysis of thousands of trips can reveal parameterisations and specifications of sensors. The data driven comparison of functions from different manufacturers (benchmarking) has been excluded in the description of work (DOW)

6.3.3 Access to data

The considerations on the use of data have already been manifested in the DoW by limiting access to "aggregated data" for one year after end of project. Aggregation in the context of the DoW means that data is treated in a way that allows fundamental research without violating privacy and IPR, and does not allow benchmarking. The process of aggregation, however, requires the development of new algorithms to remove personal information and brand related information while still retaining a sound content of information. In short, it implies a significant effort with a decrease in value.

A task force with project stakeholders (OEMs, suppliers, researchers, and legal counsel) has studied the feasibility and efforts needed to find a suitable process for data aggregation. A breakdown of efforts to balance four goals (ensure privacy, respect IPR, avoid benchmarking, maintain reasonable information content) showed, that the funds needed would not be justifiable to, at the foremost, artificially reduce the information content and to create a second body of minimal interest. The task force adopted an altogether different procedure and included this in a 3rd amendment to the Consortium Agreement that has been signed by all project partners.

Instead of providing aggregated data on a public website for download for one year the granting of access is now based on a research proposal that will be reviewed by owners of the data. The accessible content, after successful review of the proposal, represents the almost full body of data that has been analysed to derive the results obtained in euroFOT. An exception has been made not to provide personal data, such as video and GPS position, as this would require the acceptance of all FOT participants. The site of access is at the site of the owner of the database. Derived results from database search can be extracted after their content has been reviewed and checked against the research proposal. All costs for the access - like training session, operator, and consumables - will be agreed upon before the analysis is started.

7 Conclusions

FOTs can definitely contribute to the evaluation of Intelligent Vehicles, but are not the only solution for all investigations concerning new automotive systems. The methodologies that are available need to be adapted to the specific systems, also taking into account existing constraints in time and resources.

Conditions that are necessary for a successful implementation of FOTs include a large variety of aspects:

- Industrial
- Technical
- Organisational
- Market
- Evaluation

All these steps have been addressed within euroFOT, and the results obtained so far are therefore a comprehensive guideline to support the design and deployment of a FOT.

The **industrial factor** is linked in particular to the interest of the manufacturers and their commitment to investigate a specific application with potential customers. This was always the case within euroFOT, and European industries have demonstrated a continuous and effective engagement for the deployment of advanced features to satisfy the population of drivers and passengers.

In general, it can be said that the **technical aspect** is not the limiting factor anymore. This is also confirmed by the good performance of the data acquisition equipment (including sensors) and the procedures for data management in all the five operational sites within euroFOT. Future developments on the technical standpoint will aim to a further simplification of the installation of the instruments, together with a reduction of cost and invasiveness. These will be key points for implementing future FOTs on large fleets.

The **organisational factor** appears to be the most important issue: major difficulties have been found especially in the recruitment of subjects. As identified during the project, special attention should be given to the selection of drivers, the type of incentives, and the vehicle fleets. The policy of incentives was a key factor of success for instance in the German 1 VMC. The involvement of marketing departments was very important for all the OEMs. Various approaches for the recruitment have been used, and again the choice of the most suitable concept strongly depends on the local conditions. Unfortunately euroFOT was run in a period of economic downturn, and this brought additional difficulties due to the low levels of sales and the preference of customers for the more economic vehicle models, without additional features such as ADAS. Regarding the management of the experiments, euroFOT has faced a number of challenges, due to the limited previous experience in large scale tests, but in general the implemented solutions have proven to run smoothly and effectively. Guidelines for future experiments are now available, with several detailed lesson learned for all the major topics.

Market and communication strategies have been used throughout the project for assuring the involvement of interested stakeholders. In particular, the workshops with subject drivers allowed feedback on several aspects, although it was not always easy to have a good participation.

A final very important aspect regards the methodologies for **evaluation**. euroFOT proved that very specific tools and procedures for data processing are a fundamental enabler for this type of experiment, and they are not fully available at the present stage. During the project, significant progress has been made in areas such as the enrichment of data, the extraction of events, and the automated computation of performance indicators. However, after such a diversified experience, partners of euroFOT believe that additional focused research efforts are certainly required. For example, some of the automated procedures should be more extensively validated or developed, and the criteria for using video data (which are rich in information but time consuming in the analysis phase) should be clarified.

7.1.1 Contribution toward ADAS development

The analysis of ADAS-application, shown in questionnaire and usage data from this FOT, indicates that the functions under test have been applied in everyday traffic, and in the way that was intended. Furthermore, no crucial differences were found between the use habits of ADAS in the six nations represented in euroFOT. If this is translated into customer satisfaction then it is a clear message that the European market is suitable for a wider deployment of ADAS. The findings compiled in euroFOT deliverables, and the on-going use of the data, will provide a contribution to the coming roadmaps for automotive research, system developers and vehicle manufacturers.

The first deployment of a European adaptive cruise control (Distronic) in the Mercedes S-class in 1998 and the introduction of ACC in the 2011 Ford Focus indicate the democratization of ADAS. For several years now ADAS systems have started to penetrate lower value vehicles. Technical results from euroFOT help to tackle improvement of existing ADAS, better calculations and algorithms, design of navigation systems and improvement of product specifications.

Among the objectives of euroFOT has been to show the feasibility of a European approach to Field Operational Tests. The experiences gained and the wealth of results from analysis show that this objective has been achieved. Another objective, namely to provide better information for a wider deployment of ADAS, has been reached with contributions to papers, conference presentations and the Final Event at Autoworld in Brussels. It will, hopefully, be confirmed with an increase in take-.rates of ADAS functions in new vehicles.

The common methodology, albeit not always followed to the point for technical and cultural reasons, has established a valuable practical realisation of the FESTA structure and is a building block for evaluating new ADAS functions. The evaluation with real data and a good experimental design comes with a budget and time need, which is one of the disadvantages of FOTs. The Lessons learned in euroFOT will lead to more economical experiments which still allow sampling the breadth of everyday driving experience.

Acknowledgements

euroFOT has been co-funded by the European Commission in the 7th Framework Programme (FP7) ICT-2-6.2 ICT for cooperative systems, grant agreement number 223945. The consortium of 29 partners has contributed many resources not foreseen in the original planning and their contribution towards a successful completion is greatly acknowledged.

References

This listing comprises all deliverables written in the project euroFOT (e.g. D11.1). Their access status is indicated in brackets, e.g. (PU) (PU = Public, PP = Restricted to other programme participants (including the Commission Services, RE = Restricted to a group specified by the consortium (including the Commission Services), CO = Confidential)

[1] VDI/VDE-IT and ITE University of Cologne. 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. ww.esafetysupport.org/en/esafety_activities/related_studies_and_reports, 2005

[2] Lu, M., Modelling the effects of road traffic safety measures, Accident Analysis and Prevention 38, p. 507–517, 2006

[3] Clarke, D., Ward, P., Bartle, C., Truman, W. Killer crashes: Fatal road traffic accidents in the UK, Accident Analysis & Prevention 42, p.764-770, 2010

[4] FESTA Handbook v2 (Field opErational teSt supporT Action), FP7 ICT, http://www.its.leeds.ac.uk/festa/, 2008

[5] FESTA Handbook rev. 4, http://www.fot-net.eu/, http://www.its.leeds.ac.uk/festa/, 2011

[6] FESTA Experimental Procedures Deliverable 2.3, 2008

[7] Annex I - “Description of Work” to Grant agreement no. 223945, V4.0, Dec. 2011

[8] D11.1 IP process handbook (PP), Etemad A., Kessler C., Amditis A., Bimpas M., Alessandretti G., Kessel T., Inhülsen H., Irion J., Oct. 2011

[9] D11.3 Final Report (PU), Kessler C., Etemad A., Alessandretti G., Heinig K., Selpi, Brouwer R., Czerpinszky A., Hagleitner W., Benmimoun M., Dec 2012

[10] D12.1 Communication Plan (PP), Brusselmans A., Flament M., Nov. 2010

[11] D12.2 Project website (PU), http://www.eurofot-ip.eu, 2012

[12] D12.5 Plan for use of results (PP), Csepinszky A., Etemad A., Kessler C., Alessandretti G., Kessel T., Inhülsen H., Klessen C., Benmimoun M., Brouwer R., Jamson S., Tamang P., Krökel D., Burzio G., Tadei R., Obojski M.-A., Bärgman J., Selpi, euroFOT consortium, June 2012

[13] D12.6 euroFOT Final Event Plan (PP), Balistreri A., Brusselmans A., Csepinszky A., Flament M., Silva I., Fernandez L., Wagener M., Etemad A., Apr, 2012

[14] Report on Exchanges with U.S. Organisations May 10-14, 201, Alessandretti G., Benmimoun A., Brouwer R., Aust M.L., SaintPierre G., van Noort M., May 2010

[15] D2.1 Specifications & Requirements for Testing In-vehicle Systems for Driving Support (PP), Dozza M., Heinig K., Kessler C., Petersson M., Rakic M., Tattegrain H., Sep. 2010,

[16] D2.2 Report on the results of adaptation, in-vehicle implementations and piloting (PU), Saint Pierre G., Kessler C., Malta L., Junr 2011

[17] D3.1 Selection Procedure for Data Acquisition, Sensors and Storage for euroFOT (PP), Selpi, Kanstrup L., Val C., Benmimoun M., Borgen S., Jul. 2009

[18] D3.2 Base Functionality Data Analysis Software (PP), Kovaceva J., Cano J. A. L., Sánchez D., Dozza M., Selpi, Val C., Küfen J., Pagle K., Metz B., June 2011

[19] D3.3 Data management in euroFOT (PU), Selpi, Borgen S., Bärgman J., Svanberg E., Dozza M., Nisslert R., Norell C., Kovaceva J., Sanchez D., Saez M., Val C., Küfen J., Benmimoun M., Metz B., Aug. 2011

[20] D4.1 Performance Indicators, Events and Situational Variables within euroFOT (PP), Kircher K., Heinig I., Brouwer R., WP4300 partners, Oct. 2009

[21] D4.2 Experimental Procedures (PP), Jamson S., Chorlton K., Gelau C., Schindhelm R., Johansson E., Karlsson A.S., Metz B., Tadei R., Benmimoun M., Val C., Regan M., Wilschut E., Brouwer R., Oct. 2009

[22] D5.1, Description of the VMCs and common guidelines (PP), Flament M., Cavallo F., Csepinszky A., Hagleitner W., Oct. 2009

[23] D5.2, List of drivers, vehicles, equipment, Legal aspects and report on installation of the logging devices (PP), Csepinszky A., Hagleitner W., June 2011

[24] D5.3 Final delivery of data and answers to questionnaires (PP), Schoch S., Guidotti L., Csepinszky A., Metz B., Tadei R., Tesauri F., Burzio G., Schwertberger W., Val C., Krishnakumar R., Selpi, Johansson E., Gustafsson D., Obojski M.A., Grundler W., Kadlubek S., Kessler C., Hagleitner W., Oct. 2011

[25] D5.4 + D4.3, Test and vehicle report, summary of user workshops + Methodological Lessons Learned (PU), Csepinszky A., Kessler C., Hagleitner W., Schwertberger W., Obojski M.A., Just M., Kuhn F., Metz B., Guidotti L., Gustafsson D., Heinig K., Val C., Dec 2012

[26] D6.1 Final Evaluation Results (PU), Benmimoun M., Pütz A., M., Aust J., Faber F., Sanchez D., Metz B., Saint Pierre G., Geißler T., Bühne J., Guidotti L., Malta L., Dec 2012

[27] D6.2 Analysis methods for user related aspects and impact assessment on traffic safety, traffic efficiency and environment (PU), Faber F., Jonkers E., Aust M.L., Benmimoun M., Regan M., Jameson S., Dobberstein J., June 2011

[28] D6.3 Final results: User acceptance and user-related aspects (PU), Sanchez D., Garcia E., Saez M., Benmimoun M., Puetz A., Aust M.L., Gustafsson D., Metz B., Saint Pierre G., Tattegrain H., Guidotti L., Schindhelm R., Heinig I., Malta L., Obojski M.-A., Nov. 2012

[29] D6.4 Final results: Impacts on traffic safety (PU), Malta L., Aust M.J., Faber F., Metz B., Saint-Pierre G., Benmimoun M., Schäfer R., Nov. 2012

[30] D6.5+D6.6 Final results: impacts on traffic efficiency and environment (PU), Faber F., Jonkers E., van Noort M., Benmimoun M., Putz A., Metz B., Saint-Pierre G., Gustafson D., Malta L., Nov. 2012

[31] D6.7 Overall Cost-Benefit Study (PU), Geißler T., Bühne J., Dobberstein J., Gwehenberger J., Knieling M., Nov. 2012

[32] D6.8 FOT Data (PU), Kessler, C., Etemad A., June 2012

[33] Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen et al. (2006). The 100-Car Naturalistic Driving Study, Phase II – Results of the 100-Car Field Experiment.

Cambridge: U.S. Department of Transportation, John A.Volpe National Transportation Systems Center.

[34] Baum, H. et al. (2008): eIMPACT, Socio-economic Impact Assessment of Stand-alone and Co-operative Intelligent Vehicle Safety Systems (IVSS) in Europe, Deliverable D6, Cost-Benefit Analyses for stand-alone and co-operative Intelligent Vehicle Safety Systems, Cologne 2008.

[35] Wilmink, I. et al. (2008): eIMPACT, Socio-economic Impact Assessment of Stand-alone and Co-operative Intelligent Vehicle Safety Systems (IVSS) in Europe, Deliverable D4, Impact assessment of Intelligent Vehicle Safety Systems, Delft 2008.

[36] Grover, C. et al. (2008): Automated Emergency Braking Systems: Technical requirements, costs and benefits, Published project report PPR 227, Berkshire

[37] Benmimoun, M. et al. (2011). Incident detection based on vehicle CAN-data within the large scale field operational test “euroFOT”, 22nd Enhanced Safety of Vehicles Conference (ESV 2011), Washington, DC/USA.

[38] Nilsson G. (1981). The effects of speed limits on traffic accidents in Sweden. Proceedings, International Symposium on the Effects of Speed Limits on Traffic Crashes and Fuel Consumption.

[39] Summala, H. (2007). Towards Understanding Motivational and Emotional Factors in Driver Behaviour: Comfort Through Satisficing. In P. C. Cacciabue (Ed.), Modelling driver behaviour in automotive environments: critical issues in driver interactions with intelligent transport systems (pp. 189-207). London: Springer.

[40] Summala, H. (1996). Accident risk and driver behaviour. Safety Science, 22, 103–117.

[41] Shadish, R., Cook Th., Campbell, D. (2001). Experimental and Quasi-Experimental Designs for Generalized Causal Inference, Wadsworth Publishing.

Annex 1 Glossary

In this Annex the reader will find a selection of words coming from the official euroFOT glossary which is particularly important for the understanding of this deliverable.

The euroFOT glossary started inside euroFOT SP2 and is based on the FESTA glossary. Every time the glossary is updated, the parallel European supporting initiative FOT-NET[6] is notified; the glossary is then updated on the FOT-NET website and other FOT projects such as Tele-FOT are notified of the new available version of this glossary. For this reason we invite the readers of this report to also consult the glossary on http://www.fot-net.eu/.

CAN	Controller Area Network, a protocol for communicating electric signals in a vehicle bus system
MOST	**M**edia **O**riented **S**ystems **T**ransport, a protocol for high speed communication in a vehicle bus syste,
LIN	Local Interconnect Network, a protocol for communicating electric signals in a vehicle bus system
HMI	Human machine interface, which allows the communication between man and machine
ECU	Enigne or Electronic Control Unit. Embedded computer in a vehicle which controls the operation of subsystems.
EBA	Event Based Analysis
ABA	Aggregation Based Analysis
PRM	Physical Risk Modelling
ADAS	Advanced Driver Assistance Systems
CBA	Cost Benefit Analysis
GEE	Generalized Estimated Equation
GLMM	Generalized Linear Mixed Models
FESTA	Field opErational teSt supporT Action. FP7 research programme to collect knowledge on FOTs.

[6] FOT-NET: the FOT-Net project aims to create a networking platform for anyone interested in Field Operational Tests, their set-up and their results. More information on this project can be found at http://www.fot-net.eu/

FOT aka Field Operational Test	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.
Function	Implementation of a set of rules to achieve a specified goal
System	A combination of hardware and software enabling one or more **functions**
Use case	A specific **event** in which a **system** is expected to behave according to a specified **function**
Situation	One specific level or a combination of more specific levels of **situational variables**.
Situational variable	An aspect of the surroundings made up of distinguishable levels. At any point in time at least one of these levels must be valid.
System state	The current setting of a **system**.
Scenario	A **use case** in a specific **situation.**
Research question	General question to be answered by compiling and testing related specific **hypotheses**
Hypothesis	A specific statement linking a cause to an effect and based on a mechanism linking the two. It is applied to one or more **functions** and can be tested with statistical means by analysing specific **performance indicators** in specific **scenarios**. A hypothesis is expected to predict the direction of the expected change.
Baseline period/phase	The part of the data collection during which the **function(s)** operate in "silent mode", that is, they collect data, but do not give any signals to the driver. From the viewpoint of the driver the **function(s)** is/are off.
Treatment period/phase	The part of the data collection during which the **function(s)** are switched on by the experimental leader, such that they are either active all the time, or can be switched on or off by the driver.
Baseline within comparison situation	**Scenario** with **system** under evaluation "turned off"
Treatment within comparison situation	**Scenario** with **system** under evaluation "turned on"
Controlled factors	Are those factors that are kept constant within one **analysis.** The data is filtered such that only occurrences in which the controlled factors assume the intended values are selected.
Variable factors	Are covariates, they are not kept constant within one analysis, but their values are logged and their influence on the results is considered.
Performance indicator	Quantitative or qualitative indicator, derived from one or several **measures,** agreed on beforehand, expressed as a percentage, index, rate or other value, which is monitored at regular or irregular intervals and can be compared to one or more criteria.