Report describing the proposed evaluation methodology for adaptive safety systems and remote sensing systems

AP-SP13-0022

Project no. **FP6-PLT-506503**

APROSYS

Integrated Project on
Advanced Protection Systems

Intrument: **Integrated Project**

Thematic Priority 1.6. Sustainable Development
Global Change and Ecosystems

Due date of deliverance: 09/2006

Actual submission date: 01/2007

Start date of project: **1 April 2004**

Duration: **60 months**

BASt

Revision: Final

European funded project
TIP3-CT-2004-506503

APROSYS SP1

Proposed Evaluation Methodology for Advanced Safety Systems

Deliverable 1.3.2

AP-SP13-0022

Confidentiality level: Free / ~~Core / SP / Consortium / Other~~

Rev.	Issuing date	Pages	Written by	Visa	Verified by	Visa	Approved by	Visa
A	27/10/06	28	A. Eggers (BASt) R. de Lange (TNO)	√ √	T Wohllebe	√	M Edwards	√
Modifications :								
B								
Modifications :								
C								
Modifications :								

BASt

Table of contents

1 Introduction

1.1 BACKGROUND

The European 6th Framework Programme Integrated Project (IP) on Advanced Protection Systems (APROSYS) focuses on developments in the field of vehicle safety. Sub-project 1 (SP1), titled 'Car Accidents', investigates the development and validation of evaluation methods for advanced protection systems. These are mainly focussed on secondary safety related to frontal and side impact.

Within Work Package 1.3 (WP1.3), titled 'Advanced safety functions', the focus is on the development of evaluation methods for the assessment of adaptive safety devices that employ pre-crash information from environmental sensor systems.

As defined in APROSYS deliverable D1.3.1 [8], Advanced Safety Systems are:

Advanced Safety Systems are considered to be primary and secondary safety systems adaptive to different scenarios to reduce accident severity and/or injury risk. These systems employ

- *individual occupant data and/or*
- *pre-crash information obtained from environmental sensors, dynamic car data and infrastructure*

The systems considered in WP1.3 contain two main aspects:

- remote sensing systems that scan the vehicle exterior for dangerous traffic situations
- adaptive character of safety devices, including exterior sensing and interior sensing

As a first step, an inventory study was performed in Task 1.3.1 with the overall objective to investigate methodologies defined by previous research projects to evaluate advanced safety systems and to identify issues preventing a vehicle fitted with advanced safety systems from complying with the existing European legislation.

Using the results of this work (task 1.3.1) as described in APROSYS deliverable D1.3.1 [8], a strategy has been developed to assess these systems (task 1.3.2). This has resulted in a generic evaluation methodology for advanced safety systems. The methodology will be applied to a pre-crash safety system developed in SP6 and/or PReVENT (task 1.3.3) and based upon the results, a final proposal for future methods to evaluate advanced safety systems will be considered (task 1.3.4). This evaluation methodology being developed is a code of practice which may be used to evaluate the performance of advanced safety systems and which can be used in the case that Article 8(2) of EC Directive is invoked to demonstrate that at least the minimum level of safety is maintained by the advanced safety system. In the future, this code of practice evaluation method may provide a basis for changes to regulation; however, at present the tests being developed are guidelines and represent a recommended methodology for the evaluation of advanced safety systems.

This document describes the generic method developed in task 1.3.2 that can be used for the evaluation of advanced safety systems. The methodology is described using a flowchart together with a description of the separate stages of the flowchart. Note that the development of the method is still in progress. A final version of the evaluation methodology will be defined in task 1.3.4.

1.2 TYPES OF BENEFIT STUDIES AND OTHER RESEARCH

Table 1 shows different categories of benefit study as published in [11]. These show the different categories that can be used for an estimation of the safety potential of a car or a safety system. The different categories of benefit studies differentiate between operational analysis, socio-economic evaluation and strategy assessment. The method currently developed in APROSYS SP1.3 and reported in this document is a technical assessment method that can be used to assess the safety performance of a specific in-vehicle safety system. As such, the method is comparable with the method currently used by EuroNCAP to assess the safety performance of a vehicle. Therefore, it will indicate the effectiveness of the system in real world conditions, but it does not incorporate the socio-economic or long term issues.

	Categories		
Features	**Operational analysis**	**Socio-economic evaluation**	**Strategy assessment**
Kind of evaluation	Technical assessment of operational effectiveness	Economic evaluation of social impacts	Long term strategic assessment on a political level
purpose	Determiniation of technically superior solution	Indication of social worth	Estimation of fundamental opotentials and long term risks
Alternatives to be evaluated	Individual technical options	Contrete, public investment projects	Entire technologies
Perspectives	Control and optimise a technical solution	Optimal allocation of scarce resources	Provision of a basis for comprehensive progress
Result	Statements on technique-specific operational performance	Indentication of contrecte social gains and losses	Appraisal of far-reaching consequences

Table 1. Categories of benefit studies, TRL Report 220 [11].

In other research projects, benefit studies are used to identify, for example, where the focus of the future investments could or should be (e.g. vehicle technology or infrastructure etc.). On an operational level, such studies are performed in the TRACE and PReVAL projects, whereas the eIMPACT project performs an evaluation on socio-economic level.

Besides the work performed in APROSYS WP1.3, several other groups have discussed and developed methods for the technical evaluation of advanced safety systems. For instance in the RESPONSE 3 project of the IP PReVENT, a code of practice for the design of advanced driver assistance systems (ADAS) was developed. Here the focus is on the development of these systems. Also, a separate project PReVAL is currently under development for the evaluation of the systems developed in PReVENT . Furthermore, groups like EEVC WG19 and the Beyond NCAP initiative aim to develop methods for the evaluation of advanced safety systems that are not covered by current regulations or standards.

Within APROSYS WP1.3 the relevant work performed by the other groups (especially PReVENT and CONVERGE) was considered in the development of a generic evaluation methodology. Although the different groups do have a specific focus, discussion of the proposed methodologies in the other groups is seen as necessary to make sure that the work done in the different groups is complementary and to avoid unnecessary overlap.

2 Generic evaluation methodology

2.1 FLOWCHART EVALUATION METHODOLOGY

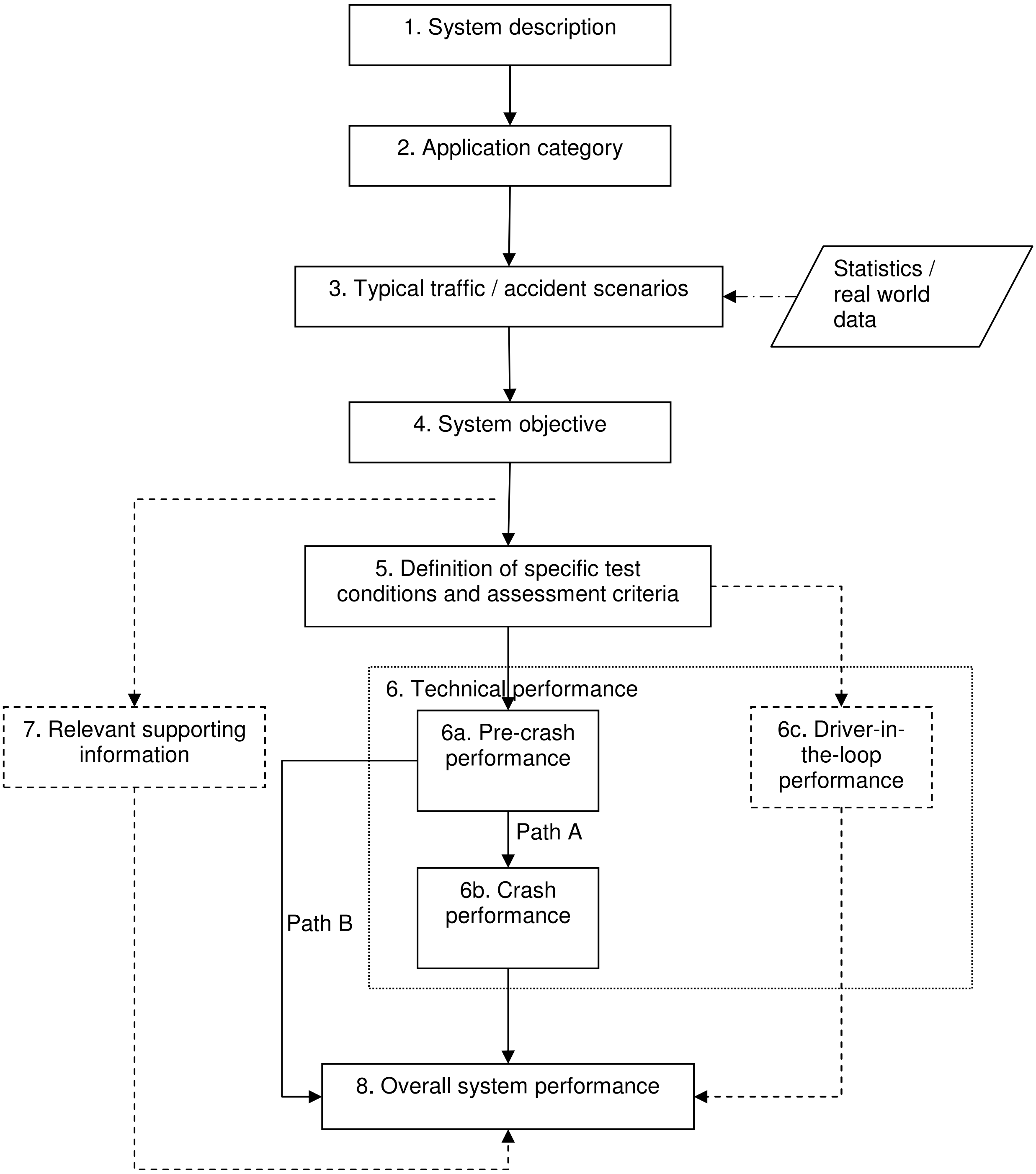


Figure 1. Generic evaluation methodology for advanced safety systems

2.2 DESCRIPTION OF THE FLOWCHART COMPONENTS

In the following sections, each of the stages of the flowchart shown in Figure 1 are described in more detail. In some cases examples are provided for clarity and open issues for further investigation are added where applicable. The following sections should therefore be read and interpreted with reference to Figure 1 (above),

1. System description

The system description is a brief description of the product or function to be evaluated. The following key characteristics of the system should be addressed:

- The application name and type;
- The major technologies and application which is going to be validated (especially relevant for integrated systems);
- The functionality or service offered;
- The technical system limitations in terms of the conditions under which the technology is not expected to function effectively. For example, some sensor systems may in general not be able to function at night or with fog.

Different integrated systems will have different functions, but they could use the same sensors to provide input to different actuator systems. An example is for instance the ABS and the ESP systems that partly use the same information. For the evaluation, these systems will be handled as independent systems. In general most advanced safety systems consist of three main components: a sensing system, an evaluation unit, which combines and evaluates the signals from different sensors, and finally some kind of actuator system , for example a pedestrian airbag or a driver warning device. Most advanced safety systems do not rely on a single sensor or type of sensor, but utilise several kinds of sensor. For example a pre-crash system might evaluate signals from radar, laser-based and vision-based sensors. At the same time, some of those sensors might be also used by other systems; for example driver assistance systems, lane-departure-warning or parking assistance systems.

The evaluation process for advanced safety systems proposed in this project should always be performed for one complete system which consists of all the relevant subsystems and devices . For example, this will include sensors, a decision-making unit and actuator.

The information in this box is closely related to box 2 of the CONVERGE evaluation methodology [5], [13]. For the CONVERGE evaluation flow chart, please refer to the appendix.

2. Application category

In the application category, information is provided regarding which road-users are protected and in which accident types the protection is offered. Note that there might be a difference in the accident-causing situation and the accident type.

The categories of road users protected by the safety system are:
- car occupants (self protection or partner protection)
- vulnerable road users (pedestrians, bicyclists)
- motorcyclists (considered separately from non-motorised vulnerable road users)

The categories of accident types for which the safety system is considered relevant are as follows. Note that this is not necessarily the direction which the sensor is looking:

- frontal impact
- lateral impact
- rear-end impact
- rollover

The vehicle category onto which the safety system is implemented:

- passenger car
- heavy goods vehicle (truck)
- bus
- motorcycle

This application category (including the vehicle category) is used as input for box 3 where the accident scenarios are selected. This box is related to the ‘use cases’ as defined by CONVERGE [5], [13].

3. Typical traffic / accident scenarios

In box 3, a set of generic scenarios should be defined based on accident data or real world situations. These scenarios should include relevant accident scenarios (linked to box 6a and 6b) and, depending on the type of system and actuator, relevant traffic scenarios that may include situations with a high accident risk, near miss scenarios and other critical situations.

As the focus in the current project is on the traffic situation in Europe, there is a need for an accepted accident database on a European basis. However, as long as there is no European database, it should be reviewed which databases are suitable in terms of relevance and level of detail. A possible source of suitable databases could be identified from the SafetyNet project (see figure below).

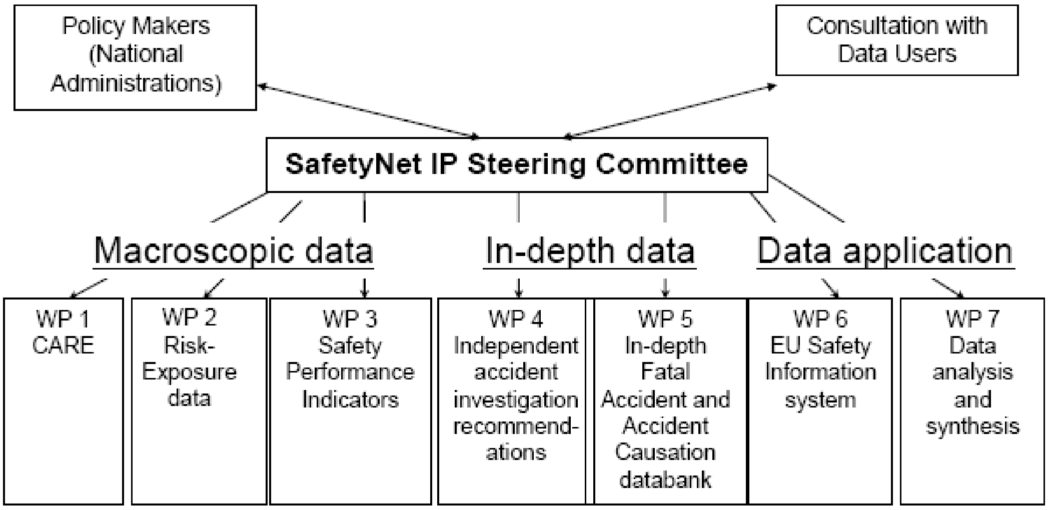


Figure 2. Project structure of SafetyNet", [12]

In general, two main sources for accident data are available. One type is descriptive national accident data (microscopic data), which is collected in all European countries by the police. In some European countries, also various in-depth accident databases exist containing so called macroscopic data. Both sources of data have their specific advantages and disadvantages for defining the relevant scenarios to evaluate advanced safety systems. In order to get a comprehensive insight into relevant scenarios which enable definition of testing methods, both data sources should be taken into consideration.

Descriptive / microscopic data

Every country has a national traffic accident database containing all accidents reported to the police. Within CARE (Working Package 1 of the SafetyNet project) an effort is made to combine the different data from national traffic accident statistics in one single European database. However, several problems still exist in realising this aim of a single European dataabase. One problem is that the level of detail of the collected data varies between countries. Also differences in certain definitions used in the coding of accident parameters exist which results in some difficulties comparing or combining data from different countries. This is even true for basic definitions, e.g. definition of road traffic related fatality. In some countries traffic victims are only counted as road traffic fatalities if the death occurs within 24 hours after the accident happened. In other countries also the death of a person dying 30 days after the accident is counted as road traffic related death. Corrections factors are being developed to account for these differences. Furthermore, many other differences in definitions exist between the different databases and this makes it very difficult to combine the data from different countries. Another problem is that not all countries can provide their data due to restrictions by data protection laws. For example Germany has very strict laws protecting personal data, which prohibit German participation in the CARE database.

Another disadvantage of descriptive accident data source is the level of detail recorded is often too low for defining relevant evaluation scenarios. The reason is that the main focus of the accident parameters collected by the police for these kind of databases is not for determining how the accident happened (accident reconstruction), but to determine the more basic accident data related to the outcome of the accident..

However, despite these facts microscopic accident data are often very valuable, because In-Depth data normally consists only of a low number of accidents and might not be representative of all accidents in the respective country. Thus, evaluating the official accident statistics is necessary to get a representative overview of the statistical importance of specific types of accident.

In-Depth-Accident Investigation

In-depth accident data is collected by specialised teams of accident researchers, consisting of physicians and technicians, who gather detailed information about every accident at the scene. Using detailed protocols the researchers collect parameters about environment conditions during the accident, detailed information about damage to the car and about injuries of persons involved in the accident. Some in-depth databases also include a detailed accident reconstruction. The level of detail of most in-depth databases should be sufficient to get all necessary information needed to derive all available test parameters for the relevant accident scenarios. This is because of the focus on accident reconstruction and accident causation which can be linked to the accident outcome.

In a report of EEVC WG19 on the interaction of primary and secondary safety [2] the following parameters were identified which could be useful to define accident scenarios:

- Impact type
- Impact velocity
- Type of injury
- Severity of injury
- Collision object
- Driver reaction (braking / steering).

For most safety systems additional parameters will be required to define the relevant scenarios, e.g. environmental data. For some systems, very detailed information about accident parameters and boundary conditions could be necessary. It has to be investigated whether there is an accident database which can provide information on the required level of detail. Finally, the number and type of required accident parameters will be very different depending on the specific safety system under evaluation.

An example, for a very detailed in-depth database such as GIDAS (German In-Depth-Accident Investigation), which is based on more than 10,000 cases collected by investigation team in the Hannover and Dresden area of Germany, the level of detail would be sufficient to define accident scenarios and all necessary information about testing conditions. However, it might not be representative for accidents in other European countries.

In many European countries in-depth accident investigation is done by various companies and institutions. Some car makers have their own in-depth accident research team which is collecting this type of microscopic accident data. Also insurance companies and federal institutions are collecting in-depth accident data. Due to the high number of different databases for this kind of data, it seems even more difficult to identify one accepted database, which is relevant for the accident situation in Europe.

Within the SafetyNet project in WP5 it is planned to built up an in-depth database of about 1,300 fatal accidents in Europe similar to the US FARS database (Fatal Accident Reporting System). A second in-depth database, which will be established within SafetyNet WP5, will consist of 1,000 accidents and provide information on key risk factors. These databases, once completed and available, could be a valuable source to define accident scenarios representative for Europe. However, due to the limited number of cases, addressing a specific type of accident, which might be the case with the evaluation of some advanced safety devices, may result in the number of relevant cases present in this database being insufficient. In this case, it is recommended that the in-depth data bases of different countries should be evaluated.

Another approach to develop an in-depth accident database for Europe was done within the EACS project (European Accident Causation Survey) [3]. The level of detail is comparable with GIDAS and also contains information about the pre-crash phase. However, the number of cases is limited, and although companies from Spain, Italy, France Germany and the Netherlands where contributing to the project, 67% of the accidents where provided by Germany. Thus, a representative European in-depth accident database is still not available.

Defining relevant accident scenarios

The basis for developing the assessment scenarios from the accident data is the information about the system and its application. Based on the intended application category of the system, the frequency of different injury severities which occur in different accident types and collision partners could be identified as a basis for defining scenarios. As the scenarios form the basis for the specific test conditions that have to be defined in box 6, a certain level of detail is necessary. This includes parameters from in-depth accident investigation like road layout, relevant objects involved (e.g. the presence of trees), type of traffic involved, impact conditions (angles, velocities), weather conditions, etc. It should be investigated how a limited set of relevant scenarios representative of the majority of the real world accidents could be derived.

The definition of relevant scenarios is a critical step in defining the test conditions, since these must be representative of the most important situations in the real world for which the system is expected to provide a benefit. Thus, it is recommended that the scenario selection process should consider the following general steps:

1. Identify accident data source applicable in scope and detail representative of the region for which the assessment is being made
2. Identify accidents which occurred in general conditions where the system should be functional (i.e any limitations of the system are excluded)
3. Identify accidents involving the user group for which the system is designed to offer a benefit
4. Assess frequency of fatal, serious and slight casualties by relevant accident parameters, including, but not limited to:
    i. Road type
    ii. Impact speed
    iii. Impact angle
    iv. Impact partner
    v. Road conditions

In addition, attention should be paid to accident and injury causation. Although this information may not be necessary to define the test scenarios, it is important for estimating the real world relevance of the system. Relevant injury mechanisms should be identified to be able to find out which proportion of injuries could be avoided by the advanced safety system. For example, if accident data showed that a certain percentage of pedestrians are severely injured by contact with the A-pillar, and the safety systems can only avoid injuries caused by contact with the hood and windscreen, this gives important information about the effectiveness of the system.

A ranking of the scenarios should be determined according to their importance or relevance. Possible criteria to identify the relevance are:

a) Accident scenarios resulting in the most killed and seriously injured people
b) Accident scenarios resulting in the greatest injury cost (severity multiplied by societal injury cost value)

According to method a) for example, only accidents with killed or severely injured victims (AIS3+) are considered. The scenarios fulfilling this condition are ranked according to their frequency to find out the most relevant.

The second method also takes into account the societal-costs of injuries. Several European countries are annually calculating these cost factors. But differences in calculation procedure exist and not all countries are providing this type of information.

For example in Germany the Federal Highway Research Institute calculates the costs of road accidents on an annual basis. For the year 2004 personal injuries and property damages in road accidents caused socio-economic costs of 30.9 billion Euro. Personal injuries made up 49.3 % of this figure, property damages 50.7 %, [4]. The costs are calculated according to a model taking into account all consequences of an accident according to the severity level of personal injury (slightly injured, severely injured, killed). The most recent figures of the year 2004 are shown in the table below:

Injury severity	Costs in Euro/person
Fatally injured	1,161,885
Severely injured	87,269
Slightly injured	3,885

Table 2 "Socio-economic costs of personal injury due to road traffic accidents in Germany 2004" [4].

Similar figures are reported by the Italian National Institute for Statistics. The calculated social costs for each road traffic victim are 1,281,778 Euro and 24,726 Euro as an average value for each injured traffic victim.

4. System objective

The objective of the system shall be defined on a general level (generic safety function). This definition will provide the key indicators to determine the parameters which must be assessed in the evaluation process of the system. For this, it is necessary to specify the manner in which the system is expected to reduce the injury risk. A list of possible examples is given below. Some are taken from the guidelines described in a report by EEVC WG19 [2].

- reduction of speed (for example by a pre-crash braking system),
- preparation of the occupants for the impact (move the occupant to the design position for an appropriate functionality of the safety systems)
- preparation of the vehicle for impact, e.g. by pre-arming or pre-firing the actuators or restraint systems (both reversible and non-reversible)
- optimisation of safety systems to the impact angle of the vehicle (e.g. motorised steering wheel);
- collision mitigation or collision avoidance by alerting the driver using audible, visible and/or haptic cues (no direct link to injury reduction).

The description should also include limitations that relate to the intended system objective, as both the system objective and the limitations are required inputs for the evaluation of the system. Unlike box 1, where general technical limitations are described, such as situations in which the sensor cannot function at all, the limitations described in this box should focus on performance risk. For example, the performance of an emergency braking system might be reduced by a wet or icy road surface, despite the fact that the sensing system functions appropriately and the obstacle is detected correctly. Any identifiable performance risks or limitation associated with the system should be mentioned here.

It should also be investigated, as described in the deliverable D11.1 by the RESPONSE Project [7], whether or not the system limits are predictable for the driver in different environments, weather and visibility conditions (e.g. fog).

5) Definition of specific test conditions and assessment criteria

In the box 'definition of specific test conditions and assessment criteria', the following are defined:

- The information from box 3 and the system description and objective in box 1 and 4 will be transformed into reproducible test scenarios to assess the performance of the advanced safety system
- Both the test conditions and the assessment criteria should be defined for the pre-crash tests and the crash tests. Box 5 could also provide input to the definition of the driver in the loop tests.
- In addition to test conditions developed from generic scenarios, it might be useful to also develop test conditions which are not derived directly from box 3.

All details about the test conditions should be defined in this box including possible new test devices which might be necessary, for example new test environments, new barriers for the sensor tests or new kinds of impactors or dummies for the crash tests.

In this box the assessment criteria used to evaluate the benefit of the advanced safety system should be defined. Standard legislation or consumer crash test criteria could be used for this purpose. However, it is also possible to apply other criteria to assess the system.

A possible method to assess the benefit of a safety system is the computation of a shift of injury risk curves described in a report by EEVC WG19 [2]. This could be a possible method used to evaluate the effectiveness of the system. Other methods to assess the system benefit will be developed in the TRACE project, [10], and could be used once they are available.

Note that the definition of the test conditions and criteria itself is not part of this project and this should be addressed by the party intending to use this generic method. Within WP1.3 this will be performed only as an example for the system used in task 1.3.3 to evaluate the applicability of this methodology.

6. Technical performance

The assessment of the technical performance of the safety system is achieved by the assessment of the pre-crash performance and the crash performance, respectively. In order to ensure that the pre-crash system does not have a negative influence on the driver during trigger events, “driver in the loop” tests are performed to ensure that the level of safety in these situations is not less than that without the system fitted. This is the case if this methodology is being used to obtain compliance existing regulation or if this methodology is adapted in the future for new legislation.

Within the methodology proposed in this report, the assessment of the safety function using driver-in-the-loop tests should be considered optional. During the development process a car manufacturer or supplier will perform a sufficient number of driver-in-the-loop tests to account for all possible risks of the system. Within the assessment methodology presented here, evaluation of the driver-in-the-loop performance should only be mandatory, if the driver behaviour could influence the performance of the system, e.g. brake assistant system which requires driver input. It should be noted that evaluation of the whole HMI performance is not the intention of the methodology presented here.

Finally a generic methodology to evaluate the possible risk for the driver is required in the future.

6a) Assessment pre-crash performance

For pre-crash performance, test scenarios and criteria defined in box 5 will be used in non-destructive tests to assess the pre-crash performance of the system.

A range of scenarios will be assessed ranging from very relevant to less relevant. The number of test scenarios passed will be used to rate the precrash performance. In this way it will be shown that the system works as intended in a certain percentage of the real world accident cases. An option could be that a minimum amount of scenarios has to be passed before it is allowed to perform the crash test (box 6b) with the system activated.

The definition of criteria to decide if a specific scenario is passed or failed should be achieved by experts of the respective organisations or institutions applying the assessment methodology. For each individual advanced safety system, different specific kind of criteria might be useful to evaluate the pre-crash performance in the defined accident scenarios.

For instance, one possible convenient assessment criterion which could be applicable for some safety systems is the trigger time. If this time is not available the system-in-function time might also be used as assessment criterion. The system-in-function time is defined as the time that the (actuator or the) system is fully activated.

It will be an issue how to measure the system-in-function time, especially as it might be impossible to get direct access to the system (on a signal level). For the pedestrian GTR (Global Technical Regulation) the CLEPA/OICA proposed that the manufacturer provides the HIT (head impact time) and the sensor time. Based on this information it is decided if the head-to-bonnet impact tests can be performed with the bonnet in its deployed state [9].

In order to be able to decide if the system works properly, there is a need to define the minimum or reference system-in-function times for the different scenarios and the different systems. As an example for pre-crash braking the system-in-function time could be defined as the time that a vehicle has reached a certain level of deceleration.

For other systems a minimum system-in-function time might not be an appropriate criterion, because a very exact timing of the complete activation of the system could be necessary for optimal system performance. For example, for some type of pedestrian protection airbag systems, the airbag might be not efficient if they are triggered too early. Thus for these systems, not only a minimum system-in-function-time, but also a time window of activation must be assessed.

For some systems it could be difficult to measure the system-in-function-time. An appropriate specific method has to be developed for the system to be tested. For example for an airbag system, measuring the airbag pressure could give the time maximum pressure, which could represent the system-in-function-time. The pass/fail criteria (system-in-function time) for each scenario can be set by different organisations (industry / consumer type / regulatory body) depending on the focus of the evaluation being undertaken.

In case of semi-autonomous safety systems with driver input required like advanced brake assist systems the system in function time will be affected by the human reaction. It is expected that an assessment of these systems can be made using a synthetic driver.

However, how to account for this factor in the assessment of the benefit of such a system should be investigated further.

After the pre-crash performance has been assessed, two different paths can be followed to complete the system evaluation:

- path A should be followed in case the actuator is expected to have an influence on the crash performance, or if the crash performance is not available;
- path B should be followed in case the actuator will not influence the crash performance and the crash performance without the system is already available. An example could be for instance a driver warning system.

6b) Assessment crash performance

If path A is followed the crash performance must be assessed after the assessment of the pre-crash performance. The sensor system is assumed to work as intended to predict the collision and provide the correct trigger signal to evaluate the crash performance of the system. A method to define the required system-in-function time, based on the results of the pre-crash performance evaluation still remains to be developed.

Using standard test procedures from regulations or EuroNCAP as a basis, the actuator system will be triggered such that the system-in-function time complies with the results of the pre-crash performance tests or the system will be triggered by the sensor system itself.

During the test, the measurements will be made according to the standard procedures (if they are available) and the results will be compared and rated using the standard crash test as a reference.

For some new safety systems, the standard procedures from regulations or customer tests might not be able to show the expected benefit of the system. For example in some cases the dummy kinematics could be significantly different to those of a real occupant or the dummy injury criteria might not be able to show the benefit in injury reduction, which could be expected for a real human. In that case the system supplier has the opportunity to provide his own method for the evaluation of the crash performance. This new test could be possible using new testing devices as well as numerical simulation to demonstrate the benefit of the system.

6c) Assessment driver-in-the-loop performance

In the driver in the loop tests, the scenarios and criteria defined in box 5 will be used to investigate if and how the system influences the safety of the occupants during no-crash conditions. The influence can be negative (e.g. in terms of distraction in case of a false alarm) or neutral (no significant influence).

In addition, driver-in-the-loop tests might for instance be useful to evaluate the driver reaction or to generate missing input for scenarios. Therefore the suitability of such tests will be investigated in the current project as part of the development of the methodology. A possible tool which could be used to evaluate the driver-in-the loop performance is a driving simulator.

Another possible methodology to assess the effect of driver distraction by in-vehicle-information-systems (IVIS) and advanced driver assistance systems is currently developed within the EC project AIDE. Within working package 2.2 methods and tools are developed to assess driver workload and distraction by such systems, [6].

The methods developed within the AIDE project could be used for evaluating the effect of advanced safety systems when activated during no-crash condition. In the AIDE deliverable D2.2.1 [6] the Occlusion technique is presented, which can be used in order to measure driver visual distraction caused by IVIS. It should be investigated if this technique is suitable for assessing the driver-in-the-loop-performance of advanced safety systems. An excerpt on this method from the mentioned AIDE report [6] is provided in the Appendix.

So called Car Clinics with naive subjects could also be a method for evaluating the influence of driver-in-the-loop on the system performance. Car Clinics as a method to evaluate ADAS are explained in the RESPONSE3 report [7]. According to RESPONSE3 in a car clinic typically a sample of drivers will experience and assess a new developed product before market introduction. For the validation of ADAS a dynamic car clinic (driving clinic) is necessary. An excerpt on Car clinics from the RESPONSE3 report is provided in the appendix.

7. Relevant supporting information

This box is considered optional as it will give the manufacturer of the system the possibility to provide relevant additional input for the evaluation of the system. Optional studies can be retrospective (using accident data) or realistic estimations based upon field studies or numerical simulations.

Examples could be:

- driving tests with professional drivers;
- field tests (robustness of the system);
- driver simulator tests;
- monitoring the real world performance (long term effect);
- benefit analysis using statistical methods or simulation

In the current methodology, the focus is not on the manufacturer input although this is an important source of information which will be used in the future assessment/evaluation of pre-crash systems. However, the methodology being developed here is intended to work even if the manufacturer does not provide any input. Therefore, this box is seen as additional for the purposes of developing these guidelines for evaluation.

However, it is recommended that for any future regulation and for systems gaining approval under the current Article 8 (2) route, this information provided by the car manufacturer should be strongly required. This evidence will demonstrate the safety of the system in a much greater range of road situations than can be assessed in specific evaluation tests. Future regulation should define the required information further in each of the categories defined above (box 7).

Note that a methodology to estimate the benefit of advanced safety functions will be developed in the EC project TRACE, [10]. In this project , specific systems will be assessed to evaluate the developed methodology. A methodology for a priori and a posteriori evaluation of safety function effectiveness will be developed. A priori methods may be especially relevant to evaluate new advanced safety systems which not yet have been introduced into current vehicles. As the methodology becomes available, it could for instance, be used by the OEM / supplier to estimate the benefit of their system. Within TRACE a method will also be developed to determine the relation to socio-economic impact of safety functions.

8. Overall system performance

In the system performance stage, the quality of the system that has been evaluated will be addressed. In this stage the driver-in-the-loop performance, the pre-crash performance and the crash performance will be combined to calculate an overall systems performance. If optional studies or supporting information have been provided by the manufacturer, the results of these studies will also be taken into account.

The procedure to determine the overall system performance is open for discussion as it still has to be developed. It could for instance be similar to the star rating now used by EuroNCAP. In the overall system performance not only the evaluation itself, but also the relevant support information as provided by the OEM or system supplier may be taken into account. Note that the procedure to assess the overall system performance, including the performance limits, could be different for the different types of organisations using this methodology for the assessment of advanced safety systems.

For instance the methodology for legal regulations to assess the overall system performance will be different compared to the method used in a consumer test. The focus in the legal test is to guarantee a base level of safety performance which systems must fulfil to be admitted to the market, whereas the intention of consumer tests like EuroNCAP is to distinguish the performance of products already available on the market.

For OEMs or suppliers of advanced safety systems the focus will not only be on the general socio-economic benefit for the whole public, but specifically the benefit for the customer, who is going to purchase the product. Thus, the car manufacturer might develop a method to assess the overall system performance which shows the system benefit for the individual customer.

Insurance companies might want to use this methodology to distinguish the overall performance of available products, with a focus on the socio-economic cost benefit. However, this might not only be the cost resulting from injury, but also the cost benefit in general which could be achieved by advanced safety systems. Thus, a rating procedure developed by insurance companies could also take this into account. Based on the outcome of the overall system evaluation the insurance companies can support the market launch and penetration of the system by giving a discount of the insurance rate for the specific car.

Due to these differences in focus of interest at rating the overall system performance every organisation or institution might come up with a different individual methodology for the overall assessment.

3 Discussion

The methodology described previously, is a generic methodology to evaluate the complete system. Therefore it is expected to be applicable for a wide range of advanced safety systems. For all types of systems the boxes in the flow diagram may have to be filled in more specifically. In task 1.3.3 the developed methodology will be worked out more in detail for one specific system in order to evaluate the applicability of the methodology.

Although advanced safety systems will have the objective to improve the safety potential, there is the possibility that some systems may have side effects and will reduce the injury risk in one area, but increase risk in another area. Because of the complexity it was decided not to include a separate risk assessment in the method currently developed. Of course a system supplier or OEM might indicate these risks. This could be done in the system description box. Any future regulation or consumer assessment should take account of these using objective assessment.

In the case the methodology is applied by a consumer organisation or regulatory body it should be considered to make a distinction between systems providing ‘additional safety’ and systems providing ‘replacing safety. A further division of these safety types should be considered as follows:

- <u>Additional secondary safety</u>

This means that a standard (current) secondary safety devices are fitted and operational or secondary safety is achieved by passive design measurements. The pre-crash system provides extra level of safety.

An example could be a pre-crash activated pedestrian airbag system at the vehicle front. In this case there has to be a sufficient level of pedestrian protection only by passive measurements (appropriate design of the bumper and hood). The pedestrian airbag system provides additional pedestrian safety.

For current legal regulation this mean that for the certification only the crash performance will be tested without the additional safety system activated. The vehicle has to fulfil the standard legal requirements without the additional system. In tests performed by consumer organisation like EuroNCAP or insurance companies, the assessment may be performed according to standard test and rating procedures without the additional system, but also the pre-crash and crash performance tested with the additional system activated. The evaluation results of this test should demonstrate the improved performance of the additional advanced safety system using the methodology described in this report and account for the additional safety for example by additional positive modifier points in the rating.

- <u>Replacement secondary safety</u>

A pre-crash system fully replaces the function of a conventional secondary safety system. The safety system is activated on basis of pre-crash, not in-crash information. For example a front passenger airbag deployed before the crash is a replacement secondary safety system, if no other in-crash airbag activation is done. Also a pre-crash activated pedestrian airbag system can be a replacement secondary safety system, if no other conventional passive measurements are provided to address the pedestrian safety of the car.

Legal regulation tests have to assess the functionality and performance of the advanced safety system and make the decision about pass or fail of the system with the advanced safety system activated based on the pre-crash and crash performance according to the methodology described in the report. Consumer and insurance tests will also evaluate the pre-crash and crash performance according to this methodology, but might use a more differentiating rating procedure to assess the overall system performance and to encourage improvements in the level of safety.

- Replacement primary safety

This type of safety might become available for some aspects of vehicle safety in the future. A pre-crash system replaces some of the safety which today is provided by secondary systems by improving the primary safety of the vehicle.

For example a collision avoidance system based on pre-crash sensor information might be able to almost completely avoid some types of collisions, where normally some secondary safety systems are necessary to mitigate the results of this type of collision. If this type of collision can be assuredly avoided for all imaginable scenarios by the advanced primary safety system, the respective secondary would no longer be necessary. Thus, the safety of this secondary system is replaced by a primary safety system.

In consumer and legislation test procedures, the pre-crash performance of the advanced primary safety systems must be assessed to demonstrate and assure that the system works as intended in all possible scenarios and that respective secondary safety is not necessary any more. Once this has been satisfactorily demonstrated, there may be no requirement to assess the secondary safety for these types of collision.

- For systems providing additional secondary safety where the system results in either the conditions of the test being violated or the sensing system is not compatible with the test tools, the vehicle will be tested by the regulatory body according to the standard test procedures with the system de-activated. In this case the requirements are fulfilled using current (passive) safety measures and the advanced safety system is considered to provide extra safety upon the standard tests, which could be assessed according to the methodology currently developed.
- If the pre-crash performance effectiveness of the advanced safety system is very high with respect to the identified relevant scenarios, the test with the system activated could count as consumer / legislative test if the system does not alter the conditions of the test or have a sensing system incompatible with current testing procedures.
- For systems considered to be “replacing primary safety” systems, the system could be used to replace some passive safety measures in the vehicle. However, the switch from additional safety to replacing safety must be developed very carefully in order to guarantee that a vehicle with a replacing safety system does not have a worse overall performance as a vehicle with a good additional safety system. This is likely to be many years away. An option could be that a minimum amount of scenarios has to be passed, as well as satisfying a fleet trial and other manufacturer derived information (box 7), before the system is considered a replacing safety system and it is allowed to perform the crash test (box 6b) with the “replacing safety” system activated. Additionally, a rating could be based on the proportion of different scenarios where the sensing system worked as intended. This will not be addressed in the methodology currently developed.

A number of issues / open points still have to be discussed:

- The step from the pre-crash performance to the crash performance assessment has to be worked out in more detail. This is likely to be different for different advanced safety systems.
- A method to rank/weight the relevance of the defined accident scenarios needs to be agreed.
- A method has to be developed to assess the overall performance of the system (pre-crash and crash) in order to arrive at a final judgement
- It should be investigated if the results of the European project TRACE on benefit analyses of safety systems once available could be integrated into the methodology proposed in this report
- In the current proposal it is mentioned that the methodology developed can be used to evaluate an advanced safety system, whereas the overall safety in a certain accident situation (and crash test) is provided by the complete vehicle (that can contain various safety systems). This would be important for vehicles equipped with advanced secondary safety systems as well as advanced primary safety systems.
- A method for risk analysis of driver-in-the-loop has to be developed or identified to be included in the methodology presented here
- Box 7 provides now the possibility for the supplier or OEM to provide relevant supporting data. Should we consider recommending a specified fleet test as part of the evaluation methodology and provide guidance for number of vehicles and distance travelled in urban/rural situations? This trial could be performed with the actuators deactivated, with the purpose being to assess any significant negative side effects (e.g. the frequency and situations of false positive and failed activations). As the specific scenarios will not cover all the relevant situations on the road, this could help to ensure with a high degree of confidence that the system does not have a negative effect on safety in real world. Depending on the type of actuator, the implications of false positive or missed activations are more important for some systems. OEMs would do this anyway as R&D but providing this information could be at least strongly encouraged by providing guidelines and by the rewarding system. Or should it even be mandatory rather than optional?
- It was decided not to include a separate risk assessment. Therefore the evaluation method itself should be able to pick-up any reasonable negative effects. If any significant risk is not picked up in the scenario tests, this would be bad. – this might be another reason why a fleet trial would be a good idea to be included in our recommendations.
- Although the developed methodology is generic and can be used by different organisations (Industry / Consumer organisations / regulatory bodies) it is recommended to explain / describe how we expect this method to be implemented.
- A European in-depth accident database is needed which contains a representative number of accident cases of all relevant European countries and a sufficient level of detail to define the parameters and conditions of the relevant test scenarios.
- It is recommended to discuss the method with other relevant groups / initiatives (PReVENT, EEVC WG19, Beyond NCAP) to inform these groups but also to get some feedback from them.

4 References

[1] Aparicio, F., (2006) "EEVC WG 19 Review Report 1, Primary and Secondary Safety Interaction"

[2] Aparicio, F. on behalf of EEVC WG19 (2005), “EEVC WG19 Activities on Primary and Secondary Safety Interaction, Paper Number 05-0462, ESV Conference.

[3] Bauer, A. (2003). EACS. European Accident Causation Survey.

[4] Höhnscheid, K.-J., Straube, M. (2006), " Socio-economic costs due to road traffic accidents in Germany 2004"

[5] Jesty, P. H., Gaillet, J.-F., Giezen, J., Franco, G., Leighton, I., Schultz, H.-J., (1998), " CONVERGE Deliverable DSA2.3: Guidelines for the Development and Assessment of Intelligent Transport System Architectures"

[6] Johansson E., Engström, J., Cherri, C., Nodari, E., Toffetti, A., Schindhelm, R., Gelau, C., (2005), "AIDE: Deliverable D2.2.1: Review of existing techniques and metrics for IVIS and ADAS assessment"

[7] Knapp, A., Schwarz, J., (2006), ""RESPONSE3: SP Deliverable based on D11.1 (V1.2):Draft of a Code of Practice for the Design and Evaluation of ADAS"

[8] McCarthy, M., Lange, R. de, Lemmen, P. (2006), “Review of Regulations and Research regarding Advanced Safety Functions, Deliverable D1.3.1”, APROSYS Document AP-SP13-0011, DRAFT Version.

[9] OICA/CLEPA, (2005) “Certification Standard for Type Approval Testing of Active Deployable Systems of the Bonnet / Windscreen Area”, Draft document INF GR PS 141.

[10] Page, I. (2005) "Accident Causation Analysis and the Evaluation of the Safety Benefits of Technologies. The TRACE Project."

[11] Perret, K.E., (1996), “Review of Potential Benefits of Road Transport Telematics”, TRL Report 220

[12] Thomas, P., (2004) "The New European Road Safety Observatory – SafetyNet

[13] Zhang, X, Kompfner, P, White, C and Sexton, B, (1998),CONVERGE Deliverable D2.3.1 "Guidebook for Assessment of Transport Telematics Applications: Updated Version"

5 Appendix

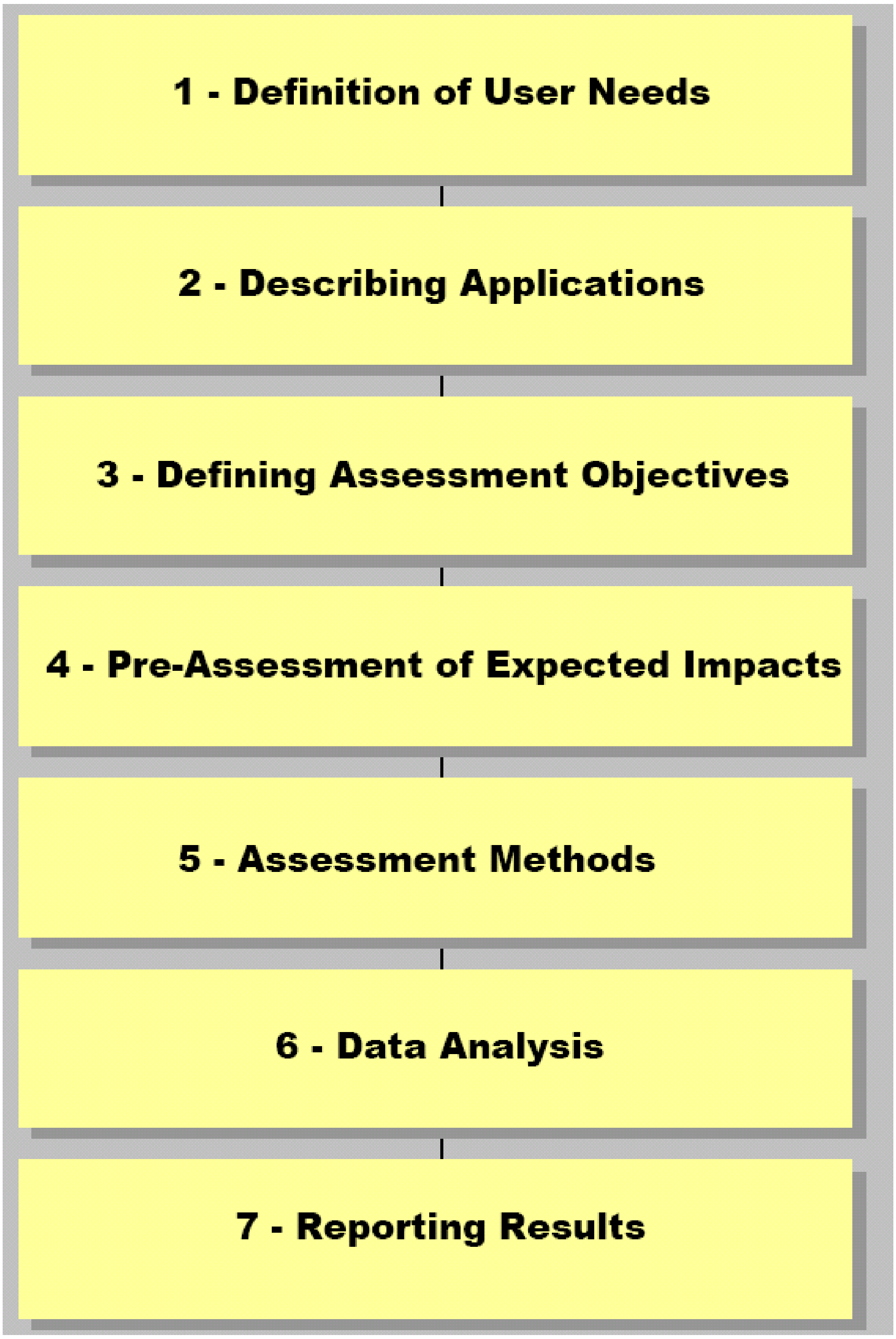
The appendix contains excerpts from other reports, which were used in this paper.

First the CONVERGE evaluation flow chart is shown, which was used as guideline for developing the evalutation methodology presented in this paper.

Second an excerpt from the AIDE project is provided explaining the "Occlusion technique", which is mentioned in box 6c as one possible method to assess the diver-in-the-loop performance.

Third an excerpt form the RESPSONSE code of practice for evaluation of driver assistance systems is given. The car clinics explained in this code of practice are proposed in box 6c as another possible method to assess the the diver-in-the-loop performance of advanced safety systems.

1. Overview of Assessment Process from CONVERGE Deliverable D2.3.1 [13]



2. Information on the "Occlusion technique", Excerpt from AIDE deliverable D2.2.1: Review of existing techniques and metrics for IVIS and ADAS assessment [6].

The measurement of driver visual distraction induced by in-vehicle tasks has become a major issue in Human Factors research (c.f. Gelau, 2004). One possibility to study effects of drivers' interaction with new on-board systems while driving could be to perform experiments in real traffic or in a driving simulator. Because this approach is very demanding and expensive, there is a need for a method that is easy to use and applicable in the very early stages of system development (e.g. Krems et al., 2004).
The *occlusion technique* has recently come under consideration as an assessment tool that fulfils these requirements (e.g. Gelau & Krems, 2004). The goal of this chapter is to briefly describe the application of the occlusion technique and discuss its validity and reliability as a method for HMI assessment of IVIS.

The primary task of driving relies heavily on the visual channel which is also the case for many invehicle tasks (see previous chapter). Thus, methods and tools are needed to assess the visual demands of the primary driving task and/or those demands imposed upon the driver by tasks performed additionally while the vehicle is in motion. As a method for HMI assessment of IVIS the occlusion technique is applicable for both assessing the visual demands of driving and simulating the interruption caused by doing some other task while driving (see the section on time sharing in the previous chapter). The occlusion technique is defined as a *"measurement method involving periodic/intermittent physical obscuration of the participant's vision or the obscuration of visual information under investigation"* (ISO/ TC22/ SC13/ N763R). The occlusion technique is based on the systematic control of the permitted time intervals for a subject to view at an object or perform a task.

In general, applications of the occlusion technique are well established in research on driving behaviour. A frequently quoted study was published by Senders, Kristofferson, Levison, Dietrich and Ward (1967), who introduced the occlusion technique as a procedure to quantify the visual demands of the primary driving task. Major progress was made by the development of so-called PLATO spectacles (Milgram & van der Horst, 1986), which made it technically very easy to control occlusion and non-occlusion sequences and eliminated interfering effects of readapting the eye after the occlusion interval. More recently the occlusion method was implemented on PCs in order to be able to use it in a laboratory context without the need of additional devices (e.g. Keinath, Baumann, Gelau, Bengler & Krems, 2001). A more detailed technical description will be given in the next paragraph. A review of the development and usage of the method in other areas than HMI assessment can be found in van der Horst (2004).

4.1 Description of the method

Studies which explored the occlusion technique as a tool for the assessment of the in-vehicle HMI investigated the application in the two different areas: First, there is evidence that the occlusion techniques reliably discriminates between visual displays of different complexity (e.g. Gelau et al., 1999; Keinath et al., 2001). This means that it can provide valid results when purely visual tasks have to be assessed. With the second area which has already been mentioned in the introduction also visualmanual tasks are addressed. For these tasks, which are perhaps more representative for the majority of in-vehicle tasks, the occlusion technique is applied to assess them under the aspect of their interruptability (e.g. Keinath et al., 2004; Krems et al., 2004) or chunkability (e.g. Noy et al., 2004; Stevens et al., 2004). The remainder of this chapter will focus on the latter area.

As already mentioned there are basically two different methods for obtaining the systematic control of the time intervals permitted for vision. The majority of published research studies have used goggles or spectacles to achieve the occluded intervals. Another means of occlusion in a number of studies is a PC that has periodic screen blanking (see Stevens et al., 2004 for a review). There is evidence which suggests that PLATO goggles or screen blanking methods are equally preferred by subjects (Weir et al, 2003). However, the PLATO goggles provide a more realistic environment for assessing the visual workload of IVIS as the driver is unable to view the vehicle display or IVIS controls whilst focusing on the road ahead. When using the screen blanking method, subjects are still able to view the touch buttons, whereas during real driving both the screen and the manual controls are not visible when the driver is viewing the road scene (Stevens et al., 2004).

The logic behind the application of the occlusion technique to assess in-vehicle tasks with respect to their interruptability is simple and straightforward. The aforementioned "obscuration of vision" when performing an HMI task is applied, in order to simulate the interruption caused by the driver's view back to the road scene during conditions of real driving. Interruptability means the degree to which subject performance on the HMI task suffers from these obscuration of vision. Basically this is determined by comparing subject performance on the task under unoccluded conditions with the performance under conditions of occlusion. In practice this means that the following parameters are calculated from the data of an occlusion experiment:

TTT (total task time): The time it takes to perform the HMI task under investigation, under unoccluded conditions.
OCCLT (occlusion time): The time for which the scene is occluded during the trial where the task has to be performed. This can be a constant value, such as 3 sec; or it can follow a certain distribution,
like a normal-distribution with AM= 2 sec and SD = .3 sec.
INSPT (inspection time): The viewing time where the shutter glasses are open during the trial where the task has to be performed. Values between 1.5 and 2 sec are often used. Also INSPT can be generated by means of a distribution.
TSOT (total shutter open time): The total time for which the scene is visible. If goggles are used this is the sum of all sequences where the goggles are open.
TSCT (total shutter closed time): The total time for which the scene is not visible. If goggles are used this is the sum of all sequences where the goggles are closed.

OCCLT and INSPT are the essential parameters of the occlusion method. However, for the aforementioned comparison of task performance under occluded and unoccluded conditions the index *R* is calculated. This index is defined as follows (c.f. ISO/ TC22/ SC13/ N763R):
R_{jk} = Mean $TSOT_{jk}$ / Mean TTT_{jk} for the j_{th} subject an k_{th} task.

From this ratio it can be derived that task performance suffers from the interruption caused by the obscuration of vision when R is greater than 1, i.e. time of vision required under conditions of occlusion is greater than under unoccluded conditions. Up to now no "critical value" of R has been defined which could be applied for a decision if a certain HMI task is "sufficiently interruptable" in order to not to interfere with the primary task of driving. The definition of such a value needs to be justified by research. Thus, the next paragraph gives a brief review on recent studies on the validity and reliability of the occlusion method.

4.2 Research on the validity and reliability of the occlusion technique

A set of studies was performed in recent years to assess occlusion technique and to show experimental validation of the method. The studies had to prove whether occlusion is able to distinguish in-vehicle dialog concepts which are tolerable concerning visual demands and chunkability, from those which are not (Krems et al., 2004).

In a study (ibid.) reading tasks of different complexity were presented to the subjects. In the simple task the subjects had to find a route from city A to city B on a map, whereas the shortest route had to be identified in the difficult task. The presentation time of the map varied in eight steps between 0.2 sec. and 1.2 sec. Results show that the probability of error was higher for the complex version across all times of the presentation. Values started to converge at a vision interval of 1.2 sec. In a second part of the study, the stimuli were subject-paced. Similar results were found in this case compared to those of the first part.

To test occlusion with regard of interuptibility of dialogs subjects were required to find a given name in simplified phone directories displayed on the PC. Subjects were interrupted at specific steps of the dialog and were given time as long as necessary to identify the correct entry. Vision interval was manipulated (0.6, 0.9 and 1.2 sec.). Furthermore, task complexity was varied (predictable place; reordered list). The number of errors was measured. It was found that probability of errors decreased with rising vision interval and was lower for the simple task.

Another study on the effects of task interruption was carried out where subjects had to find the phone number in a displayed text of 5 lines. The 'total task time' to solve the problem and 'task errors' were measured for a dual-task situation (with secondary task) and a single task situation (without secondary task). The total task time was longest for the uninterrupted situation, but no difference was found for errors. It was concluded that time of occlusion was also used for task completion. Furthermore, riteria which are only based on uninterrupted total task time (e.g. 15-second rule as proposed by Green, 1999) seem to be not sufficient for the evaluation of real driving situations."
Summarizing the experiments it was concluded that:

- occlusion technique was able to discriminate between displays and dialogs of different complexity,
- occlusion technique was able to discriminate between different conditions of task resumption and to show which of the given conditions affected an additional secondary task,
- The occlusion technique can be rated as a method that is able to evaluate the HMI design of IVIS with respect to their suitability while driving.

Whereas the validity and sensitivity of the occlusion technique has been demonstrated in numerous experiments, there is only rare evidence on its reliability. This is a clear gap in research since reliability can be interpreted as a precondition of validity. Moreover the reliability of the occlusion technique has recently been questioned in the discussions in ISO/TC 22/SC 13/WG 8. To fill this gap a project has been started by BASt where data from four occlusion experiments performed within BASt projects are re-analysed under the aspect of reliability. Results will be available by the end of 2004.

3. Information on "Car Clinics", Excerpt from RESPONSE3 deliverable D11.1 (V1.2):Draft of a Code of Practice for the Design and Evaluation of ADAS [7].

Car Clinic with Naive Subjects

The term "clinic" results from the specific investigation design: the market research location is not at the point of sale or at home, but "stationary" due to the fact that subjects are usually invited to a special test location ("clinic").
In a car clinic typically a sample of drivers will experience and assess a new developed product before market introduction.
There are several types of car clinics. For the validation of ADAS HMI and a proof of controllability a dynamic car clinic (driving clinic) is necessary.
The dynamic car clinic as a tool allows testing of driver behaviour and performance while using an ADAS system in defined situations in a realistic surrounding. It can be tested if a certain number of naive drivers are able to cope with ADAS-assisted driving situations including system failures and limitations, in a way that no accident occurs during a realistic test scenario.
In principle the car clinic can be carried out on public roads (= highest validity) or on test tracks (= better reliability and safety).
Experimental design, type of testing and test scenarios have to be defined. This includes

1. A decision for test environment

Both environments have advantages and disadvantages:

- On-the-road-trials

+ realistic environment
+ normal driving behaviour
+ high validity of results
− difficult to create and instrument
− costly and time consuming
− not fully controlled
− potentially less safe, provoking of risk situations and failure states (fault injection) is often too unsafe

- Test track:

+ Allows testing of (simulated) safety critical situations
+ Fault injection possible
+ Control of influencing variables (Reliability)
+ Artificial situation
+ No routine driving behaviour
+ difficult to create and instrument
+ expensive particularly if traffic conditions are to be simulated
+ safer than on-the-road trials, although an element of risk is still involved

2. Data to be collected

Depending on the ADAS function and the related test design different measures are necessary, e.g.:

- Speed measures
- Headway measures
- Steering control measures
- Lateral control measures
- Physiological measures
- Gaze measures
- Handling errors
- Observer ratings
- Self reports

3. Test scenarios to be used

Different test scenarios may be considered and compared during controllability assessment process.
Test scenarios might include

- type of traffic, for example light or heavy
- road type, for example urban, motorway or rural
- Weather conditions
- Road conditions
- Specific hazard situations

4. The experimental design

The design of the test trials needs to be carefully considered e.g. choosing

- Independent or dependent measures design
- repeated measures or longitudinal design
- with or without control group / Cross-over-design
- with or without base line
- Short term or long term testing / fleet testing
- Sample selection and sample size

5. The test procedure to be used

The test procedure needs to be detailed and listed e.g. in a trial timetable. The following topics might be relevant:

- Is a familiarisation drive required so that the participant can get used to the vehicle and ADAS?
- Which instructions have to be given, when and in which order.
- Is a randomisation/permutation of test scenarios in the run of events meaningful to avoid sequence effects?

6. The statistical analysis to be used

For each measure, the type of value should be specified, e.g. whether a mean or overall score, minimum or maximum will be taken. It has to be specified how to analyse this data and which statistical analysis is used.

7. A health and safety risk analysis

A health and safety risk analysis of testing shall ensure that participants are not at any risk. Areas of risks have to be identified and possible back-up plans, or countermeasures should be detailed. Legal and liability issues of testing should be regarded, e.g. the need for special insurances (participants, instructors, vehicles), and contracts.

A) Input/Requirements

As described above a

- Detailed and extensive planning by validation experts,
- Prototype equipped with ADAS system and data recorder
- A test track as the context requires
- List of ADAS-assisted driving situations including system failures and limitations, which are to be tested is needed.

B) Costs

- Time consuming (3 weeks to 3 months - depends on necessary sample size, number of test situations, duration of single test ride)
- high personnel effort (instructor, observer, various collaborators)
- high monetary effort (test track, participant fees, data recording equipment)

C) Output/Results

- Most valid results concerning driver's
    - Psycho-motor performance (reaction time, forces, etc.)
    - Mental capabilities (comprehension, learnability)
- Most valid results regarding driver's ability to control an ADAS system in assisted driving situations including system failures and limitations.

D) References

- Becker, S., Cieler, S., Johanning, T. & Schatz, A. (2000). Experimental Assessment of Driver Assistance Systems: Method (RESPONSE Project TR4022, Del. No. D5.1.1). Brussels: Commission of the European Communities, DG XIII.
- Johansson, E., Engström, E., Cherri, C., Nodari, E., Toffetti, A., Schindhelm, A., and Gelau, A. (2005): Review of existing techniques and metrics for IVIS and ADAS assessment. Del 2.2.1, Project AIDE - IST-1-507674-IP.
- Östlund, J. et al. (2004): Deliverable 2 – HMI and Safety-Related Driver Performance. Human Machine Interface And the Safety of Traffic in Europe. Project HASTE GRD1/2000/25361 S12.319626.
- Schuh, C. (1991): Die Car Clinic als Marktforschungsinstrument einer konsumentenorientierten Produktentwicklung, Köln 1991.
- Wiethoff, M. (2003). ADVISORS Final Report Annexes. EU project ADVISORS, project no. GRD1/2000/10047.