PLANNING OF A FIELD OPERATIONAL TEST ON NAVIGATION SYSTEMS: IMPLEMENTATION AND EVALUATION OF PRE-STUDIES PROVIDING BACKGROUND INFORMATION FOR THE DATA ANALYSIS

Barbara Metz[1], Stefanie Schoch[1], Mariana Rakic[2] & Friedeman Kuhn[3]
[1] Center for Traffic Sciences (IZVW) at the University of Wuerzburg, Roentgenring 11, 97070 Wuerzburg, Germany
[2] BMW Group Forschung und Technik, Hanauer Straße, Munich
[3] Daimler AG, Leibnizstrasse, Sindelfingen

ABSTRACT: In preparation of a field operational test (FOT) on navigation systems, which will be conducted as part of the European Project euroFOT, a questionnaire survey and an experiment were carried out. The questionnaire survey assessed the distribution and evaluation of navigation systems amongst drivers together with relevant driving related background information. The gathered information is planned to be used as background information for driver recruitment in the FOT as well as for up scaling the FOT-results for a larger sample. The experiment studies the influence of navigation systems on driving to unfamiliar destinations. The aim of the experiment is to implement and evaluate the data analysis planned for the FOT. The approach chosen for evaluating the data analysis is described for four selected parameters. The contribution of both studies regarding planning and implementing the data analysis for an FOT is discussed.

1 INTRODUCTION

As part of the project euroFOT, a 7th framework European project, a field operational test (FOT) on the influence of navigation systems on driving will be conducted in 2010. In an FOT, the impact of several in-vehicle safety, assistance or information systems on driving is assessed based on data that have been collected in daily, uncontrolled drives by non-professional drivers. For this reason vehicles are equipped with automatic data logging and are used by the drivers participating in the FOT like their private cars during the measurement period.

The aim of the planned FOT referred to in this paper is to study the impact of navigation systems on driving safety and efficiency. Furthermore, it is planned to investigate acceptance and usage of navigation systems in real traffic. The goal is to investigate the effect of the routing function on driving. As a second use case, the dynamic routing function will be studied which dynamically adapts the chosen route based on traffic information via TMC. Each driver will participate in the FOT for 3 months (within-subjects design): Each driver will be driving one month without a navigation system (baseline condition), one month with a mobile navigation system and one month with the built-in navigation systems provided by the participating car manufactures BMW and Daimler [1].For the FOT, both Daimler and BWM will each equip 15 vehicles with data loggers and video recording.

Prior to the FOT, a questionnaire survey has been conducted to gather background information on the usage and evaluation of navigation systems as well as information on background parameters for the recruitment of drivers for the FOT (study 1). Furthermore, as part of the piloting procedures in euroFOT, an experiment took place that compares driving with and without navigation systems on unfamiliar routes (study 2). The data collected in the experiment in real traffic is used to implement and test the data analysis planned for the FOT.

2 STUDY 1: QUESTIONNAIRE SURVEY

2.1 Methods

To gather background information about both the distribution of navigation systems among drivers and the evaluation and usage of the systems, a questionnaire survey was conducted. During short interviews (duration approx. 5-10 minutes) drivers were asked whether they owned a navigation system or not. If a driver had a navigation system in the car, user practices and the evaluation of different system aspects were of special interest. Response types were either Likert-scales or open responses. Questions with regard to system evaluation were for instance 5-point-Likert-scales (1=very useless, 5=very useful) to rate the usefulness of the two main functions of navigation systems (route guidance and dynamic routing). Questions concerning system usage were e.g. how often a navigation system was used on familiar and unfamiliar routes. Here a 7-point likert-scale ranged between 1=nearly never and 7=nearly always. Furthermore, relevant background information, like annual mileage and the frequency of driving on unfamiliar routes were collected.

The survey was conducted in November / December 2008 and in June / July 2009. N=147 drivers (male: n=123, female: n=24) were questioned at a highway service area and n=95 (male: n=50, female: n=45) at a parking lot in front of a supermarket. Mean age of the drivers was 47 years (sd = 16.52).

2.2 Results

Because of the differing sample sizes as well as differences in variance, non-parametric tests are used. Of the interviewed drivers, 56% (n=135) use navigation systems. Having a closer look at drivers who use navigation systems, 63% use a mobile device, whereas 36% have a built-in device in the car; 1% uses both mobile and built-in devices. Drivers who do not have a navigation system indicate that they have been driving less kilometers during the preceding 12 months than drivers with navigation systems. Furthermore, drivers without navigation system drive rather seldom on unfamiliar routes, whereas drivers who own a navigation system drive rather often on unfamiliar routes. Results show that drivers who use navigation systems consider both the route guidance and the dynamic routing function as more useful than drivers who do not use such systems. Drivers who use navigation systems rate the route guidance function as more useful than the dynamic routing function (Wilcoxon test: $Z=4.12$, $p<0.001$). Furthermore, they indicate that they use a navigation system only rarely on familiar ($m=2.5$, $sd=1.99$) but more or less always on unfamiliar ($m=6.1$, $sd=1.33$) routes. The difference between the two use cases is significant (Mann-Whitney-U-test: $Z=9.39$, $p<0.001$).

Table.1. Differences between users and non-users of navigation systems. Results are based on Mann-Whitney-U-tests.

	Without navigation system		With navigation system		Z-value	p
	m	N	m	N		
Mileage last 12 month	16329	103	32134	126	-5.10	< 0.001
% unfamiliar routes	12.59	103	27.12	128	-5.98	< 0.001
Usefulness route guidance	3.51	76	4.53	78	-5.58	< 0.001
Usefulness dynamic routing	3.28	74	3.69	78	-2.15	< 0.05

Drivers who own a mobile device drove less kilometres in the last 12 months (Mann-Whitney-U-test: $Z=-3.77$, $p<0.001$) and their percentage of trips on unfamiliar routes is lower than for owners of built-in navigation systems (Mann-Whitney-U-test: $Z=-2.27$, $p<0.05$). The two groups do not differ in their evaluation of the two functions.

2.3 Discussion

Not very surprisingly owners and non-owners of navigation systems differ in relevant driving-related parameters. One aim of euroFOT is to assess the impact of the systems under investigation on an EU-level [2]. One possibility to reach this aim is to gather driving data for a representative sample of drivers in the FOT. Regarding the evaluation of navigation systems this would indicate a sample which contains high as well as low mileage drivers and drivers with a high as well as a low proportion of unfamiliar trips.

At the same time, the characteristics of navigation systems have to be kept in mind. Navigation systems can be expected to have the highest impact while driving on unfamiliar routes. The data from the questionnaire survey show that this is also the use case considered most important by system users. The dynamic routing function could have an impact on driving efficiency also on familiar routes. But, as results show, this function is considered less useful and consequently, navigation systems are only rarely used on familiar routes. Thus, data collected during familiar trips like regular daily drives (e.g. to and from work) are very unlikely to contribute any substantial information for assessing the impact of navigation systems. To optimize data collection it is therefore recommended from an experimental point of view to maximize the likelihood of unfamiliar drives in the period of the data collection. Based on the results of the questionnaire survey it seems impossible to recruit a sample for the FOT that is both representative and maximizes the probability of collecting data that are relevant for the evaluation of navigation systems.

One possible strategy to combine both approaches is to recruit a sample that maximizes the likelihood of drives that are relevant for the evaluation of navigation systems. The data from the more representative questionnaire survey can then be used to upscale the results obtained in the FOT with an unrepresentative sample of high-mileage drivers. As can be seen in the results of the questionnaire study, it is not possible to upscale linearly the results obtained during unfamiliar drives of high mileage drivers. Instead the proportion of unfamiliar drives which rises together with annual mileage has to be considered. Furthermore, estimating the effect for different penetration rates and the possible benefit of higher penetrations, the relation between the likelihood of unfamiliar drives and usage of navigation systems should be kept in mind. The questionnaire survey reveals that at least in Germany drivers who regularly drive on unfamiliar routes already use and own navigation systems. At the same time, the possible benefit of navigation systems for current non-users is limited because they only rarely experience the main use-case of navigation systems (that is an unfamiliar trip).

3 STUDY 2: EXPERIMENTAL PILOT TEST

3.1 Aim

Prior to the start of the FOT, pilot tests are conducted in euroFOT. The aim of the pilots is to test the data logging and the questionnaires as well as to implement the database and the data analysis routines. To be able to evaluate the planned data analysis, the pilot study for evaluating navigation systems is designed as an experiment that compares driving with and without navigation systems. Furthermore, the drives are protocolled in detail by the experimenter to provide information on the actual occurrence of events and situations relevant for the analysis.

3.2 Methods

The route chosen for the experiment consisted of ten destinations that were combined into a circuit track from Wuerzburg to Schweinfurt and back to Wuerzburg. Since the destinations were supposed to be unfamiliar for the drivers, mainly small side roads in residential areas of Schweinfurt or destinations in small villages around Schweinfurt were selected. Drivers who came from the area around Schweinfurt were not allowed to participate. The complete route consisted of urban, rural and highway sections. The drivers did not know where they had to drive to beforehand but were told by the experimenter who was going with them during the experiment. The Daimler prototype vehicle for euroFOT was used for the experiment. It was equipped with the data logging technique to be used in the FOT. The data logging consisted of a CAN-data logger and of four-channel video data. The video includes a forward and a backward view from the vehicle, a view of the driver’s face and view showing the handling of the on-board information system COMMAND. Data logging started and stopped automatically with starting and stopping of the engine. For the FOT, an incident button has been installed which is supposed to be used by the drivers to record audio comments whenever something remarkable - e.g. an incident - happens. During the pilot study the incident button was pressed whenever one of the ten destinations was reached. Thus the reaching of a destination was marked in the data and could be identified easily in the data analysis.

N=10 subjects (2 female) participated in the pilot study. Mean age of the drivers was 32 years (sd=10.7, min=23, max=59) and mean annual mileage was 15050 km (sd=13643). The study had a between-subjects design: Five drivers were allowed to use the built-in navigation system to find the destinations, five drivers had to find their way with the help of maps or instructions of internet route planners that had been printed out for the drivers. At the beginning, after the fifth destination and at the end of the experiment the participants filled in questionnaires that were based on the questionnaires developed in euroFOT. During the drives the experimenter was protocolling driving related events relevant for the evaluation of navigation systems in the FOT. Such events are for instance turning manoeuvres or stops. In the experiment the main task of the drivers was to drive to defined but unfamiliar destinations. Due to that stops and turning manoeuvres were in general related to the way finding task (e.g. stopping to read the map, turning because of wrong direction). The protocols were used to evaluate the data processing routines implemented to automatically mark specific events based on CAN-data. The recorded CAN-data were used to implement the database and to develop routines for data quality check and data pre-processing needed for the FOT. Based on the data, some of the hypotheses developed for euroFOT were tested and the feasibility of the indicators was evaluated. All drivers except one started the experimental drive between 08:30 and 10:30 in the morning. Therefore, for nine drivers there was daylight throughout the whole drive; during the last drive (condition with navigation system) it became dark. During one drive (condition without navigation system) it rained most of the time, during the other drives if any, only little rain occurred.

3.3 Results

Due to the limited scope of the paper, the results for four selected indicators will be described: relative travel time, relative travel distance, number of turning manoeuvres per destination and number of stops per destination. All indictors relate to the hypothesis that due to navigation systems way finding will be more efficient.

Relative travel time and relative travel distance evaluate the route finding efficiency for whole trips. For both parameters, reference values were derived from start and end GPS-coordinates with a reference route planning system. For the experimental pilot, the internet based route planner by Falk (http://www.falk.de/) was used. The estimated reference was subtracted from the measured value. The parameter was calculated for travel time and travel distance separately for each destination and then aggregated per driver. The aggregation across different destinations per driver mimics the situation that in the FOT, the analysis will be based on different unrelated drives. As can be seen in figure 1 left, there is a slight increase of relative travel distance which is not significant. Relative travel time (see figure 1 right), however, increases significantly when driving without navigation system (Man-Whitney-U-test: $Z=2.61$, $p<0.01$).

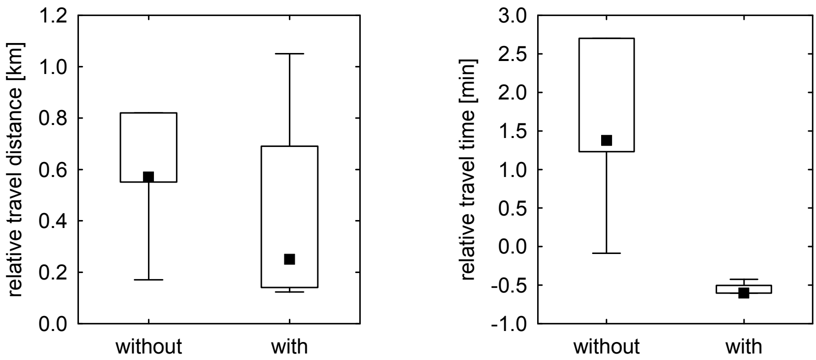

Fig.1. Influence of navigation system on the relative travel distance (left) and the relative travel time (right).

For the two other indicators, it is searched for specific events in the objective driving data which are indicative for poor way finding performance. Such events are for instance turning manoeuvres and stops that are unrelated to the current traffic situation. Based on the protocols, more turning manoeuvres occurred in the condition without navigation system (Man-Whitney-U-test: $Z=1.98$ $p<0.05$).

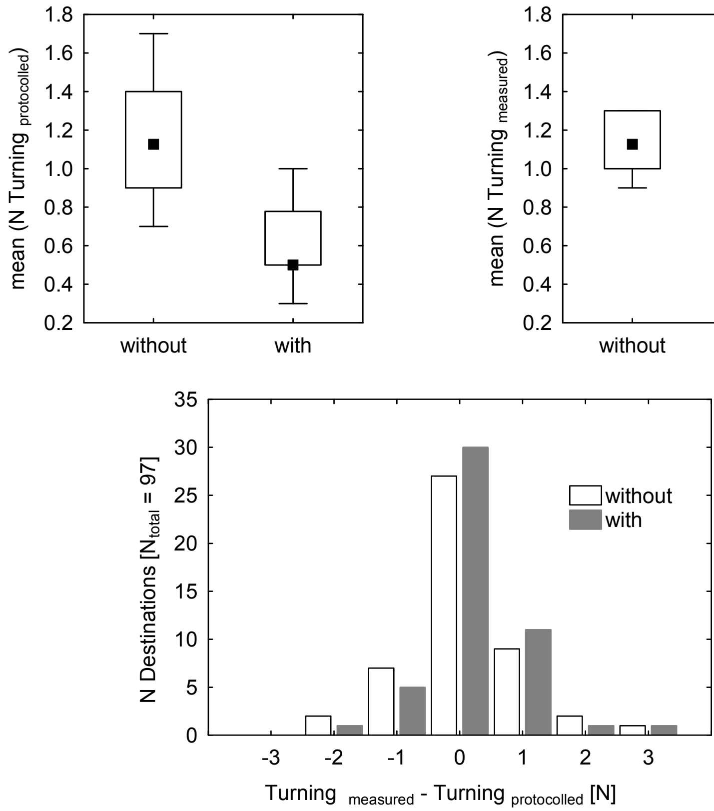

Fig.2. Influence of navigation system on measured (upper left) and protocolled (upper right) turning manoeuvres. The lower graph shows the absolute difference between measured and protocolled turning manoeuvres per destination reached.

In the CAN-data, a turning manoeuvre is coded whenever a high steering wheel angle (>90°) and reversing occur together within one minute. As can bee seen in figure 2, upper right, more turning manoeuvres are detected in the condition with navigation system compared to without. But, due to error variance, the difference between the conditions is no longer significant (Mann-Whitney-U-test: Z=1.46 p=0.144). Figure 2, lower graph, shows the absolute difference between measured and protocolled turning manoeuvres per destination. This analysis helps to evaluate the event detection algorithm by showing the difference between the number of events detected by an experimenter and the number of events detected by the automatic event coding algorithm. In the analysis, drives to N=97 destinations are considered. For 59% of the destinations the number of protocolled and measured turning manoeuvres match. In further 33%, the deviation is one turning manoeuvre. In 15% of the destinations the frequency of turning is underestimated; in 26% it is overestimated. In a next step, the event detection algorithm is evaluated by viewing the corresponding video data. The analysis of the video data shows that of the detected events 62% relate to real turning manoeuvres; most of the false positives (33%) are related to parking manoeuvres. The analysis of the video data reveals another problem of the used algorithm. The time window of 60 seconds is too long and leads to events that are much longer than the actual turning manoeuvre. In 71% of the events, the coded time period includes turning at the next intersection. The missed turning manoeuvres are probably manoeuvres that do not contain reversing. For instance, it is possible to turn around at an intersection without ever reversing. It is not possible to differentiate that kind of turn around based only on vehicle data (like steering angle) from turning right or left at intersections.

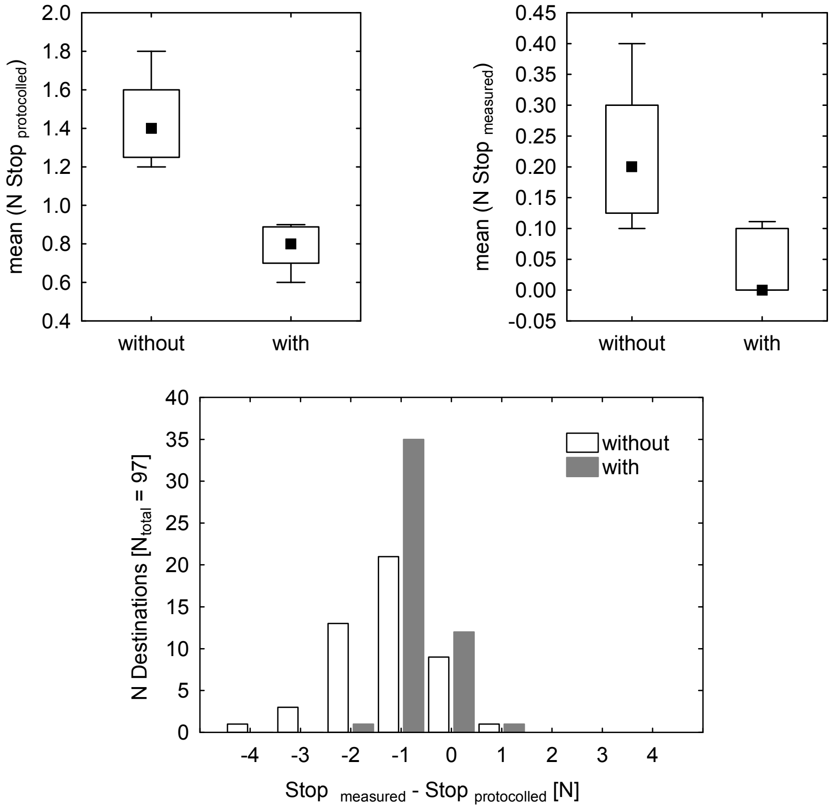

Fig.3. Influence of navigation system on measured (upper left) and protocolled (upper right) stops. The lower graph shows the absolute difference between measured and protocolled stops per destination reached.

Besides turning manoeuvres also stops during the drive that were not due to the current traffic situations were protocolled by the experimenter. For instance, drivers stopped to look on the map or to ask for the way. As can be seen in figure 3, upper graph left, drivers stopped significantly more often when driving without a navigation system (Mann-Whitney-U-test: $Z=2.61$ $p<0.01$). In the objective CAN-data, stopping is coded whenever for at least 120 seconds, speed is below 0.5 km/h or the engine is not running. Based on this definition, a significantly higher number of stopping can be found in the condition without navigation system (Mann-Whitney-U-test: $Z=2.19$ $p<0.05$; figure 3, upper graph right). As can be seen in figure 3, lower graph, the number of stops is underestimated with the used automatic event coding algorithm. For only 22% of the destinations, the number of protocolled and the number of measured stops match; in 76% of the destinations the number is underestimated and in less than 1% the number is overestimated. This result is mainly due to the long minimal duration for stops used in the event detection algorithm. The minimal duration of 120 seconds has been set to minimize the probability of wrongly classifying waiting in front of red traffic lights and similar events as stops. Therefore, reducing the durational threshold will increase the risk of including stops at intersections in the analysis.

3.4 Discussion

All the parameters that are based on objective CAN-data and that are presented in the paper are derived in a completely automatic way. To simplify the analysis of the FOT data, an analysis has to be planned that will work automatically as far as possible, without needing time-consuming video analysis. As can be shown, such an analysis is prone to be less powerful than an analysis relying on an evaluation of driving situations through an experimenter. Nevertheless, for all parameters presented in the paper the planned analysis seems at least to be promising. Based on the presented results of the video analysis, it will be tried to further strengthen the implemented event detection routines and to reduce the number of mismatches in the event detection. For instance, the time window will be reduced and each event will be split into three parts with different coding: one point in time where the turning manoeuvre has been detected, a defined time period before that point and a defined time period after that point. Then, the whole procedure of evaluating the event detection algorithm will have to start again.

The chosen approach is not suitable to evaluate the detection of very rare events like critical incidents. For such events it is only possible to check how often the event detection leads falsely to positive events. The missing of events can not be evaluated because not enough rare events occur in an experiment like the one presented here.

One further advantage of implementing the data analysis on data collected in an experimental pilot is that the feasibility and applicability of the planned data analysis can be tested. In the beginning, for instance, it had been planned to calculate the two parameters relative travel time and relative travel distance by dividing the measured through the estimated value. Doing the analysis for the experiment it became obvious that that approach is less suitable because the length and the duration of the drive impact the results too much. To give an example: having to search for the destination in a city for e.g. extra 15 minutes will have a large impact on a quotient if the whole drive consist mainly of that inner-city part. The impact will be little if there was a three hour drive on a highway before reaching the city. With calculation a difference, the impact is always 15 minutes. For an FOT, it is necessary to develop indicators that are as independent as possible of characteristics of the single drive since the situational circumstance of the analysed drives and the routes chosen by the drivers are uncontrolled. In the example given, the experimental approach was useful to evaluate the calculated indicators in detail and to assess whether the obtained values depend on characteristics of the drive (e.g. length and duration). Because the differences between diving with and without navigation system were obtained in an experiment where this was the main difference between the two conditions and other confounding factors were controlled as good as possible, the results are explained through the navigation system. Nevertheless, it is always possible that the results are not caused by the systems under investigation but by other confounding variables. This danger is especially pronounced in FOTs because here confounding factors like start and end point of the route, time of day or weather conditions are uncontrolled.

4 GENERAL DISCUSSION

The most difficult part in the analysis planned for euroFOT will be the up scaling of the results from the FOT to an EU-level. In the description of work it is described that the data from the FOT will be used to estimate the impact of the systems on driving efficiency and safety for an EU-level. Based on the results of the up scaling, a cost-benefit analysis will be carried out [2]. The data from the questionnaire survey enable the experimenter to consider system relevant background variables of drivers when recruiting the sample. Furthermore, they provide information on whether a linear up scaling of the results from the FOT to a larger sample is possible or whether certain background variables have to be considered. For instance, such background variables can influence the exposure to situations in which the system will be active.

As shown in study 2, the implementation of the data analysis for an FOT based on an experiment has several advantages:

1. It can be checked whether the planned analysis is feasible to detect system effects under controlled situational circumstances.

2. Based on a well described data set, it is possible to test the different indicators and if necessary to adopt the analysis in a way that it gets more powerful.

3. Ground truth becomes available regarding the occurrence of specific events. Thus, the event detection routine can not only be controlled for wrongly detected but also for missed events.

4. It is possible to assess how powerful a purely automatic evaluation of the data through event detection algorithms is compared to an analysis that relies on an evaluation of driving situations through an experimenter (e.g. like done with the protocols).

It is obvious that an automated event detection based on objective driving data will never be as precise as an evaluation of a driving situation by an experimenter. Nevertheless, in an FOT it is necessary to live with that imprecision since it is not possible to watch all the video. In our approach it is at least possible to describe and to quantify the possibilities and the errors of automated event detection for a well known and relevant set of data.

Both studies conducted in preparation of the FOT try to combine the expensive approach of FOTs with less costly measurement techniques. The aim is to deal with known drawbacks of FOTs in advance and to strengthen the data analysis that will be conducted with the data collected in the FOT. Based on the results that will be obtained in the FOT, it will be possible to evaluate whether the approach of combining other measurement techniques with an FOT is fruitful and helps to analyse the FOT data and to interpret the obtained results.

5 REFERENCES

[1] Jameson, 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. (2009) Deliverable 4.2 Experimental Procedures, Internal report in the European Project euroF.O.T., European Commission, Contract 223945.

[2] European Large-Scale Field Operational Test (2008) Annex I – “Description of Work”, version 1.5. Description of Work for the European Project euroF.O.T., European Commission, Contract 223945.