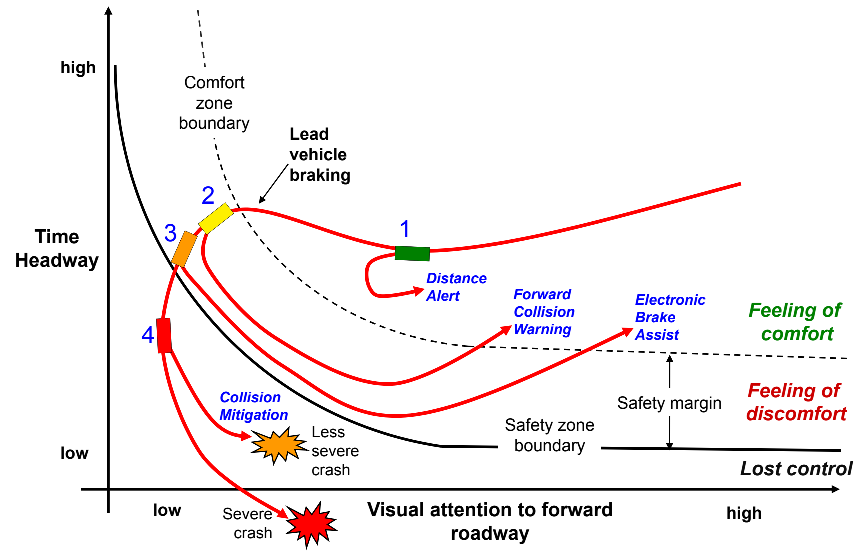
CHALMERS UNIVERSITY OF TECHNOLOGY
SE-412 96 Gothenburg, Sweden
Telephone: +46-(0)31 772 10 00
www.chalmers.se

CHALMERS

MIKAEL LJUNG AUST Improving the Evaluation Process for Active Safety Functions 2012

CHALMERS




# Improving the Evaluation Process for Active Safety Functions

## Addressing Key Challenges in Functional Formative Evaluation of Advanced Driver Assistance Systems

MIKAEL LJUNG AUST

*Department of Applied Mechanics*
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2012

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

# Improving the Evaluation Process for Active Safety Functions

## Addressing Key Challenges in Functional, Formative Evaluation of Advanced Driver Assistance Systems

MIKAEL LJUNG AUST

Department of Applied Mechanics

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2012

Improving The Evaluation Process For Active Safety Functions
Addressing Key Challenges in Functional, Formative Evaluation of Advanced Driver Assistance Systems
MIKAEL LJUNG AUST
ISBN 978-91-7385- 663-8
Ny serie nr: 3344



Doktorsavhandlingar vid Chalmers tekniska högskola
ISSN 0346-718X

Department of Applied Mechanics
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover:
Illustration of how loss of control in driving, and the intended influence of active safety functions aimed at preventing such loss, can be conceptualized as trajectories within a space defined by a set of Driver-, Vehicle- and Environment parameters relevant for controlling the sub-tasks involved in driving, such as visual attention, steering, braking, lane positioning, navigating, etc.

Printed at Chalmers Reproservice
Gothenburg, Sweden 2012

IMPROVING THE EVALUATION PROCESS FOR ACTIVE SAFETY FUNCTIONS
Addressing Key Challenges in Functional, Formative Evaluation of Advanced Driver Assistance Systems

MIKAEL LJUNG AUST
Department of Applied Mechanics
Chalmers University of Technology

**ABSTRACT**
The general aim of the present thesis was to improve key steps in the procedure for functional, formative evaluation of Advanced Driver Assistance Systems (ADAS). Five unresolved theoretical and empirical issues were identified and addressed. The first identified issue was the lack of a general conceptual framework for ADAS evaluation that can help formulate functional specifications and generate testable hypotheses on ADAS influence in critical driving scenarios. In response, a conceptual framework called Situational Control was developed. The second issue concerned the current ways in which crash data is used to specify ADAS evaluation scenarios. An improved methodology for linking a set of in-depth investigated case studies to a general crash type was developed and successfully tested. The third issue concerned the extent to which data from in-depth investigations of fatal crashes can be used to specify ADAS evaluation scenarios. Some countries have fully representative in-depth investigated datasets for this crash type, but their relevance for ADAS evaluation has not been investigated. An empirical study of causation information in fatal intersection crashes was performed. However, the information collected in these investigations was found to be limited in ways which made them less useful for defining ADAS evaluation scenarios. The fourth issue was whether sufficiently critical driving events that result in realistic driver responses can be created and repeated in driving simulator based ADAS evaluation. A study was performed in which two groups of drivers, one with and one without FCW, were exposed to repeated critical lead vehicle braking events. Results indicate that while creating a single surprise event is possible, interaction effects that compromise result generalizability occur when the critical event is repeated. The fifth issue concerned principles for how to assess the combined influence multiple ADAS when present in the same vehicle. A study of an FOT evaluated ADAS bundle consisting of FCW and ACC was carried out to empirically test whether existing conceptual models for calculating the combined effect of multiple safety functions were applicable. The results indicate that existing models were too simplistic to account for the complex modifications of driver behavior found in the data.

**Keywords: Advanced Driver Assistance Systems, Active Safety Systems, Safety Benefit Assessment, Driver Behavior Analysis, Crash Causation Analysis, Situational Control, In-depth Crash Investigation**

# ACKNOWLEDGEMENTS

The thesis work has been carried out as part of the project Factors Influencing the Causation of Accidents and Incidents II (FICA II). Project partners are Volvo Car Corporation, Autoliv Development AB, Saab Automobile AB, Volvo Technology, Volvo AB, and Chalmers University of Technology. Funding also comes from the Swedish national research program Fordonsforskningsprogrammet (FFP) under the Program Board for Automotive Research (PFF) within the Swedish Agency for Innovation Systems (VINNOVA).

I would like to thank my academic supervisor, Associated Professor Hans-Erik Pettersson at VTI/Chalmers University of Technology for great discussions and scientific advice. His ability to unerringly spot logical and methodological flaws has been invaluable for this work. I would also like to thank my examiner Professor Per Lövsund at Chalmers University of Technology for his great advice on the general structure of the work.

Much gratitude extends to Johan Engström at Volvo Technology/Chalmers University of Technology for a very fruitful collaboration in the FICA II project. Several key ideas in the thesis were worked out together with Johan, a brilliant co-wrestler when it comes to pinning down elusive concepts.

My colleagues at Volvo Cars and Chalmers have provided great support and good discussions on how to attack the problems addressed in this thesis, as well as an increasingly benevolent attitude towards the art of making good espresso.

Last but not least, thank you Karin for your love, patience, patience, then again some patience and encouragement, and Kersti and Konrad for providing me with a first hand experience of how a conceptual framework for the world is created. Dää!

# LIST OF PUBLICATIONS

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

**Paper I**
Ljung Aust, M. and Engström, J.(2011) "A Conceptual Framework For Requirement Specification And Evaluation Of Active Safety Functions", Theoretical Issues in Ergonomics Science, 12: 1, 44 - 65, First published on 15 June 2010 (iFirst).

Contribution: The basic ideas behind the proposed framework were jointly worked out by Ljung Aust and Engström. Ljung Aust authored the paper.

Paper **II**
Ljung Aust, M. (2010). "Generalization Of Case Studies In Road Traffic When Defining Pre-Crash Scenarios For Active Safety Function Evaluation". Accident Analysis & Prevention, 42, 1172-1183.

Contribution: The study was designed by Ljung Aust. Ljung Aust carried out all data reduction, performed all analyses, and authored the paper.

**Paper III**
Ljung Aust, M., Fagerlind, H., and Sagberg, F. (2012) "Fatal Intersection Crashes In Norway: Patterns In Contributing Factors And Data Collection Challenges". Accident Analysis and Prevention, 45, 782-791. doi: 10.1016/j.aap.2011.11.001.

Contribution: The study was designed by Ljung Aust, Fagerlind, and Sagberg. Sagberg extracted the relevant crash data. Ljung Aust carried out the data reduction, performed all analyses, and authored the paper.

**Paper IV**
Ljung Aust, M., Engström, J. and Viström, M. (submitted) "Effects Of Forward Collision Warning And Repeated Event Exposure On Emergency Braking". Submitted to Transportation Research: Part F.

Contribution: The study was designed by Ljung Aust, Engström, and Viström. Ljung Aust carried out the data reduction, conducted all statistical analyses, and authored the paper.

**Paper V**
Ljung Aust, M. and Dombrovskis, S. (submitted) "How To Estimate The Combined Safety Effect Of Multiple Advanced Driver Assistance Systems Deployed In The Same Vehicle During Field Operational Tests". To be submitted to Accident Analysis & Prevention.

Contribution: The study was designed by Ljung Aust. Dombrovskis extracted the relevant data. Ljung Aust carried out the data reduction, conducted all statistical analyses, and authored the paper.

# Table of Contents

## Project acronyms

| | |
|---|---|
| APROSYS | Advanced Protection Systems |
| ASSESS | Assessment of Integrated Vehicle Safety Systems for Improved Vehicle Safety |
| CAMP | Crash Avoidance Metrics Partnership |
| CICAS | Cooperative Intersection Collision Avoidance Systems |
| eIMPACT | Socio-economic Impact Assessment of Stand-alone and Co-operative Intelligent Vehicle Safety Systems in Europe |
| euroFOT | European Large-Scale Field Operational Tests on In-Vehicle Systems |
| HASTE | Human Machine Interaction and the Safety of Traffic in Europe |
| INTERACTIVE | Accident avoidance by active intervention for intelligent vehicles |
| IVBSS | Intelligent Vehicle Based Safety Systems |
| PReVAL | PReVENT Impact Assessment |
| RDCW | Road Departure Crash Warning System Field Operational Test |
| RESPONSE 3 | Code of Practice for ADAS development and testing |
| SafeTE | Safety Evaluation Of In-Vehicle Information Systems |
| SeMiFOT | Sweden-Michigan Field Operational Test |
| TRACE | Traffic Accident Causation in Europe |

## Abbreviations

| | |
|---|---|
| ACC | Adaptive Cruise Control |
| ADAS | Advanced Driving Assistance Systems |
| BRT | Brake Response Time |
| DVE | Driver-Vehicle-Environment |
| EC | European Commission |
| FCW | Forward Collision Warning System |
| FOT | Field Operational Test |
| FRAM | Functional Resonance Accident Model |
| LDW | Lane Departure Warning system |
| LKA | Lane Keeping Aid |
| STAMP | Systems-Theoretic Accident Model and Processes |
| TTC | Time-To-Collision |

# 1 Introduction

Traffic safety has long been, and continues to be, an important issue to address. For example, in 2006, there were 1.3 million road accidents involving personal injury in the EU-27, and 42 950 persons lost their lives (Eurostat, 2009).

Over time, numerous routes toward crash prevention have been invented and tried out. These include improved road design (e.g. clear zones), public awareness campaigns (e.g. don't drink and drive) and deployment of in-vehicle systems that protect drivers from injury when a crash is inevitable (seat belts, air bags, rollover protection systems, etc.).

A relatively novel approach towards accident prevention and injury reduction is the introduction of vehicle based active safety functions, also known as Advanced Driver Assistance Systems or ADAS. In contrast to protective, or passive, in-vehicle safety functions whose purpose is to mitigate crash consequences, the general goal of ADAS is to prevent crashes from occurring at all. ADAS are meant to achieve their preventive effect either by alerting the driver to potential hazards and/or by taking over the driving task to some extent, using e.g. autonomous braking in emergency situations.

An ADAS typically consist of one or more environment sensors mounted on the vehicle, e.g. radars and/or cameras, a software that based on the sensor input determines what actions the ADAS should take, and a driver and/or vehicle interface that can be used to alert the driver and/or control the vehicle. Examples of safety technologies which fall under the ADAS umbrella are FCW, ACC, LDW, and Drowsiness Warning. Several ADAS, including those mentioned, have recently been deployed on the market, and more are close to introduction.

A key issue for ADAS, like for all other means of crash and injury prevention, is to verify that they are beneficial, i.e. that they actually improve traffic safety. While the safety potential of ADAS can be affected by many factors, Carsten and Nilsson (2001) proposed that all safety implications can be classified as belonging to either of three general aspects: the function safety aspect (technical reliability of the system), the Human Machine Interaction aspect (operating, and communicating with, the system), and the traffic safety aspect (system influence on driving behavior, including changes in interactions with other road users). For a complete ADAS evaluation, verification must take place at all three levels.

All three aspects present interesting study topics. In the present work however, the scope is limited to the traffic safety aspect only. Sensor and algorithm verification, as well as interaction design issues, are left aside. Instead, the focus is on improving available procedures for assessing to what extent an ADAS influences driver behavior in relevant, often critical, driving situations, where the goal is to be able to robustly identify any ADAS driven driver behavior changes that may influence traffic safety in a positive or negative way.

## 1.1 Formalizing the formative ADAS evaluation procedure

Assessing an ADAS influence on driver behavior in particular or general traffic situations and deducing safety implications of potential behavioral changes is often referred to as functional ADAS evaluation. Functional evaluation can take place either prior to, or some time after, a function has been released into production vehicles. This is referred to as formative and summative evaluation respectively, a terminology coined by Scriven (1967). In the present context, formative evaluation thus relates primarily to laboratory based evaluation of an ADAS during its development stage, while summative evaluation refers to an evaluation of the function's performance in the real world after mass deployment in the vehicle fleet has started.

In terms of summative ADAS evaluation, there is unfortunately still a lack of relevant data. Since these functions only recently have reached the market, and because ADAS tend to be sold as options rather than standard features even on premium cars, it is yet too early to identify a sufficiently large ADAS equipped vehicle population in publicly available crash databases that can be compared to a similar, non ADAS equipped, fleet. This sets a further delimitation for the present work, i.e. this thesis discusses only formative ADAS evaluation, not summative evaluation.

The fact that for any new ADAS it will always take quite some time (often years) before sufficiently many people have bought and used the system to allow for summative evaluation has generated a lot of interest in the research community in terms of formative evaluation. Many wish to see these new and promising technologies widely deployed, but also recognize the need to try to make sure that one is not deploying something potentially harmful, i.e. that turns out to have a negative impact on traffic safety. Thus, even if the exact traffic safety impact of an ADAS only might be possible to assess in retrospect (Carsten and Nilsson, 2001), numerous large scale projects, both in Europe and the US, have tried to come up with evaluation procedures that will allow for reasonably robust prediction of which benefit(s) any given ADAS may add or subtract from the traffic system.

In the US, one of the earliest efforts was the CAMP project (Deering 2004; Shulman and Deering 2004; Shulman and Deering 2005), initiated by Ford and General Motors in 1995 to accelerate implementation of crash avoidance countermeasures in passenger vehicles. CAMP consisted of many sub-projects, but of particular relevance here is the FCW alert timing project (Kiefer *et al.* 2003; Kiefer *et al.* 2005). Other projects include several FOTs, such as RDCW (LeBlanc *et al.* 2006*a*), IVBSS (Sayer *et al.* 2011), and ACAS (Najm *et al.* 2006). Other approaches include Gordon *et al.* (2010), who developed a simulation based approach to ADAS evaluation, and CICAS (Maile *et al.* 2008) where a scenario based evaluation approach for collision avoidance systems in intersections was devised.

In Europe, projects aimed towards various aspects of formative ADAS evaluation procedures include HASTE (Carsten *et al.* 2005), AIDE (Kussmann *et al.* 2005), ASSESS (Dobberstein *et al.* 2010), INTERACTIVE (Bakri *et al.* 2011) APROSYS (Eggers and De Lange 2007), TRACE (Karabatsou *et al.* 2006), PReVAL (Scholliers *et al.* 2008b), SeMiFOT (SeMiFOT, 2009), eIMPACT (Assing *et al.* 2006), RESPONSE 3 (Knapp and Schwarz 2006), CONVERGE (Maltby *et al.* 1998; Zhang *et al.* 1998), and SafeTE (Engström and Mårdh 2007).

There has also been several attempts at standardization of ADAS evaluation procedures in Europe and the US, including The European Statement of Principles (Commission 2008), the Alliance of Automobile Manufacturers' guidelines (2003) and the checklist by Stevens *et al.* (1999).

There are many ways to describe formative ADAS evaluation, but a common approach in many of the above projects is to create a formalized flowchart that details the required steps of formative ADAS evaluation. Examples include CONVERGE (Maltby *et al.* 1998; Zhang *et al.* 1998), APROSYS (Eggers and De Lange 2007), eIMPACT (Assing *et al.* 2006), AIDE (Janssen *et al.* 2008), ACAT (Carter *et al.* 2009), and PReVAL (Scholliers *et al.* 2008*b*). Of these, one that integrated previous approaches, and which later projects often refer to, is the flowchart from PReVAL. PReVAL was a subproject of PReVENT, a large EC- funded effort to bring various ADAS under development closer to market release. The goal of PReVAL was to provide a harmonized evaluation and impact assessment methodology for the ADAS developed in PReVENT, and apply the methodology to a set of given use cases.

The PReVAL flowchart is based on CONVERGE (Maltby *et al.* 1998; Zhang *et al.* 1998) and the experiences of APROSYS (Eggers and De Lange 2007) and AIDE (Janssen *et al.* 2008). It is organized similarly to the three aspects identified by Carsten and Nilsson (2001), i.e. there is technical evaluation, human factors evaluation and impact assessment, though the human factors evaluation is more focused on functional assessment than interaction design. Technical evaluation is performed in two phases: firstly the verification to test the individual components and subsystems towards the technical specifications, and secondly the validation to test whether the goals and specifications of the complete system are met. Human factors evaluation assesses the extent to which the ADAS succeeds in generating the intended behavioral responses from the driver in the relevant critical situations. Impact assessment aims to make an aggregate-level assessment of the ADAS effects on relevant harm metrics, usually the number of fatalities, in target situations, based on the technical performance and behavioral impact of the ADAS (Scholliers *et al.* 2007*a*). The flowchart is based on the so called V design cycle (Scholliers *et al.* 2008*a*), commonly used in the automotive industry, and which here is extended by including the different steps of the evaluation process, see Figure 1.

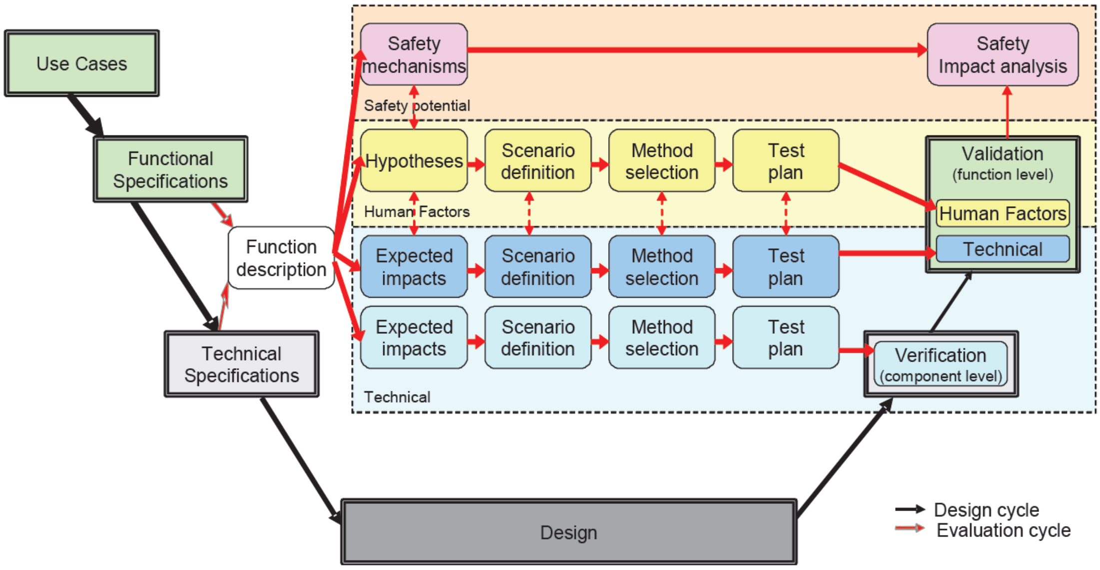



Figure 1. The PReVAL ADAS Evaluation Flowchart

The PReVAL flowchart for ADAS evaluation specifies six steps which are common for both the technical and human factors evaluation (Scholliers *et al.* 2007*b*):

1. ***Functional specification:*** *At the start of an evaluation, a sufficiently detailed functional description needs to be available, that is common for all assessments and consistent with functional specifications for other ADAS under evaluation.*

2. ***Expected impact****: Describing the objectives of the ADAS in such a way that it is possible to evaluate its performance, i.e. generate hypotheses on how the ADAS is expected to change driver behavior in relevant critical situations.*

3. ***Scenario definition****: To test the generated hypotheses, test scenarios need to be defined.*

4. ***Method selection****: Selection of the method(s) through which the test scenarios should be run, typically test track, driving simulator, or field study.*

5. ***Test plan****: Specifies the number of tests, independent and dependent variables, and actual measurements. It also describes which experimental design to use including the number of subjects, and covers all other details required to acquire statistically significant results in the hypotheses testing.*

6. ***Test execution, analysis, and reporting****: Conducting tests, analyzing results, and drawing conclusions.*

While this flowchart gives a well structured overview of the steps of an evaluation procedure, ADAS evaluation can still be said to suffer from intrinsic difficulties and a lack of harmonization (Scholliers et al. 2007a). To understand this statement, it is helpful to compare with the evaluation of in-vehicle injury protection systems like airbags and seatbelts. In this domain, the procedure which has emerged as the gold standard for testing of new functions is called a load case. The term load is to be taken very literally, i.e. a load case describes a specific test condition where vehicle and driver/passenger surrogates, i.e. crash test dummies, are exposed to a specific physical loading, through e.g. crashing the vehicle head on into a barrier or rolling it over. Measurements of how the physical loading effects the dummies and the vehicle are then used to predict the real world performance of vehicles built to the tested specification in certain crash types (Korner 1989).

The development and validation of load case procedures for evaluating in-vehicle injury protection systems has been ongoing for quite some time, and the test specifications in use today are both standardized and specified at a high level of detail regarding scenario definition, performance metrics, and pass/fail criteria, see e.g. EuroNCAP (2004; 2008; 2009). Furthermore, the laboratory processes for doing full scale and component crash testing have iteratively been refined to the point where they actually can be certified as accredited laboratories, fulfilling standards such as AIC (2010).

ADAS evaluation procedures have yet to reach this level of detail and sophistication. The projects mentioned above have all helped improve on the details of an ADAS evaluation process, but there is still no general agreement, let alone standardization, on what the evaluation procedure should look like other than on a quite abstract level such as the PReVAL flowchart.

## 1.2 General goal and structure of the thesis

The present thesis is mainly a methodology thesis, focused on issues related to ADAS evaluation. The general goal is to improve the general process for functional, formative evaluation of ADAS. The six steps of the PReVAL procedure provide a general guideline for the work., i.e. the general idea is to achieve the goal by addressing a subset of the intrinsic difficulties associated with implementing each of the six steps in the PReVAL evaluation procedure.

The thesis includes both theoretical as well as empirical work. The theoretical work addresses an unresolved issue in step 1 and 2, i.e. functional specification and expected impact. It can be argued that a prerequisite for formulating relevant functional descriptions as well as hypothesis on ADAS effects is a conceptual framework that describes the role of the driver and the role of the ADAS in the context of resolving critical driving situations. For reasons to be further reviewed below, such a framework is yet to be clearly formulated. In Section 2, criteria for, and required components of, such a framework are set up, and possible component contenders are reviewed in light of ADAS evaluation. In Paper I, a framework called Situational Control is formulated that meet the defined criteria.

The empirical work address four unresolved issues in steps 3-6. Step 3, scenario definition, is about defining relevant test scenarios for the ADAS evaluation. A data source commonly used in such work is real world crash data, since it provides a real world foundation for the test scenarios. However, for scenario descriptions to be both representative and sufficiently detailed in terms of characterizing the crash causation mechanisms an ADAS needs to neutralize to be effective, the extraction process has to solve either the generalization or the case selection problem.

The generalization problem is the question of how to determine whether findings from a particular set of crashes are representative of all crashes of that type. This is a typical problem facing in-depth investigation projects, where available resources usually only admit investigation of a limited crash set. In Section 3, the two current main approaches and their drawbacks are discussed, and in Paper II, a methodology intended to overcome those drawbacks is presented and empirically tested.

The case selection problem relates to the fact that in some countries, e.g. Sweden and Norway, there exists a data set that theoretically makes it possible to bypass the generalization problem, because the authorities commission in-depth investigations of all fatal crashes that occur. These countries thus have access to crash datasets which are both representative and detailed, and if they contain relevant information for understanding why the crashes occur, they would be very useful for defining ADAS evaluation scenarios. However, it is unclear whether these crashes are investigated in a way that supports extraction of relevant information for defining ADAS evaluation scenarios. The case selection problem is further discussed in Section 3, and in Paper III, an empirical study of a fatal crash data set is performed to assess whether the results can be useful for ADAS scenario definition. .

Steps 4 and 5, method selection and test plan, relate to how an ADAS evaluation is set up, both in terms of experimental design and choice of venue. Recent findings in naturalistic driving studies indicate that unexpectedness is a key property of real world critical events (e.g. Dingus *et al*, 2006). To properly evaluate an ADAS influence on driver behavior, one must therefore either replicate that real world state of genuine unexpectedness in the experimental environment, or show that responses from not so surprised drivers can be validly extrapolated to real world events. In Section 4, the challenges of creating and acquiring robust measurements of driver performance in ADAS relevant situations, i.e. surprise events, particularly in light of drivers inherent and quite sophisticated capacity to adapt, are discussed, along with the implications for method selection and test plan design. In Paper IV, a driving simulator study is performed that explores the extent to which critical, unexpected events can be created and repeated during FCW evaluation, and what the results imply for future ADAS evaluation.

Step 6, test execution, analysis and reporting, as indicated, concerns the analysis and reporting of evaluation results. Although old in traffic safety development in general, an issue that has recently come to the foreground in relation to ADAS evaluation is how to conduct analyses and reporting on ADAS effects when multiple ADAS are used simultaneously. Future production vehicles are likely to be equipped with bundles of ADAS rather than a single function. This should preferably be reflected in the design of ADAS evaluation studies, which means that one has to find a way of understanding and assessing the a single ADAS's influence on driver behavior when multiple ADAS are simultaneously present in the same vehicle. .

In Section 5, the challenges for identifying isolated effects of a particular ADAS when multiple ADAS are present in the vehicle are described for the case of FOT based ADAS evaluation, particularly in light of recent studies that indicate that an ADAS might influence driver behavior outside the particular scenarios where it is actively interacting with the driver.  In Paper V, an empirical study of this issue for an ACC and FCW bundle deployed in an FOT is performed.

After the problem descriptions and the literature review in Sections 2-5, the specific aims of the thesis are stated in Section 6. In Section 7, the thesis papers are summarized, and in Section 8, the strengths and weaknesses of the studies are discussed as well as more general topics. The thesis concludes with some suggestions for future work.

# 2 Functional specification and expected impacts - Understanding driver behavior and ADAS influence in critical driving situations

Haddon (1968) stated that scientific progress in general depends on, and is marked by, a transition from classifications based on descriptions of appearances to classifications based on fundamental processes. On a general level, a set of principles and concepts which capture the fundamental processes in a field of science can be called a conceptual framework (Hollnagel and Woods 2005). One way to approach the formulation of such a framework is to formulate it as a language problem. The development of a powerful descriptive language is fundamental to work and progress in any science. This language, or conceptual framework, should include a set of concepts or descriptive components that give an accessible yet powerful and scientifically valid description of the field of science it concerns, and in general fulfill three general criteria (Hollnagel and Woods 2005):

- *It must describe the important functional characteristics of human-machine systems.*
- *It must be applicable for specific purposes such as design, analysis and evaluation.*
- *It must allow a practically unambiguous use within a group of people, i.e. the scientists and practitioners who work broadly within the domain.*

In passive safety, work rests firmly on what can be called Haddon's negative energy transfer model. Injuries are understood as occurring when "energy is transferred in such ways and amounts, and at such rates, that inanimate or animate structures are damaged" (Haddon 1973). Following this model, countermeasures should focus on blocking and/or redistributing this unwanted energy in time and space so less or none of it reaches human beings in harmful doses. Haddon's energy transfer model is simple, clear, and powerful enough to be useful for mechanical engineers, biomechanical experts, behavioral scientists and/or laypersons alike, and they can communicate across their specialties using this model.

A conceptual framework for understanding crash causation and prevention that is as useful for ADAS evaluation as Haddon's is for passive safety function evaluation is thus still lacking. Active safety needs a scientifically correct conceptual framework that paints the big picture of how drivers can end up in critical driving situations, and how they would interact with ADAS in these situations. The framework should also be possible to grasp intuitively for those who work in the domain, but should also be able to specify further according to specific needs and purposes. Without such a framework, there is a clear risk of confusion and diverging definitions (Scholliers *et al.* 2007*b*).

To define a conceptual framework for ADAS evaluation, two of the general criteria from Hollnagel and Woods (2005) above can be specified more in detail. For the second criterion, which says that the framework must be applicable for specific purposes such as design, analysis and evaluation, a more specific formulation is that the conceptual framework should be applicable for generating testable hypothesis on which changes in driver behavior an ADAS is intended to generate in a critical driving situation.

Based on this, the first criteria, i.e. describing the important functional characteristics of human-machine systems, can also be given a more precise formulation. To support formulation of testable hypothesis for ADAS evaluation in critical driving situations, the conceptual framework must characterize 1) the nature of the critical situations, i.e. why they evolve; and 2) the driving task, in such a way that it becomes possible to understand why drivers sometimes fail and other times succeed in resolving these critical situations, and thus where ADAS may play a role.

Addressing the first point requires some form of accident modeling, and addressing the second point requires some form of driver model that is compatible with the accident modeling. Ultimately, the conceptual framework should merge these two into a coherent whole. In the next section, various approaches to accident modeling will be reviewed. In the section following that, approaches to driver modeling will be reviewed.

## 2.1 Accident modeling in relation to road traffic and ADAS evaluation

An accident model is an abstract conceptual representation of the occurrence and development of an accident. It describes how and why crashes happen, it defines possible causes and interactions, and it provides the basis for an associated countermeasure principle. Accident models are often implicit rather than explicit in the accident investigators' minds and guidelines, and therefore important to explicitly identify and reflect upon when discussing safety and risk control (Hollnagel 2004; Woltjer 2009).

Over the years, many accident models have been developed. On a general level, these can be divided into three categories: simple linear (sequential), complex linear (epidemiological), and systemic[1] (Hollnagel 2004). In simple linear accident models, socio-technical systems are described by their physical and organizational structure. Focus is on identifying linear cause-effect relationships between independent components, and accidents are viewed as the result of a single sequence of clearly

[1] Note that the concept systemic here refers to the nature of accidents, rather than the scope of the system in which the accident takes place. Taking a systemic view is sometimes used to indicate that all parts of a system are allowed targets for analysis and countermeasure development, as opposed to focusing on a single part or component. Here however, systemic implies that accident causation in a complex system is a question of unexpected interactions between control processes at their performance limits rather than defaulting parts, and these may occur anywhere and everywhere in the system.

identifiable component failures occurring in a specific event order, such as in Heinrich's description of an accident as a row of tumbling dominoes (Heinrich 1931).

Complex linear accident models also describe socio-technical systems by their physical and organizational structure and focus on identifying linear cause-effect relationships between components. However, the complex models look at relationships between interdependent components rather than trying to find a single chain of events. Accidents occur when certain combinations of latent conditions and active failures coincide in time and/or space in a way that overcomes existing defenses or barriers (Reason 1990). A classic example of a complex linear model is the so called Swiss cheese model (Reason 1997), which suggests that a complex system is analogous to a stack of slices of Swiss cheese. Each slice is an imperfect safety barrier in the system. While ideally a problem that passes through a hole in one layer will be stopped by the next slice because its holes are in different places, in some unpredicted instances the holes in multiple slices happen to be aligned, and an adverse event occurs. Another example, which relates directly to traffic safety, can be found in Donges (1999).

The third group, systemic accident models, focus on the natural performance variability associated with process control. In a systemic view, the purpose of a socio-technical system is to work towards relevant goals, such as reaching a destination safely or maintaining a continuous factory output. Accidents occur when one or more of the control processes used to achieve these goals coincidentally get out of hand, i.e. exceeds situation tolerance, and there is insufficient time and/or resources to regain control over them. Examples of systemic models are STAMP and FRAM (Hollnagel 2004; Woltjer 2009).

In terms of which of these modeling approaches have bearing on ADAS evaluation, one place to look for guidance is in-depth studies of crash causation. Looking at the traffic accidents analyzed in the FICA project (Ljung *et al.* 2007) and SafetyNET (SAFETYNET, 2008), it seems clear that linear accident modeling has to be ruled out. The crashes investigated in these projects were very rarely attributed to a single series of clearly identifiable events. Usually, a more complex process that involved multiple, interacting contributing factors was found.

As for deciding between complex linear models like Reason's Swiss cheese model and systemic accident models, a number of researchers have recently argued in favor of systemic models, at least when the complexity of the system where the accident occurs increases beyond a certain level (Amalberti 2001; Dekker 2005; Hollnagel 2004; Leveson 2004; Reason *et al.* 2006; Rochlin 1999). To summarize their basic argument, they say that in complex systems, the number of ways in which different factors can be combined to create an accident is practically infinite. Instead of identifying such combinations, the most promising approach to accident prevention is to try to identify reasons for, and means to reduce, process control variability (Hollnagel 2004; Rasmussen 1990).

While these authors work in many types of process control domains like aviation and software engineering, the argument has been explicitly extended to driving as well. According to Huang (2005; 2007), modern road traffic has enough of the complex and dynamic characteristics of a complex system to make systemic accident modeling the best suited approach in accounting for how and why failures occur, and to conceptualize countermeasures.

In the systemic view, loss of control is not considered an extraordinary event. It is rather a natural performance variability associated with an operator's degree of control over any process, which enables the operator to cope with complexity and uncertainty by adapting procedures and tools according to the situation. However, as Brehmer (1990) argues, there is an inherent variance in both peoples' perception and action which set a limit to how well they can adapt. It follows that loss of control will occur when natural performance variability pushes the level of control outside the tolerance limits of a given context. For any sufficiently complex system, accidents are therefore bound to occur sooner or later. This is sometimes referred to as Normal Accident Theory (Perrow 1984).

## 2.2 Driver modeling in relation to systemic accident modeling

According to Carsten (2007), two broad approaches to driver modeling can be distinguished in the literature. The first is descriptive modeling where one tries to describe parts of, or the whole, driving task in terms of what the driver has to do. The category can be further subdivided into several subcategories, i.e. task models, adaptive control models, and production models. Task models typically present the driving task as a set of hierarchically ordered subtasks. Examples include Allen *et al.* (1971) and Michon (1985). Adaptive control models describe driving as a process with inputs, outputs, and feedback, where driving is viewed as involving a continuous adaptation to a changing environment, in a way which promotes goal fulfillment (Engström and Hollnagel 2007). Examples include McRuer *et al.* (1977) and Hollnagel *et al.* (2003). The driving task can also be described as a set of rules, i.e. a production system. Michon (1985) e.g. described a formal set of rules for changing gear.

A possible fourth category to complement Carsten is resource theory models. They build on the assumption that humans have a finite amount of resources, and predict that driving performance degrades when two or more tasks demand the same resources, e.g. see Kahneman (1973). The dominant version today is Wickens' Multiple Resource Theory (MRT) see e.g. Wickens (2002; 2008). MRT has been applied particularly by researchers who study how drivers cope with multiple tasks over longer periods of time, such as in CAMP's Driver Behavior Metrics Program (Angell *et al.* 2006). MRT substituted general resource theory when substantial research showed more interference between structurally similar than structurally dissimilar tasks. In response, Wickens proposed that humans have multiple resources and interference will only be strong between tasks that claim the same resources.

However, descriptive models are generally analytical and lack predictive elements. It is therefore hard to conclude, based on descriptive models, how changes in driver motivation, capability, or decision making would affect driving performance or situational risk (Carsten 2007).

The second major type is motivational modeling, which describes how the driver manages risk or task difficulty in terms of utilities and various types of trade-off (Carsten 2007). Well known motivational models include Wilde's risk homeostasis model (Wilde 1982) and Näätanen and Summala's zero risk model (Näätänen and Summala, 1976). Other examples include attempts to characterize how drivers adapt, or fail to adapt, to road safety measures, based on perceived changes in risk (Adams 1995; Evans 1991; OECD 1990).

Some researchers have also tried to merge the two approaches. Hatakka *et al.* (2002) e.g. present a four level descriptive hierarchy model where goals for life and skills for living constitute the top layer and vehicle maneuvering constitutes the bottom layer. However, motivational models, while in theory more susceptible to testing and verification then descriptive models, have not in practice really been put to the test (Carsten 2007), and perhaps not even fully specified (Ranney 1994).

This might explain why, according to Elvik (2004), none of the theories of driver behavior proposed so far, whether motivational or descriptive, have been applied in road safety evaluation research or enjoy widespread approval among road safety evaluation researchers. The one exception might be MRT, which is used fairly often in human factors related research, such as the design of in-vehicle system interfaces.

Two things can be concluded from reviewing the literature so far. First, given that systemic accident modeling is the best approach to describe why critical situations occur, the conceptual framework needs to frame ADAS evaluation along those lines. In other words, if accidents occur due to unfortunate combinations of control variability in the processes involved in driving, then driving is best characterized as a form of process control, where accidents occur when control is lost. This perspective follows the general structure of the adaptive control models.

Furthermore, as Carsten (2007) noted, these control models describe how the multiple control processes involved in driving interact with each other but do not account for how goal states are selected. To understand why drivers choose the reference values for control that they do, the adaptive control perspective needs to be merged with an account of driver motivation. As stated by Eysenck (2004):

> *"Motivation is closely related to the direction behavior takes (which goal or goals are being pursued), the intensity of behavior (amount of effort, concentration, etc. invested in behavior) and persistence of behavior (the extent to which a goal is pursued until reached)".*

In general, two paradigms for describing human motivation exist. One is to define motivation as that which makes us strive for need fulfillment. The most well known theory in this paradigm is probably Maslow's hierarchical theory of needs, where the bottom level constitutes a need for survival and the top level a need for self-actualization (Maslow 1954; 1970).

The other paradigm defines motivation as that which drives action when an equilibrium is threatened. This view originated with the phenomenon of homeostasis, which is the tendency for an individual's internal environment, such as body heat, to remain fairly constant despite external changes. However, the homeostasis concept has been extended to cover psychological and social phenomena. For example, while people do not get hungry at exactly the same time every day, dinner time may represent a social schedule equilibrium and social norms for when dinner should take place, determined by work or other schedules, take precedence (Eysenck 2004).

The homeostasis approach is the one which has been most frequently applied in the domain of driving. Apart from Wilde's risk homeostasis theory, one early account is the zero risk theory by Näätänen and Summala (1976), which proposes that driver behavior is a balancing act between excitatory forces that push the driver to actively look for and exploit opportunities for action present in the environment, and inhibitory forces which keep the driver from acting on opportunities for action that might have regrettable consequences.

Originally, Näätänen and Summala proposed that inhibition is driven by experienced subjective risk. More recently, Vaa (2007) developed this idea by incorporating Damasio's (1994) concept of somatic markers. Somatic markers are emotional signals that attach positive or negative feelings to opportunities for action and their outcomes, based on previous outcomes from acting on them. Vaa argues that adaptive driver behavior largely is governed by somatic markers, especially in relation to threatening situations.

Following Vaa, Summala (2007) substituted the concept of subjective risk with discomfort, and argued that drivers actually strive to maintain a state of zero discomfort. Discomfort include feelings of immediate risk or threat, e.g. in a critical traffic situation, but can also be related to mobilization of effort to cope with task demands (Hockey 1997).

## 2.3 Beyond structure and motivation to expectancy and prediction - driving as proactive attention selection and action

An important dimension that is generally missing from both descriptive and motivational models, but which a long tradition of research has shown to be a key parameter of driver behavior, is an account of driver expectancies or prediction. This could be considered unfortunate, because expectancy has a large influence on how drivers respond in critical evaluation scenarios. For example, in a meta analysis of BRT from 39 different studies, Green (2000) found that BRT for expected events was on the magnitude of 0.6-0.75 s, while BRT for completely unexpected events were on the magnitude of 1.5 s and above. In other words, whether an event is expected or not might more than double the onset time of the driver response, something which clearly is important for ADAS design.

Also in a more general perspective, expectancies and prediction play a key role. As lately stated by Woods and Hollnagel (2006), but also by many before them, human action is in general oriented toward the future, in anticipation of, or as preparation for, what comes next. Driving conforms to this principle, i.e. that driving is proactive and driven top-down by current task goals and the driving context is evident in many studies across various domains and topics.

Note, however, that top-down control does not necessarily imply conscious control. As proposed by Trick and Enns (2009), a framework that can integrate current research on driver attention needs at least two dimensions; one that captures the distinction between automatic (non-aware) and controlled (aware) processes, and another dimension that reflects the origin of the process, i.e. whether the process is learned or a consequence of how humans are built (biological hardwiring). Combining these two dimensions lead to four possible modes of attention which they call reflex, habit, exploration, and deliberation. Habit and reflex represent automatic attention selection, i.e. done without conscious awareness; where habit is attention selection based on previous experience, i.e. top down driven; and reflex is attention selection based on biological hardwiring so to speak, such as braking when an animal jumps out on the road.

Examples of studies showing proactive and goal driven behavior in drivers include the studies by Cnossen *et al.* (2000; 2004), who found that drivers give more priority to tasks that serve the goals of the driving task itself. Their test participants put more effort into extracting navigation information, presented either as auditory information interspersed in a radio program or visually by means of a map, than on performing a concurrent artificial memory task.

In another study, Shinoda *et al.* (2001) found that drivers were more likely to notice stop signs located at intersections as opposed to signs in the middle of a block. They also showed that subjects spent much more time looking for signs at the intersections when asked to drive on their own and follow normal traffic rules compared to when asked to follow a lead vehicle. Such contextual selection of visual information is underscored in two reviews of theories of gaze allocation, (Hayhoe and Rothkopf 2011; Tatler *et al.* 2011), where it is clearly stated that behavioral goals is a critical factor in controlling the acquisition of visual information from the world, i.e. attention selection is proactive and goal driven.

In two studies of car-to-bicycle conflicts and accidents, Räsänen and Summala (1998), and Summala *et al.* (1996), they found that the most common conflict pattern was between drivers turning right and cyclists coming from the right on cycle paths at non-signalized intersections. Their explanation of this somewhat counterintuitive finding (if you are turning right, you should be looking right, right?), is that drivers who are about to turn right mainly focus their attention on cars coming from the left, since cars coming from the right pose no threat in right hand traffic. They therefore fail to see the cyclist coming from the right.

Martens and Micah (2007) showed that when test participants were driving the same route repeatedly, the participants' glances towards traffic signs became shorter at the same time as recollection of signs encountered along the route improved. Moreover, upon the last drive an intersection was changed so drivers had to yield, and this change went undetected by all participants although two responded correctly after crossing the yield markings. Martens (2011) also showed that the likelihood of detecting sign manipulation at the end of a series of repeated drives depends on several factors. Detection improved with a larger difference between the original and the changed sign, with the new traffic sign not fitting the traffic scene, and when attention was raised with an auditory message.

Hulst *et al.* (1998) found that drivers adopt a hierarchy of adaptive strategies to control time pressure in driving. In normal visibility conditions, drivers had a highly anticipatory, proactive driving strategy, i.e. they increased time headway at the particular times when they expected a lead vehicle to brake. When the possibilities for anticipation were reduced by introducing fog, drivers compensated by means of general speed reduction and time headway increase. When the latter compensatory strategy was made impossible/undesirable, i.e. when drivers were asked to follow a certain time schedule in their driving, they instead compensated by increasing their general level of alertness, as testified to by faster and more precise responses to unpredictable hazardous events.

## 2.4 Summary

The concepts of natural performance variability, adaptive control, discomfort avoidance, and proactive attention selection and action are closely related. In principle, if Summala's (2007) discomfort avoidance principle is applied to the dynamic control approach, then selection of reference values for control processes involved in driving can be viewed as a balancing act between the desire for goal fulfillment and discomfort avoidance. If proactive attention selection and action is added to the mix, then a key enabler of critical situations would seem to be situations where natural performance variability leads to contextually inappropriate attention selection and/or action. Consequently, the role of ADAS can be conceptualized as means of keeping such natural performance variability within safe limits.

This view seems to be corroborated by findings in recent naturalistic driving studies, which indicate that attention failures, in particular visual inattention to the forward roadway, in combination with unexpected external events such as unexpected lead vehicle braking) are key mechanisms behind the occurrence of critical events (Dingus *et al.*; 2006; Olson 2009). Although this provides a good starting point for a conceptual framework for ADAS evaluation, many details need to be worked out before a proper conceptual framework can be said to have been formulated.

# 3 Scenario definition
## - Using crash data to specify test scenarios

The third step in the PReVAL evaluation framework is specifying relevant test scenarios for evaluating the ADAS at hand. A necessary basis for defining such test scenarios is a correct and detailed characterization of the sequence of events which leads to the collision type to be prevented (Najm *et al.* 2002). Moreover, it is necessary for this characterization to include information on causal factors (Najm *et al.* 1995).

In passive safety, crash investigations has proven to be a very valuable source of information for understanding injury causation mechanisms and prevalence. For active safety, it is hoped that crash investigations could take on a similarly important role, i.e. that crash investigations for active safety functions would deliver detailed descriptions of crash causation mechanisms for the most prevalent crash types that can be used for legislation, countermeasure development, public information campaigns, etc. (SAFETYNET, 2005*b*).

To deliver detailed crash causation descriptions from crash data, at least one of two particular problems need to be addressed. The first can be called the generalization problem and the second the case selection problem. These will be discussed in turn.

## 3.1 The generalization problem

The generalization problem is the question of how to determine that findings from a particular set of crashes are representative of all crashes of that type. In passive safety, this problem has to a large extent been overcome due to a lot of hard work and the laws of physics. In very simplified terms, if a vehicle breaks in a certain way in a crash on the field, then the laws of physics predict that if a kinematically similar crash occurs, a similar vehicle would likely break in the same way, since it would be exposed to a similar onset of accelerations and forces. Thus, it is possible to extrapolate from information on how one vehicle breaks to how all similar vehicles will break.

For the humans involved in the crash, the mapping is less straightforward than for the vehicles, i.e. rather than extrapolating from one crash test dummy to all humans, particular dummies that represent certain percentiles of the population, e.g. a 95 percentile male, or a 50 percentile female, have been developed. Extrapolation to humans in general, i.e. to account for the natural, biological variation in the population, therefore requires more than one crash test dummy. However, the number of dummies needed is still in the single digit range.

For passive safety, the laws of physics thus provides an immensely powerful helping mechanism in addition to all the hard work put in when it comes to generalizing from individual cases to the whole population. A limited number of crash tests are sufficient to make a reasonably certain prediction on a protective function's impact on the injury population as a whole[2]. For ADAS evaluation however, a behavioral Newton who lays down the laws of driver behavior is yet to come. Until then, ADAS evaluation has to approach the issue of generalization somewhat differently. Other ways have to be found to show that a safety problem identified in crash data is representative for ordinary drivers in the real world.

Using crash information to define a safety problem usually means retrieving and analyzing sufficient amounts of data on contexts and causes of failure to describe what the safety problem looks like in reality. This description can then be used as a test scenario basis for ADAS evaluation. Creating such problem definitions is usually a two-step process. Data is first collected from relevant sources. Then the collected data is aggregated to create one or more representations of the safety problem which captures what is considered to be its typical attributes.

To understand what is meant by typical in this context, the distinction between a type and its tokens can be used. A type designates a general item, while a token is a particular and concrete instance of that item (Wetzel Winter Edition 2007). If a safety problem is defined by analyzing the common attributes of a set of crashes, the safety problem representation will be a general crash type rather than a crash token i.e. a particular crash. Furthermore, if the common properties are defined correctly, the representation will not just be a type, but a prototype, i.e. the most typical representation possible for that set of crashes (for more on what defines a prototype, see Rosch 1975a; 1975b; Rosch and Mervis 1975).

Defining prototype rather than token evaluation scenarios can be considered an advantage in ADAS evaluation if one assumes that solutions to a prototype problem will be of greater help to drivers than a solution customized to any one token problem. For this assumption to hold, prototype safety problems must be defined through an etiological, rather than a descriptive, approach (Haddon 1968). Etiology is the science or philosophy of causation, or the part of any special science which speculates on the causes of its phenomena (The Oxford English Dictionary, 1989). This means that if crash prototypes rather than crash tokens are used to define ADAS evaluation scenarios, their characterization should preferably be based on an understanding of their causation mechanisms rather than on their appearance.

Using an etiological approach should, however, not be confused with a never ending search for root causes. What counts as an understanding of the fundamental processes may very well be determined by the constraints of project goals and resources available. Scenarios can e.g. be characterized at a level of detail which matches the countermeasure cost and availability for a certain project, where the best

---

[2] Obviously, this argument wholly depends on the test setup being representative of real world crashes and injury mechanisms, which largely is a research field in its own right.

countermeasure would be the cheapest available solution which solves the largest part of the crash problem.

An etiological approach is thus compatible with defining prototype scenarios according to practical rather than theoretical limits. E.g. if a road administration determines that collisions with oncoming traffic is the prototypical safety problem, they could spend a lot of time investigating all causation mechanisms which make drivers leave their lane. However, another fundamental understanding of the problem is that for a collision with an oncoming vehicle to occur, the driver must be able to enter the oncoming lane. Since installing median guard rails addresses that particular problem, and is an available means for the road administration, they could settle for deploying median guard rails and leave the fundamental understanding of why drivers leave the lane to others.

## 3.2 Prototypical causation and the problem of data sources

Following Haddon's line of reasoning, an ideal crash typology for active safety should divide crashes into types based on distinct and typical causation mechanisms. After using this typology to select and sort all crash data available into their groups, the typical circumstances under which each type occurs can then be characterized using descriptive variables available in the data. The goal of the process would be a set of prototypical crash scenarios with sufficient level of detail for active safety function development and evaluation.

In order to identify the details of crash types relevant for ADAS evaluation, data from crash investigations is often used. Typically, a pre-crash typology is specified where each crash type in the typology represents a problem situation occurring before the crash. Available crash data is sorted according to the types of the typology, and the typical conditions of each crash type is established by identifying its mean and/or median values for each variable used to code crash information. These values are then compiled and presented as the typical conditions under which evaluation of an ADAS addressing that crash type should take place.

When setting out to define the causes and contexts which make a crash scenario prototypical, an often used source is official databases containing police reported crashes. These databases can be said to contain macroscopic data (OECD 1988) because the number of crashes they contain and the way these are sampled usually make the databases statistically representative for crashes occurring in a certain region, such as a country. Typical values on pre-crash conditions retrieved from a macroscopic database will therefore be statistically representative for the crash population in the sampled region, which is a very strong support of their typicality. Examples of studies using macroscopic data to characterize crash types include NCSA (2003; 2005; 2006; 2007), SAFETYNET (2005*a*; 2007*a*; 2007*b*), and SIKA (2006*a*; 2006*b*).

From an ADAS evaluation perspective, the weakness of police reported crashes is that though they can be quite detailed in describing the crash context conditions  i.e. time of day, road type, vehicle type, driver age, etc., they usually contain limited information on why crashes happen (Larsen 2004; Sabey 1990). The latter may be due to the fact that police normally aim to determine responsibility rather than to describe exactly why the crash occurred. Another reason is the tendency of drivers to underreport certain factors, such as fatigue (Anund *et al.* 2004). Pre-crash typologies developed from police reported crashes can thus be detailed in context description, but will be limited in causation description.

This means that if only police reported crashes are used to define a pre-crash typology, important parts on causation mechanism are left unspecified, and the requirement of basing the prototypical crash scenario on an understanding of its causation mechanisms is difficult to meet. Other data sources must therefore also be used.

This raises two questions. The first is where else to find information on crash contributing factors. The second question is how to determine which contributing factors should be kept and which should be discarded, i.e. which are typical for a crash type.

Neither of these questions are new. Passive safety faced both questions from the start many years ago. There, the first question was answered through the initiation and use of case studies. Case studies, or microscopic data, are normally distinguished from macroscopic data because crashes are viewed and treated as individual cases rather than as a group. In-depth case studies deliver very detailed data on contributing factors compared to police reported data, and they contain extensive, theoretically anchored efforts to understand and describe why crashes occur. According to Larsen (2004) and Midtland *et al.* (1995), qualitative in-depth crash information is the best option for identifying interactions between contributory factors, i.e. for defining causation mechanisms.

Case studies provided an excellent source of injury mechanism information for passive safety. Judging from two recent accidentology projects targeting active safety's need for information on causation mechanisms, case studies seem to be a possible route to take for scenario definition in ADAS evaluation as well. In the Swedish project FICA (Ljung *et al.* 2007) and the European project SAFETYNET (SAFETYNET 2005*b*; 2008; Wallén Warner *et al.* 2008*b*), new methodologies developed to uncover and code combinations of crash contributing factors from an active safety perspective resulted in sets of case studies that contain detailed descriptions of each crash's contributing factors. These can provide the information needed to recreate characteristics of contributing factors in ADAS evaluation scenarios.

However, when defining a prototypical scenario, one must verify that it indeed represents a general pathology rather than being specific to a particular site or set of drivers (Fleury and Brenac 2001). Here, case studies fall short. In in-depth case studies, the number of cases collected is usually small and the sampling procedure is rarely nationally or internationally representative. Which of the causation mechanisms found in case studies to consider prototypical can therefore not be established directly from case studies.

One strategy for creating a connection between macroscopic data and crash causation mechanisms in case studies, which several researchers have explored, can be called the context matching approach, (see e.g. Chovan *et al.* 1994; General-Motors-Corporation 1997; Najm *et al.* 2007; McCarthy *et al.* 2010; Wisch *et al.* 2010). Studies following this approach first select a crash type from macroscopic data. Then a set of case studies which occur in a similar context i.e. same type of road, same weather, etc. is selected from microscopic data. The case studies are analyzed to identify their common causation mechanisms, and then the researchers infer those mechanisms to be typical of the crash type.

However, the context matching approach suffers from an inherent weakness, which is that it cannot a priori be concluded that because an individual crash occurs in a context similar to that of a general crash type, its causation mechanisms will be representative of that crash type. It is possible that the case studies selected by context matching happen to represent only a few odd crashes, for which the contributing factors have nothing to do with how typical drivers get involved in the crash type.

Another approach is to start from microscopic data, using the prototype scenario approach (Fleury and Brenac 2001) applied e.g. in the project TRACE (Naing *et al.* 2007; Van Elslande and Fouquet 2007). First, the case studies are grouped based on their similarity in terms of both context properties and causal relationships. When the grouping is complete, a prototypical scenario is built for each group by identifying its main features (Fleury and Brenac 2001).

The prototypical scenario approach suffers a weakness similar to the context matching approach. Since microscopic data rarely is representative, it by definition becomes near impossible to determine whether a scenario built only from case studies is representative. To illustrate, Swedish statistics from police reported crashes identify a crash type called single vehicle crash (SIKA 2006*a*). When analyzing of set of case studies that match this macroscopic crash type, four distinct patterns of causation were identified (Sandin and Ljung 2007). The first group of crashes was due to a combination of temporary driver distraction and fatigue; the second to locally reduced friction; the third to a combination of high speed and psychological stress; and the fourth to excessive steering following a surprising event. The crashes in all four groups match the context variables for the crash type as found in the macroscopic data, and no further data is available to determine if any group's causation pattern is typical for the crash type, or if it just represents a few odd drivers that happened to be investigated in that particular project.

## 3.3 The case selection problem

The second problem regarding the use of crash data, i.e. the case selection problem, concerns the type of crashes selected for in-depth analysis. Crash investigations have traditionally focused on crashes with the most severe outcomes, i.e. crashes with fatal and/or severe injury outcomes. E.g. many projects stipulate some form of minimum damage criterion that needs to be met in order for a crash to be investigated, which is highly logical when the purpose is to develop countermeasures which target injury reduction.

Interestingly, in terms of using crash data to develop test scenarios for evaluating active safety functions, this has lead to a situation where there for some countries exist exceptions to the usual case of microscopic data not being representative. In countries like Sweden and Norway, the authorities commission in-depth investigations of all fatal crashes that occur. Thus, there exist some in-depth crash datasets which are both representative and detailed. Provided that these datasets contain relevant information for understanding why these crashes occur, the problem of linking microscopic data to macroscopic data addressed above might be unnecessary to solve, at least for countries where such datasets exist, and provided that the aim is to prevent crashes with injury outcome rather than crashes in general.

Although in-depth studies of fatal crashes clearly can be informative in terms of establishing which injuries a driver has sustained through e.g. post-mortem examination, it is by definition impossible to interview that driver in order to understand why the crash took place to begin with. Thus, even though other witnesses such as the driver of an opposing vehicle can be interviewed, and the crash site as such may provide a lot of useful information, it is unclear to what extent in-depth studies of fatal crashes really do provide a useful source of information for identifying crash contributing factors. This problem needs further study.

## 3.4 Summary

In terms of defining test scenarios for ADAS evaluation as required in step 3 of the PReVAL process, it is clear that an improved generalization process for linking information in case studies to a macroscopic crash type, covering not only context but also causation similarity, is needed. Without such a process, it will remain difficult to determine which contributing factors, alone or in combination, should form part of an evaluation scenario in order to make the evaluation scenario (proto)typical. While Fleury and Brenac (2001) have conducted initial work in this area, they also recognize that a more systematic examination is required to establish the link between prototypical scenarios built from case studies and crash types represented in macroscopic data.

As for the use of data from fatal crashes, their potential use as a source for crash causation information is yet to be evaluated. On one hand, they form a highly relevant, and in some countries fully representative, set of crashes. On the other hand, it is also clear that a number of potential caveats exist which the analysis must overcome. However, provided that the crash information is available, it seems warranted to explore how far it goes in terms of understanding crash causation and setting up ADAS test scenarios.

# 4 Method selection and test plan - Measuring performance in adaptive drivers

Most road safety measures have to influence human behavior in order to be effective (Elvik 2004). Seat belts must be worn in order to protect from injury; headlights must be turned on in order to make the car more visible; drivers must stop at red traffic signals for these to function as intended; and so on. This also apply to most ADAS. The influence can either be explicit, i.e. the function gives a warning and the driver is expected to respond accordingly, or implicit, i.e. the vehicle acts on its own, but the driver can override it at any time.

A key aspect of ADAS evaluation is therefore correct measurement of driver performance, with and without assistance from the ADAS, collected from relevant driving events (Kiefer *et al.* 1999; Ljung Aust and Engström 2011; Najm *et al.* 2000).

This leads to a basic methodological challenge in active safety function evaluation which is the driver's adaptive capacity, as described above. In ADAS evaluation, successful measurement of driver performance depends on the evaluators' ability to create test scenarios which 1) are sufficiently similar to the real world pre-crash scenarios that they are intended to represent, yet 2) lead to a sufficiently measurable degree of driver response which can be used to identify performance changes driven by the presence of an ADAS.

As reviewed above, driving is largely a self-paced task where drivers actively control their own vehicle based on how they expect the traffic situation to develop in the near future. If the driver anticipates that the traffic situation somehow might turn critical, s/he will normally adapt to avoid this by slowing down, increasing headway and/or pay more attention to, and invest more effort in, the driving task (Hulst *et al.* 1998).

The power and success of this innate capacity to adapt is testified to by the rarity of accidents in terms of frequency of occurrence over total driving time. In Swedish data on police reported crashes for 2007 found in STRADA (2007, the latest year for which final data has been released), there were 21 838 drivers involved passenger car crashes, while passenger cars travelled a total of ~64 billion km (SIKA 2007) This translates into approx. 3 million km driven per driver involved in a police reported crash. In addition, the statistics show that the average yearly mileage for passenger cars is roughly 13 000 km. Thus, one needs to drive, on average 223 years, to end up in a police reported crash.

It is thus clear that drivers have a powerful adaptive capacity. Moreover, they will not leave this capacity behind when entering the laboratory. On the contrary, since test participants normally have to fulfill some minimum driving experience criteria, they bring lots of experience in predicting and resolving potentially critical traffic situations with them into the laboratory. This is testified to by numerous studies, which have found e.g. that drivers reduce speed and/or increase headway if there is a lead vehicle present when they are doing a secondary task (Antin *et al.* 1990; Jamson and Merat 2005; Engström *et al.* 2005). Furthermore, Rauch *et al.* (2008) showed that drivers slow down before engaging in a secondary task and/or before they expect the driving situation to become demanding.

Note that a proper definition of adaptivity really needs to refer to some type of time frame. When evaluating ADAS in a laboratory setting which is the focus here, that time frame is usually limited, i.e. participants come to the laboratory for a few hours at the most, and adaptivity in this setting is thought of as very quick alterations of behavior within the ongoing study. Another common way of studying behavioral driver adaptation is to look at how drivers respond to general countermeasure implementations, such as laws on the use of studded tires (Elvik 1998); the introduction of airbags (Sagberg *et al.* 1997); the presence of road lighting (Assum *et al.*1999); or bans of cell phone use (Nikolaev *et al.* 2010). In these studies, the time frame of the study is often quite long, since e.g. reliable numbers on accident frequency takes months rather than hours to accumulate. This might be the reason why the adaptation itself seems more or less implicitly assumed to occur over a longer period of time as well. This certainly is a possibility, though as far as the author is aware, the individual time frame within which people change their behavior as a function of putting on studded tires or driving on a lit rather than dark road is neither known nor well studied. For the present discussion, this aspect of adaptive behavior will be left aside, although it certainly constitutes an important field of research in its own right.

Coming back to ADAS evaluation, in order to assess the influence of an ADAS, one must often place the driver in a situation that is rarely encountered and which the driver normally would do his/her best to avoid. It follows that the test scenarios have to be carefully crafted in order not to provide test drivers with clues that could initiate premature adaptation, while still maintaining all relevant properties of the targeted real world critical event. Conversely, if the evaluation scenario fails to generate a driver adaptation failure similar to those found in the real world pre-crash situations which the ADAS is intended for, the validity of the evaluation results is questionable.

If e.g. drivers in a driving simulator study assessing the influence of FCW on BRT expect that the lead vehicle will be braking, the BRT measures will not be representative of BRTs for truly unexpected and urgent situations where drivers panic brake to avoid crashing. As discussed above, Green (2000) clearly showed that BRT depend heavily on the level of expectancy in the driver. Hence, drivers should respond faster and faster as events repeat if expectancy increase with exposure. This was also shown empirically by among others Lee *et al.* (2002). To properly replicate this real life scenario in an experimental setting where FCW is to be evaluated, each critical event should thus ideally come as a complete surprise to the driver, since that type of event occur very unexpectedly from the driver's point of view (Dingus *et al.* 2006). How to replicate such scenarios, and perhaps even more importantly, how to know that one has achieved a successful replication, is a topic that merits further study.

Furthermore, even if the first trial scenario is successful in surprising the driver, the question is whether the surprise effect can be repeated with the same driver. In experimental studies where a group of individuals' responses to different conditions are to be tested, the most economical design, i.e. with the most statistical power in relation to how many test subjects are required, is a within group design where the same persons experience different conditions in some counterbalanced order (Field, 2009). This approach has been repeatedly used in ADAS evaluation.

A common setup in e.g. FCW studies is to repeatedly expose subjects to critical events, and then use each driver's average response when comparing the performance of drivers with and without FCW (Abe and Richardson 2004; 2005; 2006*a*; 2006*b*; Cheng *et al.* 2002; Jamson *et al.* 2008; Kramer *et al.* 2007; Scott and Gray 2008)., Provided that driver behavior is inherently adaptive, the question arises of whether a within group experimental design is at all suitable for ADAS evaluation. Even if measures like random event timing and catch trials are put in place to mask event reoccurrence, it is hard to avoid an increased expectancy and response readiness in the test subjects.

## 4.1 Summary

Regarding the method selection and test plan design in steps 4 and 5 of the PReVAL flowchart, it is clear that given the recent findings which indicate that unexpectedness is a key property of real world critical events, the possibilities of creating genuinely unexpected events in a controlled experimental setting is key to successful ADAS evaluation. To what extent it is possible and not the least verifiable needs further study. Furthermore, the use of experimental designs that involve repeated measures need careful consideration. In particular, the possibilities of repeating critical events without compromising result generalizability must be studied further.

# 5 Test execution, analysis and reporting - Interpreting results when evaluating multiple ADAS at the same time in the same vehicle

Recently, an alternative approach to laboratory testing of ADAS, e.g. test tracks and driving simulators, has been tried out. This approach is called Field Operational Tests (FOT) (FESTA-Consortium 2008). In a FOT, data is collected from a large number of ADAS-equipped vehicles driven in naturalistic conditions over an extended period of time. Large-scale FOTs have been conducted in the US and Japan on a variety of active safety systems (see e.g. Ervin *et al.* 2005; LeBlanc *et al.* 2006*b*; Sayer *et al.* 2011), and similar initiatives are underway in Europe, e.g. the EU project euroFOT (euroFOT 2009), and the Swedish-US collaboration project SeMiFOT (SeMiFOT 2009).

Conducting a FOT to evaluate an ADAS impacts on driver behavior solves many of the problems related to trying to actively place adaptive drivers in undesired critical situations. Since the drivers' experiences during the study arguably are consistent with events in the real world, critical events that occur during the study have very high ecological validity (FESTA-Consortium 2008).

However, with the deployment of recent FOTs like euroFOT, IVBSS, and SeMiFOT, another methodological concern has emerged; the issue of how to estimate the influence on driver behavior of one ADAS when multiple other ADAS are also present in the vehicle (Ljung Aust *et al.* 2011). In FOTs, hypotheses on ADAS influence are usually formulated on a per function basis, i.e. researchers assume that it is possible to determine the individual safety impact of any given function.

Depending on the project, this assumption comes from one of two possible underlying reasons. One is the experimental setup. If they have equipped the FOT vehicles with only one ADAS, it is quite certain that it is the impact of that function alone that is being tested. The second reason is what can be called the local influence conjecture. Though rarely explicitly formulated in the research plans, it is assumed that the influence of an ADAS is limited to the particular times and situations when it is active and interacting with the driver. In other words, drivers will respond locally to an ADAS, but not show signs of adaptation or behavioral change outside those interactions. If this conjecture is accurate, then a research vehicle can be equipped with multiple ADAS, and as long as their interaction with the driver does not overlap in time, each function's individual influence can be evaluated.

However, the standalone ADAS experimental design will most likely not be valid in the future since modern production vehicles are to an increasing degree being equipped with multiple ADAS functions (Faber *et al.* submitted). Since different ADAS can use the same sensor once the sensor is in place, the cost attached to developing additional ADAS is limited. Radar e.g. is needed for FCW, there is limited extra cost attached to developing and offering ACC or other functions that require lead vehicle monitoring.

The local influence conjecture is challenged. In two driving simulator studies, drivers were first visually distracted and then exposed to a surprise lead vehicle braking event in which they received a FCW (Ljung *et al.* 2007). Results showed that drivers who had received LDW warnings in the vehicle prior to the surprise FCW event responded well to the FCW, while the drivers who did not experience vehicle warnings before only responded partially to the FCW. Previous interaction with LDW thus improved driver responses to FCW, presumably because those drivers had learned to look back to the forward roadway and be prepared to respond in the case of a warning. Although this effect might be inflated due to the limited driving time, i.e. total driving time was approx. 20 minutes so the LDW interaction was in the very recent past for the drivers who did well with FCW, it still opens up the general possibility that ADAS influence is non-local.

This presents a scientific challenge for ADAS evaluation. In the case of an FOT where vehicles are equipped with an ACC and FCW bundle, both FCW and ACC may influence longitudinal control of the vehicle. Now, provided that ADAS influence potentially is non-local, and assuming that the frequency of critical lead vehicle braking events goes down in the treatment phase, the question is which ADAS to credit with this improvement.

Unfortunately, the available advice on how to deal with this problem is limited. Even though road safety programs consisting of a large number of road safety measures have been developed in many countries, empirical studies dealing with combined safety systems effects are scarce, and ideas on how to approach this problem are limited (Elvik 2009; Siegrist 2010). In terms of multiple ADAS assessment, the only explicit proposal which the author is aware of comes from Gordon *et al.* (2010). They propose a model which suggests that if there is enough information to do a detailed sequential, analytical breakdown of the typical flow of events leading up to a particular crash type, it is possible to tell the relative impact of two or more ADAS apart by analyzing which contributing factors in the flowchart each ADAS is targeting and assessing their relative success in this endeavor. However, the model was not empirically verified in that study.

Elvik (2009) suggests four general conceptual models for calculating combined effects of traffic safety measures. The additive effects model assumes that safety effects are independent and simply can be added. The dominated effects model assumes that the most efficient countermeasure will dominate all others. Last, the independent effects model and the correlated effects model both assume that the effectiveness of measure B only can be applied to the set of crashes that is left after measure A is introduced; a principle which assures that the combined effect never can exceed 100 %. The difference between the two latter is that the correlated effects model also assumes that introducing B in the traffic system partially weakens the effect of A. A's effectiveness is thus lower in combination with B than standalone.

In Elvik's study, the available data which mainly comes from studies of junction improvements, fit both the independent and the correlated effects models and could not distinguish between the two. However, Elvik argued in favor of the more conservative correlated effects model, based on the fact that risk factors tend to be correlated rather than orthogonal, e.g. drivers who are likely to drink and drive are also less likely to wear a seatbelt (Nilsson 2004). Andersson (1998) and Gordon *et al.* (2010) seem more oriented towards the independent effects perspective. Siegrist (2010) ends up somewhere in between Elvik and Gordon, i.e. he argues for a correlated effects model based on the overlapping properties of various countermeasures, but also allows the effect reduction due to countermeasure overlap to be offset by positive synergy effects, e.g. deployment of multiple countermeasures lead to a general increase in public awareness about the problem.

## 5.1 Summary

The multiple ADAS issue clearly present a challenge in conducting step 6 of the PReVAL process, i.e. test execution, analysis and reporting. Since both theories and empirical data on how to address this problem are limited, in particular for bundled ADAS; an initial step would seem to be to empirically test whether data from an FOT that evaluates a bundled ADAS supports either the explicitly ADAS oriented independent effects model from Gordon *et al.* (2010), or any of the other general conceptual models proposed by Elvik (2009).

# 6 Aims

The general aim of the present thesis is to improve key steps in the procedure for functional, formative evaluation of ADAS. This general aim can be broken down into the following five particular aims:

- *To develop a general conceptual framework for ADAS evaluation that can be used for formulating relevant functional descriptions and generating testable hypothesis on ADAS influence in critical driving scenarios*

- *To define a methodology for linking information on crash causation found in case studies to crash types defined using macroscopic data*

- *To assess whether in-depth studies of fatal crashes yield crash causation information that is relevant for setting up ADAS evaluation requirements*

- *To study whether critical driving events that result in realistic driver responses can be created and repeated in laboratory based ADAS evaluation*

- *To empirically test whether existing conceptual models for assessing the combined effect of multiple safety functions are applicable to FOT data collected from bundled ADAS functions*

# 7 Paper summaries

## 7.1 Summary of Paper I

**A conceptual framework for requirement specification and evaluation of active safety functions**

**Introduction**
To verify that active safety functions have the intended effect on driver performance and ultimately reduce the number of crashes, adequate evaluation methods are needed. A prerequisite for such evaluation is a conceptual framework, i.e. a set of concepts and principles which give an accessible yet powerful and scientifically valid description of the research field in which the evaluation takes place. For ADAS evaluation, such a framework is lacking.

**Method**
A general conceptual framework containing concepts and principles suitable for ADAS evaluation was described. The framework is intended to be applicable to the whole evaluation process, including the “translation” of accident data into generally applicable evaluation scenarios, definition of evaluation hypotheses, and selection of performance metrics and criteria. The framework is also meant to be generic, so while it mainly is intended for use in the context of controlled experiments in driving simulators and test tracks, it should also be applicable in other evaluation environments, such as large-scale FOTs. The framework’s applicability to ADAS evaluation was then tested by applying it in the context of writing and implementing requirement specifications for ADAS evaluation.

**Results**
A general conceptual framework containing concepts and principles suitable for ADAS evaluation could be described. The framework is called Situational Control, and describes driving as a control process motivated by a desire to avoid discomfort. The test of its applicability to ADAS evaluation illustrated that the framework can be used to define the necessary steps in writing and implementing requirement specifications for ADAS evaluation.

**Discussion**
While many details remain to work out, the Situational Control framework shows initial promise for work related to creating a conceptual framework that can fulfill a similar role for active safety function evaluation as Haddon’s energy transfer model does for passive safety function evaluation.

## 7.2 Summary of Paper II

**Generalization of case studies in road traffic when defining pre-crash scenarios for ADAS evaluation**

**Introduction**

To define pre-crash scenarios for evaluation of active safety functions, data from crash investigations is often used. Typical sources of crash data are official databases containing police reported crashes and in-depth case studies. The official databases can often be considered representative, but they contain little detail on causation. The opposite it true of in-depth case studies, i.e. they have much detail on causation but are rarely representative. Since pre-crash scenarios for evaluation of active safety functions need to be both representative, as well as contain detailed information on causal factors, combining data from the two sources when defining pre-crash scenarios would be ideal. However, a general difficulty in such work is how to establish that causation information identified in a set of case studies actually is representative of a crash type identified in an official database.

**Method**

In this study, a new methodology for linking causation information in case studies to a crash type selected from official databases is proposed and tested. The new methodology, called integrated context and cause matching, uses an intermediate layer of representatively sampled crash information to evaluate whether the causation information identified in a set of case studies is representative of a selected crash type. The intermediate layer, which in this study is based on questionnaire responses from crash involved drivers, thus act as glue, or a bridge, between the case studies and the crash type.

**Results**

The study shows that it is possible to create such an intermediate layer, and then compare it to the case studies on causation patterns and the selected crash type on context properties. The comparisons showed similar patterns of crash contributing factors in the case studies and crash questionnaire responses, as well as a high level of context similarity between questionnaire responses and the macroscopic crash type.

**Discussion**

The results indicate that the information on crash causation present in the selected case studies are representative of the crash type. They also indicate that while some methodological issues remain to be addressed, the proposed methodology shows good promise for work related to defining pre-crash scenarios for ADAS evaluation.

## 7.3 Summary of Paper III

**Fatal intersection crashes in Norway: patterns in contributing factors and data collection challenge**

**Introduction**

Fatal motor vehicle intersection crashes occurring in Norway during the years 2005-2007 were analyzed to identify causation patterns among their underlying contributing factors, and also to assess if the data collection and documentation procedures used by the Norwegian in-depth investigation teams produce the information necessary to conduct causation pattern analysis.

**Method**

28 fatal accidents were analyzed. Causation charts of contributing factors were first coded for each driver in each crash using the Driving Reliability and Error Analysis Method (DREAM). The charts were then aggregated based on a combination of conflict types and whether the driver was going straight or turning.

**Results**

Analysis results indicate that drivers who were performing a turning maneuver in these crashes faced perception difficulties and unexpected behavior from the primary conflict vehicle, while at the same time trying to negotiate a demanding traffic situation. Drivers who were going straight, on the other hand, had less perception difficulties but largely expect any turning drivers to yield, which led to either slow reaction, or no reaction at all.

**Discussion**

In terms of common contributing factors, those often pointed to in literature as contributing to fatal crashes, e.g. high speed, drugs and/or alcohol, and inadequate driver training, contributed in 12 of 28 accidents. This confirms their prevalence, but also shows that most drivers end up in these situations due to combinations of less auspicious contributing factors. In terms of data collection and documentation, there was an asymmetry in terms of reported obstructions to view due to signposts and vegetation. These were frequently reported as contributing for turning drivers, but rarely reported as contributing for their counterparts in the same crashes. This probably reflects an involuntary focus of the analyst on identifying contributing factors for the driver held legally liable, while less attention is paid to the driver judged not at fault. Since liability often is irrelevant from a countermeasure development point of view, this underlying investigator approach needs to be addressed to avoid future bias in crash investigation reports.

## 7.4 Summary of Paper IV

**Effects of forward collision warning and repeated event exposure on emergency braking**

**Introduction**

Many experimental studies use repeated lead vehicle braking events to study the effects of FCW systems. It can, however, be argued that the use of repeated events induce expectancies and anticipatory behavior that may undermine validity in terms of generalizability to real-world, naturalistic, emergency braking events. The main objective of the present study was to examine to what extent the effect of FCW on response performance is moderated by repeated exposure to a critical lead vehicle braking event. A further objective was to examine if these effects depended on scenario criticality, here defined as the available time headway when the lead vehicle starts to brake.

**Method**

A critical lead vehicle braking event was implemented in a moving-base simulator. The effects of FCW, repeated event exposure, and initial time headway on driver response times and safety margins were examined.

**Results**

The results showed that the effect of FCW depended strongly on both repeated exposure and initial time headway. In particular, no effects of FCW were found for the first exposure while strong effects occurred when the scenario was repeated. This was interpreted in terms of a switch from closed-loop responses triggered reactively by the situation, towards an open-loop strategy where subjects responded proactively directly to the warning. It was also found that initial time headway strongly determined response times in closed-loop conditions but not in open-loop conditions.

**Discussion**

While these results do not necessarily imply a lack of effect of FCW in real world situations, they raise a number of methodological issues pertaining to the design of experimental studies with the aim of evaluating the effects of active safety systems. In particular, scenario exposure and criticality must be carefully considered when designing studies intended to assess the effects of active safety systems and interpreting their results.

## 7.5 Summary of Paper V

**How to estimate the combined safety effect of multiple advanced driver assistance systems deployed in the same vehicle during field operational tests**

**Introduction**

The FOT has emerged as an important research methodology for assessing the impact of ADAS on driver behavior and traffic safety. Usually, hypothesis formulation and testing in FOTs is based on the assumption that ADAS effects can be tested in isolation, either because the vehicles are equipped with only one ADAS or because researchers assume that the influence of each ADAS is limited to the local time and place where it interacts with the driver. However, future vehicles are likely to be equipped with ADAS bundles rather than standalone systems. Furthermore, recent research indicates that drivers' response to one ADAS can depend on whether other ADAS are present in the vehicle. Models suitable for estimating combined effects of bundled systems are therefore needed. The aim of this study was to empirically test whether data from a FOT that deploys vehicles with ADAS bundles conform to any in a set of currently proposed models, or whether something else is needed.

**Method**

Data from the Swedish test center in euroFOT was used. At this center, 102 vehicles equipped with an ADAS bundle that includes FCW and ACC were deployed for approx. a year per vehicle. Data preparation is not completely finished, but the existing dataset, three months of baseline and three months of treatment, was mined for critical lead vehicle braking events. For these events, BRT and minimum TTC values were calculated. Exposure in terms of km driven per critical event was also calculated. Comparisons were then made between baseline and treatment, with data also separated based on whether the vehicle was actively exercising some form of longitudinal control, i.e. whether cruise control (for baseline) or adaptive cruise control (for treatment) were engaged.

**Results**

Results indicate that neither the Gordon *et al.* (2010) model nor the other general models in Elvik (2009) fit the data well. Furthermore, for FCW, results violated the assumption of only local influence, i.e. FCW effected driver behavior outside the particular traffic situations it is designed to address. Lastly, the frequency of critical events per km driven was significantly different when regular or adaptive cruise control was in use, i.e. when driving was partially automated, compared to when it was not.

**Discussion**

Existing models for assessing the combined influence of multiple ADAS in a single vehicle seem to be too simplistic. Instead, some form of multidimensional model that is built around drivers' behavioral adaptations to ADAS influence seems necessary to develop.

# 8 Discussion

As stated in the introduction, from both a theoretical and empirical perspective a lot of work remains before the procedure for ADAS evaluation reaches the same level of detail and standardization as those used for evaluation of in-vehicle injury protection functions. Views on what constitutes relevant theory; what is relevant treatment of empirical data; how drivers should be prepared before a study; and how the study itself should be designed have yet to converge on common definitions. Discussing these theoretical and empirical problems and pursuing solutions to them is therefore a key issue for improving the for ADAS evaluation procedure.

The general goal of the present thesis was to improve on the six steps defined in the PReVAL procedure for functional, formative ADAS evaluation. The extent to which this goal has been met will be discussed below. After that, some general remarks and suggestions for future work will be presented.

## 8.1 Improving steps 1 and 2 – developing a conceptual framework that facilitates formulation of functional specification and expected impacts

The first aim of the thesis was to develop a general conceptual framework for ADAS evaluation that can be used for formulating relevant functional descriptions and generating testable hypothesis on ADAS influence in critical driving scenarios. In relation to other conceptual frameworks related to traffic safety in general and ADAS in particular, Situational Control indeed seems to fill a gap, i.e. there does not seem to exist any other framework that explicitly addresses ADAS evaluation in specific test scenarios. Conceptual frameworks in traffic safety seem more oriented toward more general issues, such as general views on the driver's role and responsibilities in the traffic system (Larsson *et al.* 2010); how to efficiently support prioritization between different countermeasures (Runyan and Yonas 2008); or how to do quantitative modeling of driver behavior that can potentially be used for ADAS assessment (Cacciabue and Carsten 2010).

Another way of assessing whether the Situational Control Framework presented in Paper I fulfills this aim is to check if it meets the three criteria for conceptual frameworks specified in Section 2. The first criterion was that the framework should capture the characteristics of what driving is, as well as the reasons for why it may fail, by describing the three components cognition, motivation and failure in suitable ways. The Situational Control framework does explicitly account for all these components.

The second criterion was that the framework should be applicable to the specific purpose of generating testable hypotheses on the changes a proposed ADAS may create in a problematic driving situation. The analysis in Paper I indicates that the Situational Control framework is well suited for this purpose. By characterizing function influence as changes in relevant DVE parameters in relation to a safety zone boundary, hypotheses on function influence are easy to formulate, and also to give a quantitative form. The latter is very important, since it provides the basis for evaluation based on objective rather than subjective assessment.

The third criteria states that the conceptual framework must allow a practically unambiguous use within a group of people, i.e. the scientists and practitioners who work broadly within the domain. Whether Situational Control fulfills this criteria is less easy to determine. The most pragmatic measure of success is probably the extent to which the framework actually is put to use among those who work in the domain. Obviously, only time can tell, but at least initially the prospects look quite good. Judging by recent work in the project INTERACTIVE (Bakri *et al.* 2011), the framework has been successfully used to formulate the problems and answers the projects and studies set out to define and obtain.

It should be recognized that the Situational Control framework does not address potential issues with drivers' basic motor capabilities. This aspect might need incorporation in the future. Older drivers e.g. are quite aware of their somewhat slower response times, and this may influence the way they select target DVE values, i.e. choice of speed, route, etc. However, one way of describing such adaptive behavior is that it constitutes a change in the comfort or safety zone boundaries of that particular population. Such effects can therefore most likely be incorporated without having to change the basics of the framework itself.

## 8.2 Improving step 3:I - Stronger links between causation information in a case studies and general crash types

The integrated cause and context matching principle proposed in Paper II seems to be a promising way forward when it comes to linking in-depth case study information to general crash types. It overcomes the limitations inherent to the context matching principle and the prototypical scenario approaches by offering a way of generalization from a limited crash set that is based not only on context but on causation information as well. This is clearly an improvement compared to existing approaches.

However, the success in applying the methodology clearly depends on researchers being able to collect or access relevant questionnaire data in sufficient quantities. There also seem to be some very basic limitations to questionnaire data, stemming from its self-reported nature and the fact that crash involved drivers rarely are professional road or vehicle analysts. The nature of these limitations suggests that they may be difficult to overcome solely by improved questionnaire design.

In the end though, the fact that the empirical data in Paper II showed that general trends on causation in both in-depth and questionnaire data were similar, and that the differences between the macroscopic crash type and the questionnaire dataset were explainable without breaking the connection, makes a strong case for continued exploration of the methodology.

## 8.3 Improving step 3:II – using in-depth studies of fatal crashes to capture causation information

Fatal crashes form a relevant, and in some countries fully representative, set of in-depth investigated crashes. However, the study in Paper III shows that this dataset might not be relevant for ADAS evaluation, or at least must be analyzed with caution. Two main reasons for this are that driver fatigue and driver distraction did not show up as contributing factors in the study, despite being identified as main contributors to loss of situational control and thus to traffic accidents in numerous other studies (e.g. Anund *et al.* 2004; Hickman 2010).

There are two ways to explain why these factors do not show up; either the accident investigators failed to identify instances where they contributed, or these factors do not contribute to fatal crashes. In case of the former, investigators might need more training to recognize them. However, it is more likely that they are just hard to capture through fatal crash investigations. The information necessary for coding them as contributing factors usually comes from driver interviews, which is inherently impossible for at least one subject in a fatal crash. Furthermore, the Norwegian in-depth investigation teams rarely conduct their own interviews with those involved who can be interviewed; instead they use protocols from police conducted interviews. Here it can be hypothesized that crash survivors are not always completely forthcoming when describing crash circumstances to the police, but also that the police might not be asking the type of questions a crash investigator would ask. Thus, if the first explanation is right, then fatal crash investigations will always have an inherent omission bias for these factors.

On the other hand, if the second explanation holds, then one has to conclude that the causation mechanisms underlying loss of situational control in fatal crashes are at least partially different from those for crashes with less severe outcomes, and thus what separate ADAS for fatal and other crashes might need to be developed and tested. This possibility is further supported by the fact that not only distraction and drowsiness, but many more of the possible contributing factors available to the crash causation coder in DREAM, were never applicable to the data, despite DREAM having been put through fairly extensive validation work to corroborate it with other researchers' findings on crash causation (Wallén Warner *et al.* 2008*a*).

For the second explanation to hold, however, a prerequisite is that there exists causal mappings between pre-crash behavior and post-crash bodily injury which prevent distraction, fatigue, as well the other unused factors from contributing to fatal outcomes. The only mapping that comes to mind in this case is between distraction and speed selection, where the idea would be that as speed, and thus the amount of energy to be dissipated in a crash, increases, the influence of distraction on the driver's situational control is reduced. In other words, at high speeds the driver's full attention is required to keep the car on the road, and the driver will therefore not engage in secondary tasks. However, while this conjecture might apply to some of the involved drivers, e.g. the motorcycle drivers who were speeding for sheer excitement, a majority of the crashes in Paper III took place at normal speeds within the legal limits of the traffic system, and thus would need another explanation of why distraction does not contribute.

Which of the explanations above that best suits the crash type is a topic for further study. However, based on the study in Paper III, in-depth investigations of fatal crashes does not seem to be the best source of information for defining ADAS evaluation scenarios.

## 8.4 Improving steps 4 and 5 - creating and repeating critical driving events in driving simulator based ADAS evaluation

The recent findings in naturalistic driving studies which point out unexpectedness as a key property of critical events in the real world (Dingus *et al.* 2006), together with the long tradition of research showing that expectancy heavily influences response times, clearly indicate that successful ADAS evaluation depends on the possibility to create genuinely unexpected events in a controlled experimental setting.

The study in Paper IV suggests that creating a single unexpected critical event is feasible, at least in a driving simulator environment where the laws of physics can be manipulated without the test subject realizing it. However, repetition of such critical events generate a relatively complex adaptation process, where adaptation also occur at different rates for drivers with and without the evaluated ADAS. Similar findings, and the problem they pose for ADAS evaluation requirements, have as far as the author is aware, not been discussed in the literature, although the implications are quite profound.

Firstly, the use of repeated critical events to evaluate ADAS in an experimental setting might not be a viable approach towards generalizability. Secondly, the lack of a FCW effect on response times in the first event shows the importance of relevant driver preparation. To get valid results, ADAS evaluation would benefit from using subjects with previous experience of the ADAS in everyday driving. Drivers with such experience can be expected to have a well formed association between the warning and the intended response, which would increase the likelihood of a more representative response in the first and thus genuinely unexpected and critical event. In addition, they would be less likely to respond directly to the warning, as they would understand that such warnings generally require a situation assessment prior to response. This would reduce the likelihood of interaction effects between ADAS and event repetition if the critical event was repeated, assuming that these are the product of the altered response strategy displayed in this study.

Another experimental design issue highlighted in the study is that the available room for ADAS driven improvement is determined by the gap between when a warning is given and when looming cues that trigger a braking or steering response become critically strong. With the exception of Engström (2010), the author is not aware of this fact being discussed in the literature either. However, it has important implications for ADAS evaluation. If e.g. the road authorities in a country decide to develop an FCW evaluation scenario that can be used for official approval or rating but which does not address this issue, manufacturers might lower their warning thresholds in order to trigger the warning early, and thus maximize the measured response time improvement provided that at least some drivers respond to warning timing rather than looming cues.

Application of such strategies would have unfortunate consequences for real world driving. Framed within the Situational Control framework, if an official test rewards sensitive FCW systems that give early warnings and such systems therefore are the ones deployed on the market, then many of the warnings in real life will occur while the driver is well within his/her comfort zone boundary. These warnings would therefore be perceived as nuisance warnings, even though the driver might fully understands why they were given, and hence reduce the likelihood that the driver will treat each alarm as indicating a potentially critical situation that requires immediate and accurate assessment. Early warnings might therefore reduce the likelihood of the function leading to an actual reduction in the number of crashes.

## 8.5 Improving step 6 – selecting a model for assessing combined ADAS effects

An appropriate analysis of combined ADAS effects requires an empirically validated prediction model. As Paper V indicates, this deserves attention in future research, since none of the models from the literature review fit the empirical data particularly well.

To fit the data within the modeling paradigm suggested by Elvik (2009), one has to formulate a model beyond the suggested end point referred to as the dominant effects model. Since critical events occurred more often when both ACC and FCW were active compared to the average baseline, but less often when FCW only was active compared to the average baseline, an interpretation within the paradigm would state that the positive effect of a standalone FCW is not only negated but cancelled with margin when bundled with ACC.

Such a model would however probably be inaccurate for several reasons. First, as the analysis of other variables in the dataset showed, the circumstances under which drivers opt for partially automated control, i.e. use CC or ACC, are different compared to those under which drivers use fully manual control. Second, conflict rates could also differ due to the way semi-autonomous driving might influence drivers' attention selection and situational control.

This illustrates the importance of careful selection of comparison conditions when doing ADAS evaluation. The frequency decrease critical events was 23% when FCW only was active, while it was 52% when FCW and ACC were simultaneously active. Provided that it is appropriate to separate partially automated driving from fully manual driving as done in Paper V, the results can instead be said to indicate that FCW in general lessens the frequency of involvement in critical lead vehicle braking events, and that ACC adds even further reduction under partially automated driving conditions, i.e. when some form of cruise control is being used.

The study also corroborated the findings in Ljung *et al.* (2007), i.e. the influence of FCW did not follow the local influence conjecture given the accident causation mechanism description used. A key assumption underlying the only explicit combined effects model available, i.e. the independent effects model from Gordon *et al.* (2010), was thus not supported. Rather than acting as a local response trigger in critical situations, it seems like FCW more played the role of a driver habituator that through its driver interaction helped improve headway keeping in a manner that lessened lead vehicle conflict frequency.

On the limiting side, it should be pointed out that the interpretations above rest on the assumption that the selected dataset is relevant for prediction of rear end crash involvement. However, none of the critical events identified in this study resulted in an actual crash. One could therefore question whether a dataset like the present one actually is relevant for making predictions about ADAS effects on crash involvement. This is clearly a topic that needs further study.

## 8.6 General remarks and suggestions for future work

It is clear that generally speaking, it is still early days for ADAS evaluation and much further development is to be expected. While the work performed in the present thesis arguably to some extent has improved on the steps of the PReVAL ADAS evaluation process, it is also clear that much more is necessary before ADAS evaluation becomes a fully mature and highly standardized domain.

On a general note, it is clear that when approaching ADAS evaluation in a formative and functional manner, the natural variability and complexity of some of the evaluated components have to be simplified. This holds in particular for the drivers and the traffic environment. For example, in Sweden, out of a population of ~9.5 million (SCB, 2012), approximately 6.2 million hold a driver's license. Thus, even if an evaluation of an ADAS unique to Sweden was to include 40, 60 or 100 subjects, there is a clear risk that the behavior of the selected subjects might not be representative of the 6.2 million. The same argument can be made for representations of the traffic environment, e.g. intersection layout and design is not uniform throughout Sweden, so a couple of generic representations e.g. in a driving simulator would not capture the full variability of Swedish intersections.

The extent and severity of these generalization problems in ADAS evaluation are largely unknown. One could hypothesize that the variability is very large due to the population size, but one could also hypothesize that drivers on average have so many hours behind the wheel when they encounter a truly critical situation that their responses are highly overlearned and more or less identical. As far as the author is aware, no one has yet attempted to calculate how large a driver or road design sample must be in order to capture say 95 percent of the total variability in driver responses to critical situations and corresponding ADAS input at those locations.

However, judging by the variability in driver response from the empirical studies in the present work, safety margins seem to be a personal thing. For ADAS evaluation, it might therefore be the case that the population variability is so large that use of mean values for response times, etc., will not be meaningful. As ADAS developers are likely to recognize this differentiation and invest efforts into personalizing ADAS properties and interfaces, future ADAS evaluation procedures have to take this into account. It carries numerous implications for key issues like test subject preparation, individual tailoring of test scenarios to test drivers' own comfort zones, and the like.

Another conclusion that the author would like to draw from the studies is this thesis is that the time seems to be right to embrace the fact that driver behavior is inherently complex and adaptive in nature. Simple models of driver behavior do not fit empirical data well; they are too static representations of what goes on in driving. More complex models that come closer to real behavior are necessary, and this is likely to have an impact on how testable hypothesis for ADAS evaluation are formulated in the future.

The study in Paper IV pointed out that ADAS evaluation would benefit from using subjects with previous experience of the ADAS in everyday driving. The topic of test driver preparation in terms of previous ADAS exposure is very interesting to pursue. It is an aspect of a more general topic that sometimes is called driver adaptation over time. The basic idea is that learning and adjusting behavior based on ADAS input is an iterative process, so driver responses will vary over time. Furthermore, it is not clear whether drivers reach a stable state over time, or whether responses continue to oscillate. From an ADAS development perspective, this means that optimal settings and interaction will depend on where in that iterative process the driver is when a truly critical event occurs. Finding ways to measure the adaptation process and making sure that the test subject used are at a relevant stage of their adaptation is therefore key both to ADAS improvement and evaluation, and worthy of further study.

In relation to data collection, it seems clear that each data type has its own limitations. Questionnaire data suffers from problems related to self-reporting and the fact that crash involved drivers rarely are professional analysts. Data on driver behavior from driving simulators comes from time compressed exposure to very particular events in an artificial environment. FOT data shows high degrees of variability and low levels of exposure to truly critical events. Video data might create a bias toward visible driver behaviors, while on scene, in-depth analysis of crashes focus a lot more on what the driver was thinking and expecting. Neither of these limitations are new nor surprising, but they clearly indicate that the support data for ADAS evaluation cannot come from a single source only. A full and deep understanding of the problem and how a suggested ADAS will effect it requires combining these data collection methods in an intelligent way.

It is clear that an important challenge in future ADAS evaluation is the measurement of driver performance. In the situational control framework, it is stated that all relevant situational control indicators should be measured. This is relatively easy for driving performance as indicated through kinematics but much more difficult when it comes to measuring how the driver mentally adapts to a development of the situation or the input of an ADAS. Operational definitions of e.g. what it means to have seen an object such as an oncoming vehicle, and what it means to understand the consequences of its presence, have to be developed and tried. However, some caution against making these definitions too rigid or final is also advised, since societal learning might make today's strategies obsolete tomorrow. Much as the general population e.g. now relies on everyday cell phone use in a way that was inconceivable 15 years ago. The same could be hypothesized for ADAS, i.e. as they become commonplace and most drivers get them, it will be more acceptable to rely on these systems for safe transport. Thus, the variability in terms of what the population expects in terms of warning timing, etc. might actually shrink. At the same time, expectations on stability and accuracy may increase. Any such trends are important to follow, as they set clear limits and requirements for ADAS performance.

# 9 References

Abe, G., and Richardson, J. (2004). The effect of alarm timing on driver behaviour: An investigation of differences in driver trust and response to alarms according to alarm timing. *Transportation Research Part F: Traffic Psychology and Behaviour, 7*(4-5), 307-322.

Abe, G., and Richardson, J. (2005). The influence of alarm timing on braking response and driver trust in low speed driving. *Safety Science, 43*(9), 639-654.

Abe, G., and Richardson, J. (2006*a*). Alarm timing, trust and driver expectation for forward collision warning systems. *Applied Ergonomics, 37*(5), 577-586.

Abe, G., and Richardson, J. (2006*b*). The influence of alarm timing on driver response to collision warning systems following system failure. *Behaviour and Information Technology, 25*(5), 443-452

Adams, J. (1995). *Risk*. London: UCL.

AIC (2010). ILAC Policy for Uncertainty in Calibration

Allen, T. M., Lunenfeldt, H., and Alexander, G. J. (1971). Driver information needs. *Highway Research Records 366*, 102 – 115.

Alliance of Automobile Manufacturers (2006). Statement of principles, criteria, and verification procedures on driver interactions with advanced in-vehicle information and communications systems (Version 2.2). . Southfield, MI.

Amalberti, R. (2001). The paradoxes of almost totally safe transportation systems. *Safety Science, 37*(2-3), 109-126.

Andersson, G. (1998). Trafiksäkerhetspotentialer och trafiksäkerhetsreformer 1994-2000. In S. v.-o. transportforskningsinstitut (Ed.). Linköping

Angell, L., Auflick, J., Austria, P. A., Kochhar, D., Tijerina, L., Biever, W. *et al.* (2006). Driver Workload Metrics Task 2 Final Report - Appendices: National Highway Traffic Safety Administration.

Antin, J. F., Dingus, T. A., Hulse, M. C., and Wierwille, W. W. (1990). An evaluation of the effectiveness and efficiency of an automobile moving-map navigational display. *International Journal of Man-Machine Studies, 33*(5), 581-594.

Anund, A., Kecklund, G., and Peters, B. (2004). *Min trötta resa (in swedish)*. Linköping: Väg- och transportforskningsinstitutet.

Assing, K., Baum, H., Bühne, J.-A., Geißler, T., Grawenhoff, S., Peters, H. *et al.* (2006). eIMPACT - Socio-economic Impact Assessment of Stand-alone and Co-operative Intelligent Vehicle Safety Systems (IVSS) in Europé.

Assum, T., Björnskau, T., Fosser, S., and Sagberg, F. (1999). Risk compensation - the case of road lighting. *Accident Analysis & Prevention, 31*(5), 545-553.

Bakri, T., Blanco, R., Fahrenkrog, F., Koskinen, S., Larsson, A., Larsson, P. *et al.* (2011). INTERACTIVE Deliverable 7.1 - Requirements for the Evaluation Framework.

Brehmer, B. (1990). Variable error set a limit to adaptation. *Ergonomics, 33*, 1231-1240

Cacciabue, P. C., and Carsten, O. (2010). A simple model of driver behaviour to sustain design and safety assessment of automated systems in automotive environments. *Applied Ergonomics, 41*(2), 187-197.

Carsten, O. (2007). From driver models to modelling the driver: what do we really need to know about the driver? In P. C. Cacciabue (Ed.), *Modelling Driver Behaviour in Automotive Environments* (pp. 105-120).

Carsten, O., Merat, N., Janssen, W. H., Johansson, E., Fowkes, M., and Brookhuis, K. A. (2005). Human Machine Interaction and the Safety of Traffic in Europe (HASTE): Final Report.

Carsten, O., and Nilsson, L. (2001). Safety assessment of driver assistance systems. *European Journal of Transport and Infrastructure Research, 1*(3), 225-243.
Carter, A. A., Burgett, A., Srinivasan, G., and Ranganatha, R. (2009). *Safety impact methodology (sim): evaluation of preproduction systems* Paper presented at the International Technical Conference on the Enhanced Safety of Vehicles (ESV), Stuttgart, Germany.
Cheng, B., Hashimoto, M., and Suetomi, T. (2002). Analysis of driver response to collision warning during car following. *JSAE Review, 23*(2), 231-237. Chovan, J. D., Tijerina, L., Everson, J. H., Pierowicz, J. A., & Hendricks, D. L. (1994). Examination Of Intersection, Left Turn Across Path Crashes And Potential IVHS Countermeasures, Final Report: United States. Joint Program Office for Intelligent Transportation Systems.
Chovan, J. D., Tijerina, L., Pierowicz, J. A., and Hendricks, D. L. (1994). Examination Of Unsignalized Intersection, Straight Crossing Path Crashes And Potential IVHS Countermeasures, Final Report: United States. Joint Program Office for Intelligent Transportation Systems.
Cnossen, F., Meijman, T., and Rothengatter, T. (2004). Adaptive strategy changes as a function of task demands: A study of car drivers. *Ergonomics, 47*(2), 218-236.
Cnossen, F., Rothengatter, T., and Meijman, T. (2000). Strategic changes in task performance in simulated car driving as an adaptive response to task demands. *Transportation Research Part F: Traffic Psychology and Behaviour, 3*(3), 123-140.
Commission, E. (2008). Commission Recommendation of 26/V/2008 on safe an defficient in-vehicle information and communication systems: Update of the European Statement of Principles on Human-Machine Interface. Brussels: Commission of the european communities.
Damasio, A. R. (1994). *Descartes' error: emotion, reason, and the human brain*. New York: G.P. Putnam.
Deering, R. K. (2004). *Annual Report, April 2001 - March 2002* (Vol. DOT HS 809 531): Crash Avoidance Metrics Partnership, National Highway Traffic Safety Administration.
Dekker, S. (2005). *Ten questions about human error : a new view of human factors and system safety*. Mahwah, N.J.: Erlbaum.
Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J. *et al.* (2006). The 100-Car Naturalistic Driving Study, Phase II – Results of the 100-Car Field Experiment. Cambridge: U.S. Department of Transportation, John A.Volpe National Transportation Systems Center.
Dobberstein, J., McCarthy, M., and Lüdeke, A. (2010). ASSESS D2.1 - Final - Integrated Methodology Framework.
Donges, E. (1999). Conceptual framework for active safety in road traffic. *Vehicle System Dynamics, 32*(2), 113-128.
Eggers, A., and De Lange, R. (2007). APROSYS - Proposed evaluation methodology for advanced safety systems
Elvik, R. (1998). The effects on accidents of studded tires and laws banning their use: A meta-analysis of evaluation studies. *Accident Analysis & Prevention, 31*(1-2), 125-134.
Elvik, R. (2004). To what extent can theory account for the findings of road safety evaluation studies? *Accident Analysis & Prevention, 36*(5), 841-849.
Elvik, R. (2009). An exploratory analysis of models for estimating the combined effects of road safety measures. *Accident Analysis & Prevention, 41*(4), 876-880.

Engström, J. (2010). *Scenario criticality determines the effects of working memory load on brake response time.* . Paper presented at the Proceedings of the European Conference on Human Centred Design for Intelligent Transport Systems (HUMANIST), Lyon, France.

Engström, J., Johansson, E., and Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*(2 SPEC. ISS.), 97-120. Engström, J., & Mårdh, S. (2007). SafeTE: final report (pp. 45). Borlänge, Sweden: Vägverket.

Ervin, R., Sayer, J., LeBlanc, D., Bogard, S., Mefford, M., Hagan, M. *et al.* (2005). *Automotive Collision Avoidance System Field Operational Test Report: Methodology and Results*: University of Michigan Transportation Research Institute, General Motors Corporation, National Highway Traffic Safety Administration.

euroFOT (2009). *European Field Operational Test on Active Safety Systems.*

EuroNCAP (2004). Side Impact Testing Protocol, version 4.2. Retrieved from http://www.euroncap.com/files/Euro NCAP Side Protocol Version 4.2---291ce83b-24e7-4f1d-98e1-6e6d27a9e1d4.pdf

EuroNCAP. (2008). Assessment Protocol and Biomechanical Limits v4.2 Retrieved from http://www.euroncap.com/files/Euro NCAP Assessment Protocol Version 4.2 June 2008---5ab781cd-4e7c-4542-b6f7-1d46733289ef.pdf

EuroNCAP. (2009). Frontal Impact Testing Protocol, version 4.3 Retrieved from http://www.euroncap.com/files/Euro-NCAP-Frontal-Protocol-Version-4.3---b8b12883-bd24-4b6e-b663-7898cda98e56.pdf

Eurostat (2009). Panorama of Transport - 2009 edition Luxembourg: Office for Official Publications of the European Communities.

Evans, L. (1991). *Traffic safety and the driver*. New York: Van Nostrand Reinhold.

Eysenck, M. W. (2004). *Psychology : an international perspective*. Hove: Psychology.

Faber, F., Jonkers, E., Ljung Aust, M., Benmimoun, M., Regan, M. A., Jamson, S. *et al.*, (submitted). Analysis methods for user related aspects and impact assessment on traffic safety, traffic efficiency and environment. *Deliverable 6.2 of European Large-Scale Field Operational Test on Active Safety Systems (euroFOT) project.* . Brussels: European Commission.

FESTA-Consortium. (2008). FESTA Handbook Version 2 *Deliverable T6.4 of the Field opErational teSt supporT Action (FESTA)*. Brussels: European Commission.

Field, A. (2009). *Discovering statistics using SPSS : (and sex and drugs and rock 'n' roll)*. Los Angeles: SAGE.

Fleury, D., and Brenac, T. (2001). Accident prototypical scenarios, a tool for road safety research and diagnostic studies. *Accident Analysis & Prevention, 33*(2), 267-276.

General-Motors-Corporation. (1997). 44 Crashes, Version 3.0. Warren, MI: North American Operations, Engineering Safety and Restraints Center, Crash Avoidance Department.

Gordon, T., Sardar, H., Blower, D., Ljung Aust, M., Bareket, Z., Barnes, M. *et al.* (2010). Advanced Crash Avoidance Technologies (ACAT) Program – Final Report of the Volvo-Ford- UMTRI Project: Safety Impact Methodology for Lane Departure Warning – Method Development And Estimation of Benefits.

Green, M. (2000). How long does it take to stop? Methodological analysis of driver perception-brake times. *Transportation Human Factors, 2*(3), 195-216.

Haddon, W. (1968). Changing approach to epidemiology prevention and amelioration of trauma - transition to approaches etiologically rather than descriptively based. *American Journal of Public Health and the Nations Health, 58*(8).

Haddon, W. (1973). Energy damage and 10 countermeasure strategies. *Journal of Trauma-Injury Infection and Critical Care, 13*(4), 321-331.

Hatakka, M., Keskinen, E., Gregersen, N. P., Glad, A., and Hernetkoski, K. (2002). From control of the vehicle to personal self-control; broadening the perspectives to driver education. *Transportation Research Part F: Traffic Psychology and Behaviour, 5*(3), 201-215.

Hayhoe, M. M., and Rothkopf, C. A. (2011). Vision in the natural world. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(2), 158-166.

Heinrich, H. W. (1931). *Industrial accident prevention : a scientific approach* (1. ed.). New York: McGraw-Hill.

Hickman, H. B. (2010). *Distraction in commercial trucks and buses: Assessing prevalence and risk in conjunction with crashes and near-crashes.*

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive energetical framework. *Biological Psychology 45*, 73-93.

Hollnagel, E. (2004). *Barriers and accident prevention*. Burlington, VT: Ashgate.

Hollnagel, E., Nåbo, A., and Lau, I. (2003). *A Systemic Model for Driver-in-Control*. Paper presented at the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Park City, Utah.

Hollnagel, E., and Woods, D. D. (2005). *Joint cognitive systems : foundations of cognitive systems engineering*. Boca Raton, FL: CRC Press.

Huang, Y.-H. (2005). *A systemic traffic accident model*. Linköping: Department of Computer and Information Science, Linköpings universitet.

Huang, Y.-H. (2007). *Having a new pair of glasses : applying systemic accident models on road safety*. Linköping University, Linköping. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-8189

Hulst, M. V. D., Rothengatter, T., and Meijman, T. (1998). Strategic adaptations to lack of preview in driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 1*(1), 59-75.

Jamson, A. H., Lai, F. C. H., and Carsten, O. M. J. (2008). Potential benefits of an adaptive forward collision warning system. *Transportation Research Part C: Emerging Technologies, 16*(4), 471-484.

Jamson, A. H., and Merat, N. (2005). Surrogate in-vehicle information systems and driver behaviour: Effects of visual and cognitive load in simulated rural driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 8*(2 SPEC. ISS.), 79-96.

Janssen, W. H., Nodari, E., Brouwer, R. F. T., Plaza, J., Östlund, J., Keinath, A. *et al.* (2008). Specification of AIDE methodology.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, N.J.: Prentice-Hall.

Karabatsou, V., M. Pappas, et al. (2006). A-priori evaluation of safety functions effectiveness - Methodologies. D4.1.3. TRACE.

Kiefer, R., LeBlanc, D., Palmer, M., Salinger, J., Deering, R., and Shulman, M. (1999). Development and validation of functional definitions and evaluation procedures for collision warning/avoidance systems (pp. 736 p.): Crash Avoidance Metrics Partnership, National Highway Traffic Safety Administration.

Kiefer, R., Shulman, M. A., Deering, R. K., Palmer, M. D., LeBlanc, D. J., Flannagan, C. A. C., *et al.* (2003). *Forward collision warning requirements project: refining the CAMP crash alert* *timing approach by examining "last-second" braking and lane change maneuvers under various kinematic conditions*.

Kiefer, R. J., Cassar, M. T., Flannagan, C. A., Jerome, C. J., and Palmer, M. D. (2005). *Surprise Braking Trials, Time-to-Collision Judgments, and "First Look" Maneuvers Under Realistic Rear-End Crash Scenarios*: Crash Avoidance Metrics Partnership, National Highway Traffic Safety Administration.

Knapp, A., and Schwarz, J. (2006). D11.2 Code of Practice for the Design and Evaluation of ADAS. Retrieved from http://www.prevent-ip.org/download/deliverables/RESPONSE3/D11.2/PR-11300-SPD-061031-v30-CoP.pdf

Korner, J. (1989). A Method for Evaluating Occupant Protection by Correlating Accident Data with Laboratory Test Data.

Kramer, A. F., Cassavaugh, N., Horrey, W. J., Becic, E., and Mayhugh, J. L. (2007). Influence of age and proximity warning devices on collision avoidance in simulated driving. *Human Factors, 49*(5), 935-949.

Kussmann, H., Modler, H., Engström, J., Agnvall, A., Piamonte, P., Markkula, G., *et al.* (2005). Requirements for AIDE HMI and safety functions.

Larsen, L. (2004). Methods of multidisciplinary in-depth analyses of road traffic accidents. *Journal of Hazardous Materials, 111*, 115-122.

Larsson, P., Dekker, S. W. A., and Tingvall, C. (2010). The need for a systems theory approach to road safety. *Safety Science, 48*(9), 1167-1174.

LeBlanc, D., Gordon, T. J., Goodsell, R., Bareket, Z., Hagan, M. R., Buonarosa, M. L. *et al.*, (2006*a*). Road departure crash warning system field operational test: methodology and results. Volume 1: technical report. 307 p.

LeBlanc, D., Gordon, T. J., Goodsell, R., Bareket, Z., Hagan, M. R., Buonarosa, M. L. *et al.* (2006*b*). Road departure crash warning system field operational test: methodology and results. Volume 2: appendices. 358 p.

Lee, J. D., McGehee, D. V., Brown, T. L., and Reyes, M. L. (2002). Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. *Human Factors, 44*(2), 314-334.

Leveson, N. (2004). A new accident model for engineering safer systems. *Safety Science, 42*(4), 237-270.

Ljung Aust, M., Fagerlind, H., and Sagberg, F. (2012) “Fatal Intersection Crashes In Norway: Patterns In Contributing Factors And Data Collection Challenges”. Accident Analysis and Prevention, 45, 782-791. doi: 10.1016/j.aap.2011.11.001.

Ljung Aust, M., and Engström, J. (2011). A conceptual framework for requirement specification and evaluation of active safety functions. *Theoretical Issues in Ergonomics Science, 12*(1), 44 - 65.

Ljung Aust, M., Regan, M. A., and Benmimoun, M. (2011). *Disentangling the Effects of Advanced Driver Assistance System Functions in Field Operational Tests: Recommendations from the European EuroFOT Project*. Paper presented at the 8th European Congress and Exhibition on Intelligent Transport Systems and Services, Lyon, France.

Ljung, M., Blommer, M., Curry, R., Artz, B., Greenberg, J., Kochhar, D. *et al.* (2007). *The influence of study design on results in hmi testing for active safety*. Paper presented at the 20th International Technical Conference on the Enhanced Safety of Vehicles, Lyon, France.

Ljung Aust, M. (2010). "Generalization Of Case Studies In Road Traffic When Defining Pre-Crash Scenarios For Active Safety Function Evaluation". Accident Analysis & Prevention, 42, 1172-1183.
Ljung, M., Fagerlind, H., Lövsund, P., and Sandin, J. (2007). Accident investigations for active safety at CHALMERS - new demands require new methodologies. *Vehicle System Dynamics, 45*(10 ), 881 - 894.
Maile, M., Neale, V., Ahmed-Zaid, F., Basnyake, C., Caminiti, L., Doerzaph, Z. *et al.* (2008). Cooperative intersection collision avoidance system limited to stop sign and traffic signal violations (CICAS-V).
Maltby, D., Morello, S., Perrett, K., and Hopkin, J. (1998). Checklist for Preparing a Validation Plan: Updated Version.
Martens, M., and Micah, R. J. F. (2007). Do familiarity and expectations change perception? Drivers' glances and response to changes. *Transportation Research Part F, 10*, 476-492.
Martens, M. H. (2011). Change detection in traffic: Where do we look and what do we perceive? *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(3), 240-250.
Maslow, A. H. (1954). *Motivation and personality*. New York: Harper & Row.
Maslow, A. H. (1970). *Motivation and personality* (2. ed.). New York: Harper & Row.
McCarthy, M., Fagerlind, H., Heinig, I., Langner, T., Heinrich, S., Sulzberger, L. *et al.*, (2010). ASSESS D1.1 - FINAL - Preliminary_Test_Scenarios_2010-04-06.pdf: EUROPEAN COMMISSION DG RTD.
McRuer, D. T., Allen, R. W., Weir, D. H., and Klein, R. H. (1977). New results in driver steering control models. *Human Factors, 19*(4), 381-397.
Michon, J. A. (1985). *Critical view of driver behavior models: what do we know, what should we do?*, Warren, MI, USA.
Midtland, K., Muskaug, R., Sagberg, F., and Jørgensen, N. O. c. (1995). Evaluation of the In-depth Accident Investigations of the Swedish National Road Administration.
Näätänen, R., and Summala, H. (1976). *Road-user behavior and traffic accidents*. Amsterdam ; New York,. Which Factors and Situations for Human Functional Failures? - Developing Grids for Accident Causation Analysis (2007).
Najm, W. G., daSilva, M. P., and Wiacek, C. J. (2000). *Estimation of Crash Injury Severity Reduction for Intelligent Vehicle Safety Systems*. Paper presented at the SAE 2000 World Congress, Detroit, MI.
Najm, W. G., Koopmann, J., Boyle, L., and Smith, D. (2002). Development of test scenarios for off-roadway crash countermeasures based on crash statistics (pp. 82 p.): Volpe National Transportation Systems Center
National Highway Traffic Safety Administration.
Najm, W. G., Mironer, M. S., and Fraser, L. C. (1995). *Analysis of target crashes and its countermeasure actions*: ITS America.
Najm, W. G., and Smith, J. D. (2007). Development of Crash Imminent Test Scenarios for Integrated Vehicle-Based Safety Systems (pp. 57p): Volpe National Transportation Systems Center, National Highway Traffic Safety Administration.
Najm, W. G., Smith, J. D., and Yanagisawa, M. (2007). Pre-Crash Scenario Typology for Crash Avoidance Research Washington U.S. Department of Transportation, National Highway Traffic Safety Administration.
Najm, W. G., Stearns, M. D., Howarth, H., Koopmann, J., and Hitz, J. (2006). *Evaluation of an Automotive Rear-End Collision Avoidance System*: Research and Innovative Technology

Administration, Volpe National Transportation Systems Center, National Highway Traffic Safety Administration.
NCSA. (2003). Traffic Safety Facts: a compilation of motor-vehicle crash-data from FARS and GES (U. D. o. Transportation, Trans.). Washington, D.C.: National Center for Statistics and Analysis of the National Traffic Highway Safety Administration.
NCSA. (2005). Traffic Safety Facts 2004 - A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System (U. D. o. Transportation, Trans.). Washington, D.C.: National Center for Statistics and Analysis of the National Traffic Highway Safety Administration.
NCSA. (2006). Traffic Safety Facts 2005 - A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System (U. D. o. Transportation, Trans.). Washington D.C.: National Center for Statistics and Analysis of the National Traffic Highway Safety Administration.
NCSA. (2007). Traffic Safety Facts 2006 - A Compilation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System (U. D. o. Transportation, Trans.). Washington D.C.: National Center for Statistics and Analysis of the National Traffic Highway Safety Administration.
Nikolaev, A. G., Robbins, M. J., and Jacobson, S. H. (2010). Evaluating the impact of legislation prohibiting hand-held cell phone use while driving. *Transportation Research Part A: Policy and Practice, 44*(3), 182-193
Nilsson, G. (2004). Trafiksäkerhetsåtgärder och efterlevnad: hastighetsanpassning, användning av bilbälte och nykter som bilförare. Linköping.
OECD. (1988). Road accidents: on-site investigations. Paris.
OECD. (1990). Behavioural Adaptations to Changes in the Road Transport System. *Road Transport Research Programme*.
Olson, R. L. (2009). *Driver distraction in commercial vehicle operations*. Washington, D.C.: Federal Motor Carrier Safety Administration.
The Oxford English Dictionary - OED Online. (1989, 2nd ed. ed.). Oxford University Press.
Perrow, C. (1984). *Normal accidents : living with high-risk technologies*. New York: Basic Books.
Ranney, T. A. (1994). Models of driving behavior: A review of their evolution. *Accident Analysis & Prevention [ACCID. ANAL. PREV.]. Vol., 26*(6), 733-750.
Räsänen, M., and Summala, H. (1998). Attention and expectation problems in bicycle-ìcar collisions: an in-depth study. *Accident Analysis & Prevention, 30*(5), 657-666.
Rasmussen, J. (1990). Human error and the problem of causality in analysis of accidents. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 327*(1241), 449-462.
Rauch, N., Gradenegger, B., and Krüger, H. P. (2008). User strategies for the interaction with in-vehicle devices while driving. *IET Intelligent Transport Systems, 2*(4), 266-275.
Reason, J. (1990). *Human error*. Cambridge: Cambridge Univ. Press.
Reason, J. (1997). *Managing the risks of organizational accidents*. Aldershot, Hants ; Brookfield, Vt.: Ashgate.
Reason, J., Hollnagel, E., and Paries, J. (2006). Revisiting the "Swiss cheese" model of accidents. Brétigny-sur-Orge, France: EUROCONTROL Experimental Centre.
Rochlin, G. I. (1999). Safe operation as a social construct. *Ergonomics, 42*(11), 1549-1560.
Rosch, E. (1975*a*). Cognitive reference points. *Cognitive Psychology, 7*(4), 532-547.
Rosch, E. (1975*b*). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*(3), 192-233.

Rosch, E., and Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*(4), 573-605.
Runyan, C. W., and Yonas, M. (2008). Conceptual Frameworks for Developing and Comparing Approaches to Improve Adolescent Motor-Vehicle Safety. *American Journal of Preventive Medicine, 35*(3, Supplement), S336-S342.
Sabey, B. E. (1990). Accident Analysis Methodology. *IATSS Research, 14*(1), 35-42.
SAFETYNET. (2005*a*). Annual Statistical Report 2005.
SAFETYNET. (2005*b*). Deliverable 5.2: In-depth Accident Causation Data Study Methodology Development Report: European Road Safety Observatory.
SAFETYNET. (2007*a*). Annual Statistical Report 2006: SafetyNet.
SAFETYNET. (2007*b*). Annual Statistical Report 2007 *Workpackage 1, Task 3*.
SAFETYNET. (2008). Deliverable 5.8: In-depth Accident Causation Database and Analysis Report Retrieved from http://www.erso.eu/safetynet/fixed/WP5/SN_D5_8_In-Depth_accident_causation_database_and_analysis_report.pdf doi:http://erso.swov.nl/safetynet/fixed/WP5/SN_D5_8_In-Depth_accident_causation_database_and_analysis_report.pdf
Sagberg, F., Fosser, S., and Sætermo, I. A. F. (1997). An investigation of behavioural adaptation to airbags and antilock brakes among taxi drivers. *Accident Analysis & Prevention, 29*(3), 293-302.
Sandin, J., and Ljung, M. (2007). Understanding the causation of single-vehicle crashes: a methodology for in-depth on-scene multidisciplinary case studies. *International Journal of Vehicle Safety 2007, 2*(3 ), 316 - 333.
Sayer, J., LeBlanc, D., Bogard, S. E., Funkhouser, D., Bao, S., Buonarosa, M. L. *et al.* (2011). Integrated vehicle-based safety systems field operational test final program report. Retrieved from http://hdl.handle.net/2027.42/84378
SCB (2012). Befolkningsstatistik 2011. Available from http://www.scb.se/Pages/TableAndChart_262459.aspx
Scholliers, J., Blosseville, J.-M., Netto, M., Mammar, S., Chen, J., Heinig, K. *et al.* (2007*a*). D16.1 Review of validation procedures for preventive and active safety functions. Retrieved from http://www.prevent-ip.org/download/deliverables/PReVAL/PR-16000-SPD-070611-v20-D16_1_RewiewOfProcedures.pdf
Scholliers, J., Heinig, K., Blosseville, J., Netto, M., Anttila, V., Leanderson, S. *et al.* (2007*b*). D16.3 Proposal of procedures for assessment of preventive and active safety functions. Retrieved from http://www.prevent-ip.org/download/deliverables/PReVAL/PR-16000-SPD-070831-v110-D16_3.pdf
Scholliers, J., Hendriks, F., Aust, M. L., Luoma, J., Netto, M., Engström, J. *et al.* (2008*a*). An integrated evaluation framework for preventive safety applications. 85-90.
Scholliers, J., Kutzner, R., Luoma, J., Sihvola, N., Leanderson, S., Netto, M. *et al.* (2008*b*). Project final report and recommendations for future assessments. Final Report of PReVENT IP subproject PReVAL.
Scott, J. J., and Gray, R. A. (2008). Comparison of Tactile, Visual, and Auditory Warnings for Rear-End Collision Prevention in Simulated Driving. *Human Factors, 50*(2), 264-275(212).
Scriven, M. (1967). The methodology of evaluation. In R. E. Stake (Ed.), *Curriculum evaluation*. Chicago: Rand McNally.
SeMiFOT. (2009). Sweden Michigan Naturalistic Field Operational Test Retrieved from https://http://www.chalmers.se/safer/EN/projects/traffic-safety-

analysis/semifot/downloadFile/attachedFile_f0/Folder_SeMiFOT_project?nocache=1237896326.73
Shinoda, H., Hayhoe, M. M., and Shrivastava, A. (2001). What controls attention in natural environments? *Vision Research, 41*(25-26), 3535-3545.
Shulman, M., and Deering, R. K. (2004). *Second annual report of the crash avoidance metrics partnership, April 2002-March 2003*: Crash Avoidance Metrics Partnership, National Highway Traffic Safety Administration.
Shulman, M., and Deering, R. K. (2005). *Third Annual Report of the Crash Avoidance Metrics Partnership, April 2003 - March 2004*: Crash Avoidance Metrics Partnership, National Highway Traffic Safety Administration.
Siegrist, S. (2010). Towards a method to forecast the effectiveness of national road safety programmes. *Safety Science, 48*(9), 1106-1110.
SIKA. (2006*a*). Road traffic injuries 2005 *Official Statistics of Sweden*. Stockholm: Swedish institute for transport and communication analysis (SIKA).
SIKA. (2006*b*). Vägtrafikskador 2005. Utvidgat tabellverk enligt äldre publicering. Stockholm: Swedish institute for transport and communication analysis (SIKA).
SIKA. (2007). Körsträckor 2007 *Official Statistics of Sweden*. Stockholm: Swedish institute for transport and communication analysis (SIKA).
Stevens, A., Board, P., A., and Quimby, A. (1999). A Safety Checklist for the Assessment of in-Vehicle Information Systems: Scoring Proforma, Project Report PA3536-A/99, Transport Research Laboratory, (Crowthorne, UK.
Summala, H. (2007). Towards Understanding Motivational and Emotional Factors in Driver Behaviour: Comfort Through Satisficing. In P. C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments: critical issues in driver interactions with intelligent transport systems* (pp. 189-207). London: Springer.
Summala, H., Pasanen, E., Räsänen, M., and Sievvänen, J. (1996). Bicycle accidents and drivers' visual search at left and right turns. *Accident Analysis -& Prevention, 28*(2), 147-153. doi: 10.1016/0001-4575(95)00041-0
Tatler, B. W., Hayhoe, M. M., Land, M. F., and Ballard, D. H. (2011). Eye guidance in natural vision: reinterpreting salience. *Journal of vision, 11*(5), 5.
Transportstyrelsen (2012). Statistik över körkortsinnehavare efter åldergrupp. Available from http://www.transportstyrelsen.se/sv/Press/Statistik/Vag/Korkort/Statistik-over-korkortsinnehavare-efter-aldersgrupp/
Trick, L. M., and Enns, J. T. (2009). A two-dimensional framework for understanding the role of attentional selection in driving. In C. Castro (Ed.), *Human Facors of Visual and Cognitive Performance in Driving* (pp. 63-73). Boca Raton, Florida: CRC press.
Vaa, T. (2007). Modelling Driver Behaviour on Basis of Emotions and Feelings: Intelligent Transport Systems and Behavioural Adaptations. In P. C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments: critical issues in driver interactions with intelligent transport systems* (pp. 208-232). London: Springer.
Van Elslande, P., and Fouquet, K. (2007). Analyzing 'human functional failures' in road accidents *Deliverable 5.1*: TRACE.
Wallén Warner, H., Björklund, G., Johansson, E., Ljung Aust, M., and Sandin, J. (2008*a*). DREAM 3.0 - documentation of references supporting the links in the classification scheme. Chalmers University of Technology Applied Mechanics.
Wallén Warner, H., Ljung Aust, M., Björklund, G., Johansson, E., and Sandin, J. (2008*b*). Manual for DREAM 3.0, Driving Reliability and Error Analysis Method *Deliverable D5.6 of the EU FP6 project SafetyNet.* Brussels: European Commission.

Wetzel, L. (Winter Edition 2007). Types and Tokens. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*.
Wickens, C. D. (2002). Multiple Resources and Performance Prediction. *Theoretical Issues in Ergonomic Science*.
Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50*(3), 449-455.
Wilde, G. J. S. (1982). The theory of risk homeostasis: Implications for safety and health. *Risk Analysis, 2*(4), 209-225.
Wisch, M., Fagerlind, H., Sulzberger, L., McCarthy, M., Roynard, M., Hulshof, W. *et al.* (2010). ASSESS D1.2 - Specifications for scenario definitions: European Commission DG RTD.
Woltjer, R. (2009). *Functional Modeling of Constraint Management in Aviation Safety and Command and Control.* Linköpings universitet, Linköping. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-17639
Woods, D. D., and Hollnagel, E. (2006). *Joint cognitive systems : patterns in cognitive systems engineering*. Boca Raton: CRC/Taylor & Francis.
Zhang, X., Kompfner, P., White, C., and Sexton, B. (1998). Guidebook for Assessment of Transport Telematics Applications: Updated Version: Converge consortium.