Aggregating Case Studies of Vehicle Crashes by Means of Causation Charts

An Evaluation and Revision of the Driving Reliability and Error Analysis Method

JESPER SANDIN

Department of Applied Mechanics
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2008
THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

AGGREGATING CASE STUDIES OF VEHICLE CRASHES BY MEANS OF CAUSATION CHARTS

AN EVALUATION AND REVISION OF THE DRIVING RELIABILITY AND ERROR ANALYSIS METHOD

JESPER SANDIN

VEHICLE SAFETY DIVISION
DEPARTMENT OF APPLIED MECHANICS
CHALMERS UNIVERSITY OF TECHNOLOGY

GÖTEBORG, SWEDEN 2008
AGGREGATING CASE STUDIES OF VEHICLE CRASHES BY MEANS OF CAUSATION CHARTS
AN EVALUATION AND REVISION OF THE DRIVING RELIABILITY AND ERROR ANALYSIS METHOD

JESPER SANDIN

© Jesper Sandin, 2008

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 2849
ISSN 0346-718X

Vehicle Safety Division
Department of Applied Mechanics
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone +46 (0)31 772 1000
URL www.chalmers.se

Printed at Chalmers Reproservice,
Göteborg, Sweden
Till Anna
Abstract

There is a need for increased knowledge about causes to motor-vehicle crashes and their prevention. Multidisciplinary in-depth case studies can provide detailed causation data that is otherwise unattainable. Such data might allow the formulation of hypotheses of causes and causal relationships for further study. By converting the data into causation charts that are aggregated, common causation patterns would give greater weight to such hypotheses. However the charts must first be compiled by means of a systematic analysis method, which requires three parts; a model, a classification scheme and a classification method.

Four general accident models were evaluated and found inadequate to form the basis for a causation analysis method. This was primarily because the models in practice treat road-users, vehicles and traffic environment as separate components, but also due to the focus on events immediately prior to the crash and either static, sequential, or absent modelling of interaction.

Two studies were carried out to evaluate whether case files could be aggregated by means of charts that had been compiled with the Driving Reliability and Error Analysis Method (DREAM). In DREAM, contributory factors (genotypes) are systematically analysed, classified and linked in a single chart for each driver that illustrate the causes of a critical event (phenotype). In the first study, case files from 38 single-vehicle crashes were examined to distinguish crashes with similar circumstances. Four types of loss of vehicle control were identified, for which the associated DREAM charts were aggregated. The results revealed common patterns within the types, as well as different patterns between them. The second study focused on 26 intersection crashes. Based on the most common violations at intersections, six risk situations were defined, and the DREAM charts associated with each risk situation were aggregated. A common pattern in each of two risk situations indicated that drivers with and without the right of way had not seen the other vehicle due to distractions and/or sight obstructions. A frequently occurring pattern for the drivers with the right of way was that they had not expected another vehicle to cross their path. The absence of clear patterns in three risk situations was a result of a low number of charts and rather unique circumstances in these cases. Parts of the aggregated charts contained an unexpectedly large variation, identified as a consequence of inconsistently compiled charts.

Prior the final study assessing intercoder agreement, DREAM was revised into a new version based on the experience from the latter aggregation study. A total of seven investigators from four European countries compiled seven DREAM charts for each driver involved in four types of accidents. The results indicated that the intercoder agreement for genotypes ranged from 74% to 94% with an average of 83%, while it for phenotypes ranged from 57% to 100% with an average of 78%. This acceptable level of agreement is expected to rise with enhanced training. The present thesis thus shows that DREAM is a highly promising method for the compilation of causation charts. Future studies are expected to benefit from aggregating DREAM charts when formulating hypotheses of general causes and causal relationships as a subject for further research, as well as to identify alternative countermeasure strategies.

Keywords: driver error, collision avoidance, pre-crash, fatigue, slipperiness, rollover, young drivers, vehicle dynamics
List of papers
The present thesis comprises the following appended papers, referred to in text in italic and with their Roman numerals:


Contribution: Huang initiated the study and wrote it together with Ljung, Sandin and Hollnagel. Sandin contributed to the evaluation of the general accident models and the examples of traffic accident models.


Contribution: Sandin initiated and wrote the paper. As two out of three members of the investigation team, Sandin and Ljung carried out the accident investigations. The DREAM charts were compiled by Ljung and aggregated and analysed by Sandin. Sandin compared the findings with previous studies.


Contribution: Sandin initiated and wrote the paper. As one out of three members of the investigation team, Sandin carried out the accident investigations. The DREAM charts were updated by three additional investigators and aggregated and analysed by Sandin. Sandin compared the findings with previous studies.


Contribution: Wallén Warner initiated and wrote the paper together with Sandin. Wallén Warner and Sandin planned the study, wrote the accident scenarios, trained the investigators, and analysed and interpreted the results.
# Table of contents

Abstract ................................................................................................................................. i
List of papers .......................................................................................................................... ii
Table of contents ................................................................................................................... iii
Acknowledgements ............................................................................................................. v

## 1 Introduction

1.1 Road traffic accidents: the need for new knowledge and measures .......................... 1
1.2 The value of multidisciplinary in-depth accident investigations .............................. 2
1.3 The traditional way of analysing causation in case studies .................................... 3
1.4 An alternative approach: aggregating causation charts .......................................... 4
1.5 Methods for the compilation of causation charts ..................................................... 5

## 2 Objectives ...................................................................................................................... 11

## 3 The Driving Reliability and Error Analysis Method

3.1 The accident and driver models .................................................................................. 13
3.2 The classification scheme ........................................................................................... 14
3.3 The method .................................................................................................................. 16

## 4 Summary of Papers I, II, III, and IV

4.1 Summary of Paper I ..................................................................................................... 21
4.2 Summary of Papers II and III .................................................................................... 22
4.3 Summary of Paper IV ................................................................................................. 25

## 5 Addendum: Revision of DREAM

5.1 Revision of the phenotypes ......................................................................................... 27
5.2 Revision of the genotypes ........................................................................................... 27
5.3 Revision of the links .................................................................................................... 29

## 6 General discussion

6.1 Two purposes behind aggregating case files by means of causation charts .......... 31
6.2 Relations between models, classification scheme and classification method ....... 32
6.3 Different causation charts for the risk and emergency phases .............................. 34
6.4 Experiences gained from aggregating DREAM charts .......................................... 35
6.5 Data collection in multidisciplinary in-depth accident investigations .................. 38
6.6 Reliability in the classification of qualitative material ............................................. 39
6.7 Further research: identifying countermeasures by means of causation charts ....... 42

## 7 Conclusions .................................................................................................................... 43

## 8 References ..................................................................................................................... 44
Acknowledgements

I would like to thank Professor Per Lövsund and Adjunct Professor Hans Norin at the Department of Vehicle Safety for accepting me as a PhD student and for their endeavours to guide my research in the chaos of causation. I give my sincere thanks to Adjunct Professor Hans-Erik Pettersson for sharing your knowledge in the field of multidisciplinary accident investigations and driver behaviour, for constructive discussions and scientific guidance.

Many thanks go to the rest of the “first” multidisciplinary accident investigation team at the Department of Vehicle Safety, Anders Flogård and Mikael Ljung Aust, who have contributed considerably to the knowledge described in this work.

The reference group responsible for the DREAM 3.0 revision, in addition to aforementioned Mikael, Gunilla Björklund (PhD), Johan Engström, and Emma Johansson, are highly appreciated for inspiring, and sometimes frustrating, discussions. I thank especially the “Head” of the reference group, Henriette Wallén Warner (PhD), for inspiring methodological and scientific discussions concerning the last paper in the present thesis.

Yu-Hsing Huang and Erik Hollnagel at the Department of Computer and Information Science at Linköping University are acknowledged for valuable cooperation.

The members of Autoliv Research, Saab Automobile, Volvo Car Corporation, Volvo AB participating in the workshops in phases 2 and 3 of the FICA project are acknowledged for valuable input and progressing discussions.

Although we have not always agreed, I would like to give my grateful thanks to the scientific reference group for their well-meant criticism during this work, Adjunct Professor Emeritus Anders Englund, Professor András Várhelyi, and aforementioned Adjunct Professor Hans-Erik Pettersson.

The Swedish Vehicle Research Program (PFF) administered by VINNOVA, Sweden, Autoliv Research, Swedish Road Administration (Vägverket), Saab Automobile, Volvo Car Corporation, Volvo AB, are greatly acknowledged for their financial support. A special thank goes to the project leader of the project Factors Influencing the Causation of Accidents and Incidents (FICA), Håkan Gustafsson.

I thank everyone at the Department of Applied Mechanics and at SAFER – Vehicle and Traffic Safety Centre at Chalmers, for lots of laughter in the lunch room when discussing a diversity of subjects. A special thank to Vehicle Safety SC - it is a long time since we met in the heat.

I give my sincere thanks and appreciation to my parents for their ability to bring matters down to earth and for providing me with abilities that you cannot acquire in school. If it were not for my brother, Fredrik Sandin, this thesis would likely not have been written. Fredrik persuaded me at the last minute to choose technology instead of economics as my study programme in upper secondary school. As you told me to expect, there have been many parties but, more significantly, there has been a great deal to learn and explore.

Last but not least, thank you, Anna, for your love and encouragement in moments of despair. You are always on my mind.
1 Introduction

1.1 Road traffic accidents: the need for new knowledge and measures

Road traffic injuries are a huge public health problem all over the world. The WHO World Report on Road Traffic Injury Prevention (Peden et al., 2004) estimated that road traffic accidents kills almost 1.2 million people annually as well as injuring or disabling between 20 and 50 million. Without appropriate action, these injuries will rise dramatically by 2020, particularly in countries that are rapidly becoming motorized. Peden et al. (2004) conclude that road safety in developing countries would benefit from adopting the basic principles that have led to a sharp reduction in crashes and casualty numbers in many high-income countries over recent decades. These basic principles include good road design and traffic management, improved vehicle standards, speed control, the use of seat-belts and the enforcement of alcohol limits.

In developed countries that have been motorised for decades, there is a need to find new measures to further reduce the number of casualties. Preferably by means of crash prevention, because “A crash prevention measure that reduces crash risk by some percent is necessarily a far more effective intervention than a crashworthiness measure with the same percent effectiveness” (Evans, 2004, p. 9). As the reduction of casualties is of major concern, preventing single-vehicle and intersection crashes, two of the most common and serious crash types, would be beneficial.

Single-vehicle crashes are the largest contributor to serious crashes, accounting for about half of all fatal and a third of all serious injury motor-vehicle crashes (Collin, 2000; NCSA, 2005; SIKA, 2005). Despite the fact that vehicle structures and passive safety features have been significantly improved during recent decades, single-vehicle crashes are still a major problem. Due to their consequences and high societal costs, a great deal of attention has been devoted to single-vehicle crashes with serious injuries. However, cost estimates for less serious and material damage crashes show that they also result in high total costs because of their frequency (Blincoe et al., 2002).

Intersection crashes are generally recognised as the second largest type of motor accidents (ERSO, 2006a; IATSS, 2005; NCSA, 2005). In European countries, more than half of the fatalities in intersection crashes involve the occupants of motor vehicles (ERSO, 2006a, b). In the United States, intersection crossing-path crashes between motor vehicles account for approximately 25% of all police-reported crashes annually, with consequences ranging from fatalities to material damage. They also account for 27% of all delays caused by crashes (Lee et al., 2004). According to Blincoe et al. (2002) and Lee et al. (2004), these consequences lead to an estimated average cost per crossing-path crash of $28,209.

These examples demonstrate that crash prevention would result in a reduction not only of casualties but also of private and societal costs as well as traffic congestion. Nowadays, there are high expectations on vehicle-mounted active safety measures (Lee et al., 2004; Najm et al., 2001; Vahidi and Eskandarian, 2003). So far, such measures have mainly been driven by engineers and technological development (Donges, 1999; Peden et al., 2004). The focus has been on collision avoidance systems that can improve occupant safety and reduce losses by preventing accidents that are beyond the control of the driver (Deering and Viano, 1994; Kawai, 1994; Vahidi and Eskandarian, 2003). For a higher level of driver warning and assistance systems, Kawai (1994), Shladover (1995) and Vahidi and Eskandarian (2003) recognise a need for extensive research of human factors in order to obtain a deeper understanding of drivers’ psychology and behavioural habits for the design of interactive
systems (Vahidi and Eskandarian, 2003). One obvious part of this research is the investigation and analysis of road traffic accidents in terms of causation.

Investigations of road traffic accidents focusing on causation have been conducted since shortly after the introduction of the motor vehicle. Between 1930 and 1950, the accident research focused upon the driver and driving behaviour (Grayson and Hakkert, 1987). During this period, the idea of accident prone drivers emerged, a concept that has frequently been refuted (McKenna, 1983), but which still seems to persist (Visser et al., 2007). During the 1960s accident causation research shifted to view traffic as a human-vehicle-environment system, where the human was seen as the bottle-neck or weak link in an over-demanding traffic environment (Englund et al., 1998). During the 1970s and ‘80s, road-users’ risk perception became the subject of attention, especially the debate on the risk homeostasis theory (Wilde, 1982, 1988). While this theory has been criticised for lack of realism and scientific value (Evans, 2004; Ranney, 1994), the debate highlighted the importance of drivers’ motives (Näätänen and Summala, 1974; Summala, 1988). A more recent addition to the theories of road-user behaviour are hierarchical control models (Hale et al., 1990; Ranney, 1994), primarily Michon’s (1985) hierarchical control structure with its assumption of concurrent activity at strategic, manoeuvring, and operational levels of control, and Rasmussen’s (1983) Skill-Rule-Knowledge (SRK) framework.

Although numerous studies focusing on accident causation have been conducted, the technology available today opens new possibilities for preventing crashes (Shladover, 1995; Vahidi and Eskandarian, 2003). However, this is likely to necessitate new perspectives and knowledge of crash causation and driver behaviour. As early as 1988, OECD (1988) stated that there is a need to intensify road safety research in order to find new theoretical perspectives and effective countermeasures. The authors of the report further argued that multidisciplinary in-depth accident investigations are highly valuable for these purposes (OECD, 1988).

1.2 The value of multidisciplinary in-depth accident investigations

A number of studies have concluded that multidisciplinary in-depth accident investigations are highly valuable when exploring aspects of road traffic accidents, especially accident causation, where the complex interactions of factors related to the driver, vehicle and road environment are of interest (Grayson and Hakkert, 1987; OECD, 1988; Sabey, 1990). It is also acknowledged that because of the many people involved and the large amount of information collected, multidisciplinary in-depth investigations are time consuming and expensive, especially those carried out on-scene. This means that only a small number of accidents can be investigated in relation to the total number that occur (Grayson and Hakkert, 1987). Consequently, multidisciplinary in-depth studies usually take a clinical or case study approach. In the clinical approach, accident investigations and/or analyses comprise a limited number of accidents with common characteristics that address a particular research question (OECD, 1988). A case study refers to the in-depth investigation of a single accident (OECD, 1988), with the aim of describing how and why it occurred and the drawing of conclusions in each individual case (Englund et al., 1978).

Clinical studies and case studies are often compared to statistical studies (see Baker, 1960; Englund, 1978; Fleury et al., 1994; OECD, 1988; Treat et al., 1977a). In the latter, statistical methods are used to analyse coded accident information in databases of statistically representative accidents (OECD, 1988). The emphasis of “statistical” is on whether or not the accident sample is representative, even though the term can be used rather ambiguously and
occasionally refers to the analytical technique used and/or to a large collection of accidents (Englund, 1978).

Nevertheless, because statistical studies are expected to generate representative findings that can be generalised, the size of problems can be quantified (McKenna, 1982). However, it is widely acknowledged that in order to handle a large number of crashes, statistical studies have to limit the level of detail (Sabey, 1990), which greatly simplifies accident information, does not provide sufficient understanding of the coded accident circumstances (Grayson and Hakkert, 1987; Midtland et al., 1995) and cannot be used to identify interactions between several accident factors (Larsen, 2004).

The FERSI group (Fleury et al., 1994) concluded that in-depth accident investigations, even those conducted without aiming for representative findings (i.e. case studies), can serve several purposes such as to:

- provide documentation leading to the discovery of “new” problems and/or formulation of hypotheses, which can be tested experimentally or on representative accident material.
- supply detailed information concerning a phenomenon or a causal relationship which has previously been established on the basis of statistical material or as a result of laboratory experiments.
- furnish ideas and suggestions concerning measures of a general nature and indications for appropriate local measures.

Consequently, although in-depth multidisciplinary investigations are seldom representative, they are nevertheless valuable, particularly for providing information which is unattainable in any other way (Grayson and Hakkert, 1987). In order to make use of case studies, in particular in relation to analysis of causes and the formulation of hypotheses, it would be useful to aggregate case study files (Midtland et al., 1995). Midtland et al. (1995) recognised that formulating hypotheses on the basis of one case is inadequate, while two cases showing the same causes would be somewhat better, and several aggregated cases showing common causes would have an even greater value for the formulation of hypotheses. Due to the large amount of collected accident information, case files as such are not easily aggregated (Fleury and Brenac, 2001; Midtland et al., 1995) making the identification of common causes difficult.

1.3 The traditional way of analysing causation in case studies

Traditionally, the qualitative information in case studies is analysed using a standard statistical approach. First, the qualitative information is coded according to a classification scheme with operational definitions of contributory factors. Then, when all cases have been collected, the coded information is analysed using standard statistical techniques, typically calculating frequency distributions and conducting correlation analyses.

One of the first and well thought out classification schemes was developed by Baker (1960), Baker and Horn (1960) and Baker and Ross (1960), who evaluated it by means of an analysis of 42 in-depth investigations of accidents. Grayson and Hakkert (1987) reviewed nine other studies conducted during the 1960s and 70s using a standard statistical approach. With the introduction of computers during the 1970s, larger samples of accidents could be analysed in terms of contributory factors. Two widely referenced larger studies were conducted during the 1970s in the UK (Sabey and Staughton, 1975; Staughton and Storie, 1977) and the US (Treat et al., 1977a). In the former, 2130 accidents were investigated and coded according to a
standard form with 400 items for each accident, where the items in the human errors category “were chosen to some extent arbitrarily but on the basis of past experience” (Sabey and Staughton, 1975, p. 6). In the US, Treat et al. (1977a) conducted case studies on the on-site level (2,258 accidents) and the in-depth level (420 accidents). The on-site investigations were conducted immediately after the accident by a team of technicians. The in-depth investigations were performed by a multidisciplinary team in order to analyse detailed causation mechanisms. The 420 accidents were selected by chance from accidents at the on-site level and investigated independently. On both levels, the contributory factors were coded according to a classification scheme divided into three major hierarchical causal factor trees for vehicle, environment and human causes. Regarding the Human direct causes, “The human factors part of the accident causation hierarchy is patterned after a stage model of human information processing consisting of at least three additive stages involving recognition, decision, and response.” (Treat et al., 1977b, p. 175)

Sabey and Staughton (1975) and Treat et al. (1977a) acknowledged that road traffic accidents are a result of the failed interaction between road-user, vehicle and traffic environment. In both studies however, when only one factor was identified, it was overwhelmingly the road user (65% in the UK study, 57% in the US study). Road user factors were found to be the sole or contributory factors in 94% of accident in the UK study, and 93% in the US study (Evans, 2004).

These conclusions have led to a great deal of discussion about the theoretical and methodological base of these studies as well as human errors in general (see e.g. Lourens, 1989, 1990; Patrick, 1987; Ranney, 1994; Rothengatter, 1987; Rumar, 1990; Sivak, 1981). Ranney (1994) argued that instead of focusing on the high percentage of human error, factors that create incompatibilities between drivers, vehicles and traffic environment should be identified.

1.4 An alternative approach: aggregating causation charts

The analysis of traffic accidents has much in common with the analysis of occupational accidents. In Rasmussen’s (1997) three hazard domains of socio-technical accidents, occupational accidents belong to the hazard domain of frequent, small scale accidents. In this domain, safety is typically monitored empirically, based on epidemiological studies of previous accidents (Rasmussen, 1997), employing coded information from case reports that are analysed using standard statistical techniques (Leplat and Rasmussen, 1987). Leplat and Rasmussen (1987) argued that the recommendations of such analyses are very general and difficult to implement.

Leplat and Rasmussen (1987) outlined an alternative approach to the analysis of occupational accidents, where measures could be identified with the help of causation charts in terms of “variation diagrams” (Leplat 1987, see sec. 1.5.2 below). Figure 1 present an example of a variation diagram of a single-vehicle crash involving a lorry (Leplat and Rasmussen, 1987, p. 159). The circles illustrate “nodes” representing variations from “normal”, and the occurrence of one node is the effect of an antecedent node. In identifying the need for measures, one principle could be to eliminate a variation node, either by changing the physical condition (fix the brakes), or changing the reason for a human act (better work conditions). Another principle could be to break the flow of events between the variation nodes, for example by informing humans concerned (maintenance personnel or drivers) so that risky situations are detected and corrected. Leplat and Rasmussen (1987) recognised that although variation diagrams may facilitate the identification of several measures that could break the sequence of events, it is inadequate to suggest them on the basis of one single accident case. Instead a
method should be developed for aggregating variation diagrams for similar accident sequences or similar working conditions. Aggregated diagrams would allow the identification of measures that have a recurrent effect, while measures that are only relevant to one single case could be rejected. In addition to the variation diagram method, other methods also make use of accident or causation charts.

![Variation diagram of a single-vehicle accident (from Leplat and Rasmussen 1987).](image)

**Figure 1: Variation diagram of a single-vehicle accident (from Leplat and Rasmussen 1987).**

### 1.5 Methods for the compilation of causation charts

A brief survey and description of accident analysis methods that make use of accident or causation charts is presented below. The survey is not intended to cover all analysis methods and begins with four categories of “chain-of-events” methods that are customary in the analysis of major accidents, typically airplane crashes, ferry accidents, train crashes and hotel fires. The survey then continues with three methods that have been suggested for the analysis of traffic accidents by means of causation charts or causal relationships. The survey ends with a description a method proposed for accident analysis and predictive risk-analysis in process industry and nuclear power-plants.

#### 1.5.1 Chain-of-Events methods

“Chain-of-events” methods (Leveson, 1995) are generally divided into four types (Figure 2); event tree and fault tree (NUREG-0492, 1981), single chain (Heinrich et al., 1980) and multilinear event sequence methods (Benner, 1975; Hendrick and Benner, 1987). The common feature of these methods is that they explain accidents in terms of multiple events sequenced as a chain over time, almost always involving some type of component failure (Leveson, 1995, 2004), which in the case of fault trees, and particularly event trees, are categorised in a binary way as either total success or total failure (Hollnagel, 1998). The events are chained together by means of connection rules, for example the AND/OR gates in fault tree diagrams (NUREG-0492, 1981) and the rule of “one actor + one action = one event” in multilinear event sequence methods (Benner, 1975; Hendrick and Benner, 1987).

In the investigation of major accidents, the chain-of-events methods are used to structure the analysis and data collection, as well as to illustrate the resulting accident sequence (Benner, 1985; Ferry, 1988; Hendrick and Benner, 1987; Sklet, 2002).
1.5.2 The Variation Diagram method

The variation diagram method (Leplat, 1987) was developed for the investigation of occupational accidents in the industrial domain (the INRS method, Leplat (1978)), but has also been used for vehicle accidents (Figures 1 and 3). An analysis by mean of the variation diagram method is conducted in two steps. First, an analyst identifies, ranks, and lists the variation nodes from the accident information and orders them in a list. Second, the listed nodes are organised into diagrams by following two connection rules; the event chain relationship and the confluence relationship (Figure 3). The event chain relationship indicates that if event X had not occurred, event Y would not have taken place. The confluence relationship states that in the absence of two independent events, X1 and X2, event Y would not have occurred (Leplat, 1987). Leplat (1987) acknowledged that the method focuses on events and “is based on execution performance which for humans lead to an analysis of behaviour rather than cognition”. According to the author, however, the diagrams could be extended with antecedent and latent factors. For example in Figure 3, factors such as tyre wear and a poorly balanced braking system could have contributed to the confluence relationship of sudden braking and wet ground leading to skidding. Generally however, latent factors are not represented in variation diagrams if they do not have a direct influence on the event (Leplat, 1987).

Figure 3. Example of a variation diagram for a single-vehicle crash. Unbroken lines: the chaining of various consequences of these actions (if successful) (from Leplat 1987).
1.5.3 The Accident Causal System

Fell (1976) described human factors in terms of antecedent human states that in turn could lead to a deterioration in performance. Although Fell’s “Accident causal system” did not take the form of chart representations, it was based on a cause-effect relationship, in which human factors could be organised in the analysis of traffic accidents. In this system, the human effects were expressed in terms of human information-processing failures or non-performances described as four distinct yet interrelated processes (Figure 4a). The human causes behind these effects were grouped into five categories. Although the author exemplified some likely human cause-effect relationships (e.g. the physiological failure of falling asleep would be a cause behind the non-perception effect), the relationships were not pre-defined. The intention was to identify the cause-effect relationships on the basis of in-depth accident investigations. While Fell focused on human factors in the proposed system, he acknowledged that other causes, e.g. vehicle, highway and ambient conditions, could lead to human effects. Figure 4b presents an example of a hypothetical crash with a human causal chain and equivalent chains for the vehicle and environment (Fell, 1976, p. 86-87).

Figure 4. a) Human causes and effects, b) A hypothetical crash according to the "Accident causal system", c = cause, e = effect (based on Fell 1976).

1.5.4 Compilation and aggregation of Accident Mechanisms

Malaterre (1990) presented one of the few studies that have attempted to compile as well as aggregate causation charts. The aggregation attempts were made during the evaluation of a classification scheme, partly inspired by Fell (1976). The classification scheme was divided into factors, antecedents, function failures, failure tasks, etc. and intended to have a high degree of compatibility with an accident model represented by four accident phases; driving, accident, emergency and collision. The classification scheme was used during the analysis of 72 in-depth accident investigations involving 115 road users. The compilation and aggregation focused on the accident phase represented by the three “accident mechanisms”: antecedents, errors and function failures, most often represented by one sequence for each road-user.
The intention behind the aggregation of accident mechanisms was to find common sequence patterns. The accident mechanisms of the 115 road-users were compared, and those with similar mechanisms were grouped together into 15 categories and subsequently aggregated. Malaterre (1990) found that the results, in terms of common sequence patterns, varied between the categories. Figure 5 presents the results from a successful aggregation, where the sequence of accident mechanisms was common to all 13 road-users in that category. Figure 6 presents a category with a less successful result, where antecedents could be identified for only two of the five road-users. Malaterre also tried to extend the scope of the aggregated charts to include the driving phase as well as the emergency phase. However, the author found that the extended and aggregated charts were too complicated, especially if they included the variety of emergency manoeuvres that occur in the emergency phase. Malaterre (1990) concluded that while the classification scheme had potential, the methods for compiling and aggregating charts required further development.

Figure 5. Aggregated accident mechanisms for category 5 (13 road users) (Labels for antecedents and function failures are added. From Malaterre 1990).

Figure 6. Aggregated accident mechanisms for category 9 (5 road users) (Labels for antecedents and function failures are added. From Malaterre 1990)

1.5.5 The Cognitive Reliability and Error Analysis Method

The Cognitive Reliability and Error Analysis Method (CREAM) was proposed by Hollnagel (1998) as an alternative to previous methods which were primarily based on fault and event tree representations used in accident analysis and predictive risk-analysis in the processing industry and nuclear power-plants. CREAM has three parts; method, classification scheme and model.

Unlike previous methods based on event or fault tree representations, the method (i.e. the classification procedure) associated with CREAM is “recursive” rather than strictly hierarchical. This is a consequence of the flexibly linked categories of contributory factors in the classification scheme. Because of the non-hierarchical structure of this scheme, the method contains clear stop rules comprising well-defined conditions that determine the point at which an analysis has come to an end; otherwise an analysis could stop prematurely or go on forever.
On the highest level, the classification scheme makes a distinction between observable effects (phenotypes) and the causes (genotypes) of those effects (see e.g. Hollnagel, 1998, p. 48). Observable effects refer to human (overt) actions as well as system events such as indicated malfunctions, releases of matter and energy, changes in speed and direction, etc (see Figure 7). In retrospective accident analysis, the effects are the starting point of the analysis. The causes (genotypes) are the categories that can be used to describe that which brought about the effect(s). In CREAM, the genotypes are divided into 14 main categories of factors, placed under three main classes according to the Man (human), Technology and Organisation (MTO) framework (Table 1).

Figure 7. The relation between causes and manifestation of effects (based on Hollnagel, 1998)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication</td>
<td>Equipment failure</td>
</tr>
<tr>
<td>Organisation</td>
<td>Procedures</td>
</tr>
<tr>
<td>Training</td>
<td>Temporary interface</td>
</tr>
<tr>
<td>Ambient conditions</td>
<td>Problems</td>
</tr>
<tr>
<td>Working conditions</td>
<td>Permanent interface</td>
</tr>
<tr>
<td></td>
<td>problems</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In addition to listing genotypes and phenotypes, the classification scheme also describes possible links between them. Because the classification scheme is not organised in a strictly hierarchical fashion, the links between categories follow a repeated antecedent-consequent (or cause-effect) relationship. The links show the different ways in which genotypes may affect each other, while the actual relationships between factors in an accident are based on available accident information.

The model refers to the Contextual Control Model (COCOM), a cyclical model of human cognition (Hollnagel, 1998). COCOM mainly serves as a basis for organising the categories of four specific human functions; observation, interpretation, planning and action. In the same way as the classification in general, a noteworthy aspect is the emphasis on the distinction between what can be observed (an action) and what must be inferred (observation, interpretation and planning). What follows is a non-sequential representation of cognition, which means that the path through the classification scheme is guided by the possible causal links between the various cognitive functions as these unfold in a particular context. These links cannot be defined à priori, but must reflect the prevailing conditions as they are known or assumed by the accident analysis (Hollnagel, 1998).
2 Objectives
The overall objective of this thesis is to evaluate whether case studies of motor-vehicle crashes can be aggregated with the help of causation charts representing causal relationships between contributory factors in the pre-crash accident phase.

The specific aims are to:

- Evaluate general accident models regarding their potential to form a basis for causation analysis of road traffic accidents
- Evaluate whether causation charts compiled with a specific method can be aggregated in order to identify common causation patterns.
- Assess whether an adequate level of intercoder agreement can be reached in the compilation of causation charts.
3 The Driving Reliability and Error Analysis Method

The Driving Reliability and Error Analysis Method (DREAM) is an adaptation of the Cognitive Reliability and Error Analysis Method (CREAM; Hollnagel, 1998). While CREAM was developed to analyse accidents within process control domains such as nuclear power plants and the processing industry, DREAM has been adapted to suit the road traffic domain.

The present chapter mainly describes the most recent version, DREAM 3.0. Three versions of the method (the first version of DREAM (Ljung, 2002), DREAM version 2.1 (Ljung et al., N.d.), and DREAM version 3.0 (Wallén Warner et al., 2008b)) were used in each of the three appended papers (II, III and IV). The reason for this is that the method has been continuously developed. A more detailed description of its development can be found in Chapter 5.

Addendum: Revision of DREAM. For the sake of comparison, two tables are provided at the end of the present chapter presenting the contributory factors of DREAM 2.1 and DREAM 3.0. The classification scheme of the first version (Ljung, 2002) is not provided, as the genotype categories are largely the same as in version 2.1.

The goal of DREAM is to enable systematic classification of accident causation information in multidisciplinary in-depth case studies. The result of a DREAM analysis is presented as a causation chart (DREAM-chart) of interlinked contributory factors. The advantage of systematically compiled DREAM-charts is that they facilitate direct comparison of accident causation, and are possible to aggregate in order to find common causation patterns. Furthermore, DREAM has been developed for the purpose of identifying traffic situations for which the development of technical solutions has the potential to reduce the number of future accidents. As can be seen in Figure 8, accident prevention systems can be roughly divided into four main types, where each type presents its own challenges in terms of accident investigation and the development of countermeasures.

Aim

<table>
<thead>
<tr>
<th>Mode</th>
<th>Collision avoidance</th>
<th>Risk avoidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autonomous systems</td>
<td>Technically possible but difficult from a legal perspective.</td>
<td>Technically possible, but efficiency is threatened by driver adaptation.</td>
</tr>
<tr>
<td>Interactive systems</td>
<td>Technically complicated, since the time needed for driver action puts extreme demands on sensor and algorithm performance in situation identification.</td>
<td>Technically possible and often easier than collision avoidance, but very demanding from an HMI perspective.</td>
</tr>
</tbody>
</table>

Figure 8. Various types of active safety systems targeting different areas of accident avoidance (from Wallén Warner et al., 2008b).

When DREAM was first developed, its main focus was on only one of the four prevention types. More specifically, the aim was to identify interactive systems for risk prevention (Figure 8: lower right quadrant). Consequently, the DREAM causation categories in DREAM, as well as the underlying accident model reflect this.
DREAM includes the same three parts as CREAM, which are presented below in the following order: accident and driver models, classification scheme and method.

### 3.1 The accident and driver models

In DREAM, the accident model and the driver model are used to organise the genotype and phenotype categories in the classification scheme. The accident model employs the human-technology-organisation (HTO) triad as a reference – in DREAM 3.0 represented by the driver (human), the vehicle and traffic environment (technology) and the organisation. Figure 9 illustrates how accidents are seen as the result of an unsuccessful interplay between driver, vehicle and traffic environment, as well as the organisation(s) responsible for shaping the conditions under which driving takes place. Failures at the sharp end (Reason, 1997) as well as at the blunt end (Hollnagel, 1998, 2004) are taken into consideration. Sharp end failures take place in close proximity to the accident (e.g. the driver fails to see a red traffic light which contributes to two cars colliding), while blunt end failures occur at other times and/or locations. For example, a mechanic fails to maintain the brakes properly, which later contributes to two cars colliding. The faulty brakes are a latent (failure) condition (Reason, 1997), hidden in the system.

![Figure 9. The accident model on which DREAM 3.0 is based (Wallén Warner et al., 2008b)](image)

The driver model is based on the Contextual Control Model (COCOM) (Hollnagel, 1998; Hollnagel and Woods, 2005), which is used to organise the four basic cognitive functions of observation, interpretation, planning and action related to the driver in the driver-vehicle/traffic environment-organisation triad. Figure 10 presents the approximate relative positions of these functions in the most recent version of COCOM (see Hollnagel and Woods, 2005). COCOM recognises that cognition includes processing observations and performing actions, as well as continuously revising goals and intentions which create a “loop” on the level of interpretation and planning.
3.2 The classification scheme

The DREAM classification scheme comprises a number of observable effects in the form of human actions and system events, represented by the phenotypes. It also contains a number of possible contributory factors, i.e. the genotypes, which may have brought about these observable effects. Besides the phenotypes and genotypes, the DREAM classification scheme also includes links between phenotypes and genotypes, as well as between different genotypes.

3.2.1 Phenotypes

The purpose of the phenotypes is to classify the observable effects into a relatively limited set of categories, which forms the starting point for the actual analysis. The general phenotypes are all linked to one or several specific phenotypes. The difference between general and specific phenotypes is the degree of information, where the latter describe more specific effects than the former. If the investigator has sufficient information about the accident, a specific phenotype should be chosen. Table 2 presents the phenotype categories in DREAM versions 2.1 and 3.0.

Table 2. General and specific phenotypes in DREAM versions 2.1 and 3.0.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>DREAM 2.1 Specific phenotypes</th>
<th>DREAM 3.0 Specific phenotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timing</td>
<td>Premature action, Late action, No action</td>
<td>Timing Too early action; Too late action; No action</td>
</tr>
<tr>
<td>Duration</td>
<td>Prolonged action, Shortened action</td>
<td>Speed Too high speed; Too low speed</td>
</tr>
<tr>
<td>Force</td>
<td>Insufficient force, Surplus force</td>
<td>Distance Too short distance</td>
</tr>
<tr>
<td>Distance</td>
<td>Prolonged distance, Shortened distance</td>
<td>Direction Wrong direction</td>
</tr>
<tr>
<td>Speed</td>
<td>Surplus speed, Insufficient speed</td>
<td>Force Surplus force; Insufficient force</td>
</tr>
<tr>
<td>Direction</td>
<td>Incorrect direction</td>
<td>Object Adjacent object</td>
</tr>
<tr>
<td>Object</td>
<td>Adjacent object, Similar object</td>
<td></td>
</tr>
<tr>
<td>Sequence</td>
<td>Skipped action, Repeated action, Reversed action, Extraneous action</td>
<td></td>
</tr>
<tr>
<td>Quantity/volume</td>
<td>Too little, Too much</td>
<td></td>
</tr>
</tbody>
</table>
Some of the phenotypes (e.g. timing, distance and speed) are very closely related even though they are conceptually distinct. If, for example, a car collides with an oncoming car when overtaking, should it be considered an effect of timing (the overtaking was initiated too early or too late), distance (the stretch of free road was too short in order to complete the overtaking) or speed (the speed was too low in order to complete the overtaking)? The answer is that the investigator has to choose the phenotype that makes the most sense, given what is known about the accident.

With regard to the example above, although all three phenotypes are logically possible, one of them is probably more appropriate in the given circumstances. Let us assume that the overtaking is performed at a speed of 100 km/h (speed limit 90 km/h) close to the crest on an uphill slope. *Speed: Too low speed* is then a less appropriate choice of phenotype as the speed was more than sufficient. *Distance: Too short distance* seems more appropriate, as the stretch of free road was too short to overtake safely. However, it is common driver knowledge (taught in driver training) that one should not overtake unless there is a clear view of a sufficient stretch of road and in this case the crest of the hill obviously blocked the view. In view of this fact, the most appropriate phenotype would be *Timing: Too early action*.

Sometimes the choice of phenotype can be quite tricky. In DREAM 3.0 (unlike DREAM 2.1), all phenotypes are linked to the same set of genotypes and therefore a less appropriate choice of phenotype will not affect the genotype choices.

### 3.2.2 Genotypes

Genotypes are factors which may have contributed to the phenotypes (the observable effects). The genotypes can usually not be observed and therefore they have to be inferred from e.g. interviews with the drivers or other information gathered in the investigation. In DREAM 3.0, there are 51 general genotypes, some of which are linked to one or several specific genotypes. As with the phenotypes, the difference between general and specific genotypes is the degree of detail in the information available, where the specific genotypes describe more specific factors than the general ones. A specific genotype should be chosen if the investigator has sufficient information about the accident. See Table 5 for examples of antecedent general and specific genotypes of the Interpretation category.

In DREAM 3.0, the genotypes are organised according to the driver-vehicle/traffic environment-organisation triad. The driver category consists of genotypes related to possible problems with cognitive functions such as observation, interpretation and planning (in accordance with COCOM). It also includes general temporary and permanent person related factors that can contribute to an accident (e.g. inattention). The vehicle/traffic environment category comprises of vehicle and traffic environment related genotypes, while the organisation category consists of genotypes related to organisation, maintenance and design. See Table 3 and 4 at the end of the present chapter for a schematic presentation of the different categories in DREAM version 2.1 and 3.0 respectively.

### 3.2.3 Links

Besides the phenotypes and genotypes mentioned above, the DREAM classification scheme also includes links between the phenotypes and genotypes, as well as between different genotypes. In DREAM 3.0 these links represent existing knowledge about how different factors can interact with each other (for a review see Wallén Warner et al., 2008b) and result in analysis chains where a genotype can be both the consequent of a previous genotype and
the antecedent of another genotype, e.g. the cause of the genotype. If, for example, genotype A results in genotype B and genotype B results in genotype C, then A can be said to be the indirect cause of C, and B can be said to be both a result of A and a cause of C (See Figure 11). The DREAM genotypes can therefore function both as forward and backward links in a chain of reasoning, which makes it possible to deduce indirect causes (such as A in relation to C in the present example).

![Figure 11. A is an indirect cause of C](image)

The links between the phenotypes and the genotypes, as well as between different genotypes, are incorporated in the classification scheme. Table 5 presents an excerpt from DREAM 3.0 with the possible backward links to antecedent general and specific genotypes of the Interpretation category.

### 3.3 The method

The method (i.e. the classification procedure) contains several stop rules, which are well defined conditions that determine when the analysis should be terminated. Stop rules are necessary, as the classification scheme represents a network (rather than a hierarchy) and the analysis could go on forever in the absence of such rules.

#### 3.3.1 Stop rules

The DREAM classification scheme is non-hierarchical, which means that no genotypes have precedence, and there are no highest or lowest levels at which an analysis must end. Stop rules are therefore necessary to avoid random or subjectively determined termination of the analysis.

Overall, general genotypes have the status of non-terminal events. If a general genotype is the most likely cause of a general consequent, that cause is chosen and the analysis must continue until one of the three stop rules below is fulfilled.

The stop rules in DREAM 3.0 are:

1. *Specific genotypes have the status of terminal events. Therefore, if a specific genotype is the most likely cause of a general consequent, that genotype is chosen and the analysis stops.*
2. *If no general or specific genotypes that link to the chosen consequent exist, the analysis stops.*
3. *If none of the available specific or general genotypes for the chosen consequent is relevant, given the available accident information, the analysis stops.*
<table>
<thead>
<tr>
<th>MAN Driver</th>
<th>TECHNOLOGY Vehicle</th>
<th>ORGANISATION Traffic environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation</td>
<td>Temporary HMI problems</td>
<td>Communication</td>
</tr>
<tr>
<td>Missed observation</td>
<td>Access limitations</td>
<td>Communication failure (between drivers)</td>
</tr>
<tr>
<td>False observation</td>
<td>Incorrect information</td>
<td>Information failure (between driver and traffic environment, or driver and vehicle)</td>
</tr>
<tr>
<td>Wrong identification</td>
<td>Temporary sight obstruction</td>
<td></td>
</tr>
<tr>
<td>Interpretation</td>
<td>Permanent HMI problems</td>
<td>Maintenance</td>
</tr>
<tr>
<td>Faulty diagnosis</td>
<td>Sound</td>
<td>Maintenance failure</td>
</tr>
<tr>
<td>Wrong reasoning</td>
<td>Illumination</td>
<td>Inadequate quality control</td>
</tr>
<tr>
<td>Decision error</td>
<td>Access problems</td>
<td></td>
</tr>
<tr>
<td>Delayed interpretation</td>
<td>Mislabelling</td>
<td></td>
</tr>
<tr>
<td>Incorrect prediction</td>
<td>Permanent sight obstruction</td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td>Equipment failure</td>
<td>Experience/Knowledge</td>
</tr>
<tr>
<td>Inadequate plan</td>
<td>Equipment failure</td>
<td>Insufficient skills</td>
</tr>
<tr>
<td>Priority error</td>
<td>Software fault</td>
<td>Insufficient knowledge</td>
</tr>
<tr>
<td>Temporary Personal Factors</td>
<td></td>
<td>Organisation</td>
</tr>
<tr>
<td>Memory failure</td>
<td>Inad. HMI</td>
<td>Inad. instructions/procedures</td>
</tr>
<tr>
<td>Fear</td>
<td>Cognitive bias</td>
<td>Overload/Too high demands</td>
</tr>
<tr>
<td>Distraction</td>
<td>Inad. ergonomics</td>
<td>Inad. management</td>
</tr>
<tr>
<td>Fatigue</td>
<td>Inad. design of communication devices</td>
<td>Inad. training</td>
</tr>
<tr>
<td>Performance variability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inattention</td>
<td></td>
<td>Road design</td>
</tr>
<tr>
<td>Under the influence of substances</td>
<td></td>
<td>Inadequate road design</td>
</tr>
<tr>
<td>Physiological stress</td>
<td></td>
<td>Obstruction to view</td>
</tr>
<tr>
<td>Psychological stress</td>
<td></td>
<td>Inad. information design</td>
</tr>
<tr>
<td>Permanent Personal Factors</td>
<td></td>
<td>Vehicle design</td>
</tr>
<tr>
<td>Functional impairment</td>
<td>Unpredictable system characteristics</td>
<td>Inad. HMI</td>
</tr>
<tr>
<td>Cognitive bias</td>
<td></td>
<td>Inad. ergonomics</td>
</tr>
</tbody>
</table>

inad. = inadequate
Table 4. Genotypes of DREAM version 3.0 (Wallén Warner et al., 2008b).

<table>
<thead>
<tr>
<th>HUMAN Driver</th>
<th>TECHNOLOGY Vehicle</th>
<th>ORGANISATION Organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Observation</strong></td>
<td><strong>Temporary HMI problems</strong></td>
<td><strong>Organisation</strong></td>
</tr>
<tr>
<td>Missed observation</td>
<td>Temporary illumination problems</td>
<td>Time pressure</td>
</tr>
<tr>
<td>Late observation</td>
<td>Temporary sound problems</td>
<td>Irregular working hours</td>
</tr>
<tr>
<td>False observation</td>
<td>Temporary sight obstructions</td>
<td>Heavy physical activity before drive</td>
</tr>
<tr>
<td><strong>Interpretation</strong></td>
<td><strong>Temporary access limitations</strong></td>
<td><strong>Inad. training</strong></td>
</tr>
<tr>
<td>Misjudgement of time gaps</td>
<td>Incorrect ITS-information</td>
<td></td>
</tr>
<tr>
<td>Misjudgement of situation</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Planning</strong></td>
<td><strong>Permanent HMI problems</strong></td>
<td><strong>Maintenance</strong></td>
</tr>
<tr>
<td>Priority error</td>
<td>Permanent illumination problems</td>
<td>Inad. vehicle maintenance</td>
</tr>
<tr>
<td><strong>Temporary Personal Factors</strong></td>
<td><strong>Permanent sound problems</strong></td>
<td>Inad. road maintenance</td>
</tr>
<tr>
<td>Fear</td>
<td>Permanent sight obstruction</td>
<td></td>
</tr>
<tr>
<td>Inattention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatigue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under the influence of substances</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excitement seeking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sudden functional impairment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological stress</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Permanent Personal Factors</strong></td>
<td><strong>Vehicle equipment failure</strong></td>
<td><strong>Vehicle design</strong></td>
</tr>
<tr>
<td>Permanent functional impairment</td>
<td>Equipment failure</td>
<td>Inad. design of driver environment</td>
</tr>
<tr>
<td>Expectance of certain behaviours</td>
<td></td>
<td>Inad. design of communication devices</td>
</tr>
<tr>
<td>Expectance of stable road environment</td>
<td></td>
<td>Inad. construction of vehicle parts and/or structures</td>
</tr>
<tr>
<td>Habitually stretching rules and recommendations</td>
<td></td>
<td>Unpredictable system characteristics</td>
</tr>
<tr>
<td>Overestimation of skills</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insufficient skills/knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>State of road</strong></td>
<td><strong>Traffic environment</strong></td>
<td><strong>Road design</strong></td>
</tr>
<tr>
<td>Insufficient guidance</td>
<td><strong>Weather conditions</strong></td>
<td>Inad. information design</td>
</tr>
<tr>
<td>Reduced traction</td>
<td>Reduced visibility</td>
<td>Inad. road design</td>
</tr>
<tr>
<td>Road surface degradation</td>
<td>Strong side winds</td>
<td></td>
</tr>
<tr>
<td>Object on road</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inadequate road geometry</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Communication</strong></td>
<td><strong>Obstruction of view due to object</strong></td>
<td></td>
</tr>
<tr>
<td>Inad. transmission from driver</td>
<td>Temporary obstruction of view</td>
<td></td>
</tr>
<tr>
<td>Inad. transmission from road environment</td>
<td>Permanent obstruction of view</td>
<td></td>
</tr>
</tbody>
</table>

inad. = inadequate
Table 5. Excerpt from the classification scheme of DREAM 3.0 showing the genotype category of interpretation and the possible links backwards to the antecedent specific or general genotypes.

**INTERPRETATION C**

Interpretation includes, for all but novice drivers, quick and automated (routine) procedures where typical situations and their associated actions are recognized and acted upon (script choice). Mistakes in interpretation occur at the sharp end – within the local event horizon.

<table>
<thead>
<tr>
<th>ANTECEDENTS</th>
<th>SPECIFIC Genotypes (with definitions)</th>
<th>Examples for SPECIFIC Genotypes</th>
<th>CONSEQUENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Late observation (B2)</td>
<td>Misjudgement of time gap due to incorrect speed estimate (C1.1)</td>
<td>Intersection</td>
<td>None defined</td>
</tr>
<tr>
<td>False observation (B3)</td>
<td>The driver misjudges the time gap due to a misjudgement of the approaching vehicle’s speed.</td>
<td>Misjudgement of time gaps (C1)</td>
<td>The estimation of time gaps (e.g. time left to approaching vehicle, stop sign, traffic lights etc.) is incorrect.</td>
</tr>
<tr>
<td>Fatigue (E3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under the influence of substances (E4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological stress (E7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permanent functional impairment (F1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expectancy of certain behaviors (F2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habitually stretching rules and recommendations (F4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overestimation of skills (F5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insufficient skills/knowledge (F6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incorrect ITS-information (G5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced visibility (J1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insufficient guidance (L1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced friction (L2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inadequate road geometry (L5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inadequate transmission from road environment (M2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unpredictable system characteristics (P4)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

19
4 Summary of Papers I, II, III, and IV

4.1 Summary of Paper I

The purpose of Paper I is to evaluate general accident models in terms of their potential to form a basis for causation analysis of road traffic accidents.

4.1.1 Methodology

Four general accident models were evaluated on the basis of six traffic accident model criteria in the analysis of causation and subsequent identification of preventive measures.

The four general accident models

The four general accident models comprised Sequential models single-event (Heinrich et al., 1980) and fault tree models, Epidemiological models such as the Host-Agent-Environment model (Gooden, 1949) and Haddon’s matrix (Haddon, 1972), Energy transfer models that form the basis of Haddon’s ten countermeasure strategies (Haddon, 1975) and Systemic accident models in which accidents occur when several causal factors, i.e. human, technical and environmental, exist concurrently within a specific time and space (Hollnagel, 1998; Leveson, 1995).

The six evaluation criteria

1. A clear definition of concepts and interactions.
2. A clear definition of scope.
3. Handling of extended time spans.
4. Handling of dynamic aspects.
5. Adequate analysis methods and stop rules.
6. Suitability for preventive work.

4.1.2 Results and discussion

Theoretically, sequential models retrospectively explore a wide range of elements in the search for causal factors. However, analysis methods based on these models tend to focus upon events involving individual road users immediately prior to the accident. Consequently, such models have limited scope, static characteristics and a one-to-one perspective on causes. Epidemiological models overcome the limitations of sequential models that focus on events by taking account of latent conditions, thereby widening the scope of actors and time. However, they are entirely descriptive with little predictive potential and are incapable of describing dynamic and interactive processes. Energy transfer models cannot function as accident prevention models in modern traffic conditions, as they do not describe how and why accidents occur. They also tend to cover very short time spans, i.e. initiation, impact and standstill. With their wide and flexible scope, Systemic models are capable of describing the complex and dynamic nature of driving. However, at present, such models cannot be used in accident prevention work, since neither analysis methods nor stop rules are defined at a sufficient level of detail for the development of countermeasures. Consequently, none of these models fulfill all six criteria.

Furthermore, the four general accident models share a common structural problem, in which the human is seen as a system component, and this perspective is transferred to the analysis methods. As the analysis involves the division of a system into its parts in order to identify the failing one(s), the interaction between contributory factors is neglected. This seriously hinders
the possibility of describing and defining the dynamic and interactive characteristics of road traffic.

4.1.3 Conclusions
The four general Sequential, Epidemiological, Energy transfer and Systemic accident models evaluated in this study are inadequate for modelling accidents in a modern traffic system. In addition, they share a common structural problem in which the human and other elements are treated as separate components. There is a need to develop a traffic accident model that can meet the criteria set out in this study, where the goal is to form the basis for an analysis method that can be used for the identification of measures. The method needs to take account of the variety of interactions that can occur between traffic elements as well as incorporate stop rules to ensure that the analysis remains structured.

4.2 Summary of Papers II and III
In Paper I it was concluded that the goal of a traffic accident model is to form the basis for an analysis method that can be used for the identification of measures. Such a method needs to take account of the variety of interactions that can occur between traffic elements. Paper II and III evaluated the use of such a method, namely the Driving Reliability and Error Analysis Method (DREAM), in which contributory factors are systematically analysed, classified and linked in a causation chart. A causation chart is thus able to demonstrate the interaction between several factors in the course of an accident.

The purpose of Papers II and III is to evaluate whether causation charts, compiled using DREAM and DREAM 2.1, respectively, can be aggregated in order to identify common causation patterns.

4.2.1 Methodology
For the purpose of Papers II and III, 100 case files of motor-vehicle crashes were examined in order to find single-vehicle crashes and intersection crashes. These 100 crashes occurred during 2003 and 2004 in the Gothenburg area of Sweden. In-depth, on-scene crash investigations were conducted by a multidisciplinary team, independent of the police within a limited geographic area, on weekdays during working hours.

The case files contained information about the time and day of the week, month, weather conditions, visibility, driver, vehicle and road environment, calculated or estimated approach speeds, etc. The circumstances of the accident were described in the form of a narrative, including the drivers’ goals and intentions during the journey and if and how they had reacted in the emergency phase. One DREAM causation chart was compiled for each driver.

For the purpose of Paper II, 38 of the 100 accident cases were regarded as single-vehicle crashes, 28 were ordinary single-vehicle crashes, while the ten other cases resulted in head-on collisions when the vehicle in question entered the opposite traffic lane. Despite the involvement of a second vehicle, the ten cases were treated as single-vehicle crashes due to the fact that the focus of Paper II was crash causation. For the purpose of Paper III, 26 of the 100 crashes were found to have occurred at urban crossing-path intersections.

The DREAM charts related to the single-vehicle crashes (Paper II) were compiled with the first version of DREAM (Ljung, 2002). Those related to the intersection crashes (Paper III) were also compiled using the first version of DREAM, but updated by means of DREAM
version 2.1 (Ljung et al., N.d.) before being aggregated. The update was carried out by three accident investigators who were unaware of the purpose of the study.

In Paper II, the 38 DREAM charts were aggregated on the basis of four different types of loss of vehicle control, termed “scenarios” (Table 6). The 52 DREAM charts in Paper III (one chart for each of the drivers in the 26 intersection crashes) were aggregated for six defined intersection crash risk situations (Table 7).

<table>
<thead>
<tr>
<th>Scenario number</th>
<th>Type of loss of vehicle control</th>
<th>Number of crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Vehicle drifts out of lane</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>Loss of control in curves with locally reduced road friction</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>Excessive speed in curves</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>Alarmed drivers react with excessive driver manoeuvres</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 7. Definitions of risk situations and distribution of the 52 drivers from the 26 intersection crashes in Paper III

<table>
<thead>
<tr>
<th>Driver without the right of way</th>
<th>Number of drivers</th>
<th>Driver with the right of way</th>
<th>Number of drivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>The driver failed to observe a red or amber light or a sign.</td>
<td>6</td>
<td>V</td>
</tr>
<tr>
<td>II</td>
<td>The driver observed a red or amber light or a sign but continued driving.</td>
<td>4</td>
<td>VI</td>
</tr>
<tr>
<td>III</td>
<td>The driver failed to observe the vehicle with the right of way.</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>The driver observed the vehicle with the right of way but continued driving.</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

4.2.2 Results and Discussion

The results of the aggregations revealed common causation patterns, although they were less clear for the intersection crashes. See Figures 2 a-d in Paper II (p. 321-322) and Figures 1 - 6 in Paper III.

In Paper II, which focused on the 38 single-vehicle crashes, the results indicated common patterns within the categories, as well as different patterns between them. In the first scenario, vehicles drifted out of lane due to driver fatigue or distraction. In the second, an undetectable reduction in road friction caused experienced drivers to lose control of the vehicle in curves. Loss of vehicle control in curves was also a factor in scenario three, partly due to high speed, but also because the drivers had overestimated their driving skills and/or had limited experience of the vehicle or the curve. In the fourth scenario, frightened drivers lost control of the vehicle as a result of excessive steering-wheel manoeuvres.

In Paper III, which dealt with the 26 intersection crashes, clear patterns were found in three of the six risk situations, i.e. III, V and VI. A common pattern in risk situations III and V revealed that drivers with and without the right of way had not seen the other vehicle due to distractions and/or sight obstructions. A frequently occurring pattern for the drivers with the
right of way (risk situations V and VI) was that they did not expect another vehicle to cross
their path. The absence of clear common patterns in risk situations I, II and IV was due to the
low number of charts and the relatively unique crash circumstances.

In comparison with the aggregated charts in Paper II, those in Paper III generally exhibited
greater variation. In the latter, three investigators had prepared the charts using DREAM 2.1.
When the investigators were questioned about their assessment of factors, two main problems
related to inconsistency were revealed; firstly, the assignment of genotypes belonging to the
“Interpretation” category (Table 3) and secondly the assignment of phenotypes in risk
situations I and II. There were no obvious inconsistencies in Paper II, most likely because
only a single investigator had prepared the charts for the single-vehicle crashes.

The first problem related to inconsistency in Paper III concerned the assigned factors
belonging to the Interpretation category, which was due to lack of clear distinctions between
them. This was particularly problematic in cases with lack of, or ambiguous, accident
information. In addition, the team could not rule the possibility that their assessment may
have been influenced by the factors assigned in the original chart, despite the fact that they
were allowed to change them. Taken together, this led to a large variation in the assigned
Interpretation factors (indicated by the letter C), which can be seen in the aggregated charts.

The second problem of inconsistency in Paper III concerned the phenotypes that had been
assigned in different ways in the crashes at traffic-light controlled intersections i.e. risk
situations I and II. The problem was that the phenotype was placed at different positions in the
course of the accident, which affected the appearance of the causation chart. Figure 12
demonstrates the problem in the case of a fictitious crash, in which the causation chart for
Vehicle A differs depending on the position (A1 or A2) the phenotype is placed. Out of eight
charts in Paper III, five were assigned a phenotype related to the A1 position and three an A2
position (Paper III: p. 8 and Figures 1 and 2). As demonstrated in Figure 12, the assigned
phenotype sets the point at which the DREAM analysis should start. When starting the
analysis at A1, the factors Distraction leading to Missed observation of the red traffic light are
included. When starting the analysis at A2, the same factors are included, in addition to Sight
obstruction, which led to Missed observation of vehicle B. Consequently, the phenotype
position influences which factors that are included from the phenotype and backwards.
Paper III contained several common patterns demonstrating the influence of distraction. However, the type of distraction could not be identified without closer examination of the text excerpts. Although this had no consequence in terms of consistency, the appearance and interpretation of the patterns would benefit from a more detailed classification of the type of distraction.

In Paper II, and to some extent in Paper III, the results of each aggregation were compared with findings from previous accident studies, observational studies, etc. It emerged that several of the contributory factors could be identified in other studies, either as a causation factor in crashes or as a risk factor in the observational studies. However, while the previous studies focused on one or a few factors at a time, the DREAM charts were able to illustrate how several interacting factors were needed to cause a crash.

### 4.2.3 Conclusions
Papers II and III demonstrate that the rather unique approach of aggregating causation charts is a methodology that can explain how interacting factors has caused crashes. The results indicate that the approach of aggregating DREAM charts has potential, provided that the charts are compiled in a reliable and consistent way.

### 4.3 Summary of Paper IV
The results of Papers II and III indicated that the approach of aggregating DREAM charts has potential. However, a prerequisite for identifying common patterns in aggregated charts is that, when several investigators are involved, the charts are compiled in a consistent manner. The inconsistencies identified in Paper III call for improved formal operational definitions of the phenotypes and some of the genotypes. The experience gained in Paper III led to DREAM being upgraded into a new version known as DREAM 3.0 (Wallén Warner et al., 2008b). See Chapter 5. Addendum: Revision of DREAM for further details.
The purpose of Paper IV is to allow coders (i.e. accident investigators) from several European countries to analyse and classify the causes of the same accident scenarios in order to assess the intercoder agreement of DREAM 3.0.

4.3.1 Methodology
A total of 11 coders were contacted, and nine agreed to participate in the study, of whom seven were deemed to have sufficient experience of DREAM analyses to take part in the study. The seven coders ranged in age from 27 to 50 years, with a mean of 34 years. Six were male and one was female. The investigators were located in Sweden, the Netherlands, Great Britain and Italy.

The coders were trained independently at a distance using written material (by e-mail), consisting of test instructions, a DREAM 3.0 manual and three training cases in the form of accident scenarios to be analysed by compiling DREAM-charts. When the DREAM-charts for the training cases had been completed and returned to the authors of Paper IV, the coders were sent solutions to the cases together with comments on their proposed solution. They also received a new updated version of the DREAM 3.0 manual (with some minor adjustments) and the real test cases in the form of four accident scenarios for the coders to solve by compiling DREAM-charts. These accident scenarios included one catching-up, one single-vehicle, one intersection and one traffic light-controlled intersection accident. Each coder compiled seven DREAM-charts, one for each driver involved in the four accident scenarios.

4.3.2 Results and Discussion
The results showed that the intercoder agreement for genotypes ranged from 74% to 94% with an average of 83%, while for phenotypes it ranged from 57% to 100% with a mean of 78%. In comparison with other studies, which recommend a minimum of 85% agreement, the intercoder agreement in Paper IV can be considered somewhat low. However, when comparing agreement from different studies, it is important to take into account the amount of coder training and the point in the development/research process at which intercoder agreement is tested. The coders in Paper IV came from different organizations in various countries and only some of them had previously worked together. This, together with the fact that the coders only received distance training (via e-mail) and had no contact with each other, means that the intercoder agreement of the present study must be considered acceptable.

4.3.3 Conclusions
The results reveal that a high level of agreement can be reached, provided adequate training is provided. In addition, the testing of intercoder agreement can play an important role in identifying weaknesses in the classification scheme, the training of coders and the presentation of accident information.
5 Addendum: Revision of DREAM

DREAM 3.0 is the result of a method update of DREAM that took place as a consequence of the improvements in the operational definitions of phenotypes and genotypes identified in Paper III. The experiences gained as a result of Paper III led to a thorough examination of the DREAM versions, as well as CREAM (Hollnagel, 1998). Other input came from the practical experience of using the DREAM concept in both a national and a European project.

The first two versions of DREAM (Ljung 2002) and DREAM 2.1 (Ljung et al., N.d.) were originally used in the Swedish national project Factors Influencing the Causation of Accidents and incidents. When DREAM was to be employed in the European cooperation SafetyNet road safety project, DREAM 2.1 was translated into English and adapted to suit the traffic environment in the participating countries. This adapted version was called SafetyNet Accident Causation System (SNACS 1.1; Ljung (2006)) and has the same method, accident model and main classification scheme structure as DREAM 2.1, although some of the individual genotypes have been altered.

DREAM 2.1 and SNACS 1.1 were upgraded into the latest DREAM 3.0 version by a reference group headed by the authors of the DREAM 3.0 manual (i.e. Wallén Warner et al., 2008b). In DREAM 3.0, the main changes concern the phenotypes, genotypes and links classification schemes, which are described in the following sections.

5.1 Revision of the phenotypes

In Paper III it was found that the positioning of phenotypes was not consistent for crashes at traffic-light controlled intersections. This inconsistency led to different charts for similar crashes, which made it difficult to identify common causation patterns in the aggregated DREAM charts. On closer examination of the phenotypes in DREAM 2.1, it was found that their operational definitions were not explicit enough to consistently place them at the point at which the course of an accident turns from a risk situation to an emergency situation. A clear definition of what constitutes a risk situation was also lacking.

In DREAM 3.0, the operational definitions of phenotypes are more clearly related to the transition between the risk phase and the collision/emergency phase Figure 8. In order to facilitate the correct assignment of the phenotypes, their operational definitions are complemented by examples and illustrations that show where the phenotype is to be placed for different accident/traffic situations. The number of phenotypes was also reduced, as some of them were not used very much (See Table 2).

5.2 Revision of the genotypes

In Paper III, an inconsistency was found in the assigned genotypes belonging to the Interpretation category. In DREAM 2.1 this category comprises as many as five genotypes; Faulty diagnosis, Wrong Reasoning, Decision error, Delayed interpretation and Incorrect prediction (Table 3), which are the same as in CREAM. While the latter was primarily developed for processing industry, the operational definitions of the five interpretation genotypes were adjusted to fit the traffic domain in DREAM 2.1. However, in Paper III, it was found that the large variations were due to an inadequate distinction between the definitions of the interpretation genotypes.

The Contextual Control Model (COCOM) is a central model in both DREAM and CREAM. In both methods, COCOM is used to organise the four basic cognition categories;
Observation, Interpretation, Planning and Action, the latter being represented by the phenotypes.

According to Hollnagel (1998), the main categories in the CREAM classification scheme (Table 1) is generic, i.e., it is not intended as a classification scheme for a particular domain. At the same time however, Hollnagel (1998) acknowledges that the influence of nuclear power plants can probably be found in several parts of the classification scheme.

An example of the generic aspect is no doubt the four cognition categories, which, according to Hollnagel (1998), can be used for almost all applications. The influence of nuclear power plants is, however, to be found in the five genotypes under the Interpretation category, which are implicitly designed for a problem-solving operator of a complex industrial process, e.g. a nuclear power-plant. The situations faced by such an operator are likely to be far more complex than those faced by an average driver.

In DREAM 3.0, the four cognitive functions were kept the same as in the COCOM model, while the number of genotypes under the Interpretation category was reduced from five to two; Misjudgement of time gaps and Misjudgement of situation (Table 4). The two interpretation genotypes are thus more appropriate for the driving task and the occurrence of a risk situation and imply that, had the driver been aware of the circumstances of the situation, he/she might have acted differently.

While COCOM is used to organise the classification of human cognition in DREAM/SNACS and CREAM, the Man-Technology-Organisation (MTO) triad serves as a frame of reference for the main genotype categories. The MTO triad in DREAM 2.1/SNACS corresponds to Driver - Vehicle-Traffic environment/Organisation (Table 3) and is complemented by the notion of sharp-end, latent failure conditions and blunt end failures. When examining the DREAM 2.1 classification scheme however, these two theoretical frameworks do not completely overlap. For example, it could be assumed that Organisation in the MTO triad contains mainly blunt end failures, although the Communication category is clearly a sharp end failure. Obstruction to view under the Road design category is also a sharp end failure, or possibly a latent failure condition. Experience/knowledge is best described as a latent failure condition, which may have been the results of Inadequate training under the Organisation category, i.e. a blunt end failure. In general, in the same way as the categories under Vehicle, some categories under Traffic environment are better categorised in terms of Technology rather than Organisation.

In DREAM 3.0, the Man-Technology-Organisation (MTO) triad was retained as a frame of reference for the main genotypes categories (but with the term Human instead of Man, i.e. HTO). Traffic environment was moved from the Organisation to the Technology and therefore the HTO triad in DREAM 3.0 corresponds to the Driver – Vehicle/Traffic environment –Organisation (Table 4). The main genotype categories have been rearranged, so that they better correspond to the notion of sharp-end, latent failure conditions and blunt end failures. The operational definitions of the genotypes have been improved and supplemented by more examples.

In Paper III, the differences in types of distraction could not be identified without closer examination of the text excerpts (see Figures 1-6, Paper III). In DREAM 2.1, Distraction comprises two sub-genotypes; Internal distraction and External distraction. Although this had no consequence in terms of consistency, the appearance and interpretation of the patterns would benefit from a more detailed classification of the type of distraction.

Therefore, DREAM 3.0 comprises a more detailed classification of types of distraction, from two in DREAM 2.1, to five in DREAM 3.0; Driving-related distracters inside/outside vehicle, Non driving-related distracters inside/outside vehicle and Internal distracters. (The
types of distraction cannot be seen in Tables 3 and 4 because they are sub-genotypes of Distraction and Inattention respectively).

5.3 Revision of the links

The links were subject to many discussions during the practical work using DREAM 2.1 and SNACS, which mainly concerned two issues. Firstly, the predefined links between the phenotypes and genotypes, and secondly, the background to the links.

Regarding the first issue, each phenotype in DREAM 2.1 and SNACS is linked to its own predefined set of genotypes. Consequently, the choice of phenotype decides the scope of genotypes that can be selected in the next steps. In practice, this has occasionally resulted in investigators being required to choose a phenotype that may be less appropriate for the observable accident event, but which must be selected in order for the chart to end up in a required genotype at a later stage. Theoretically, however, the choice of phenotype should not be influenced by the required genotypes.

In DREAM 3.0, the links are arranged so that all phenotypes are linked to the same set of genotypes. In this way, the phenotype that best fits with the observable accident events can be chosen, yet does not influence the subsequent choices of genotypes. However, when compiling a chart using DREAM 3.0, in most cases it is necessary to choose one of the Interpretation genotypes; Misjudgement of time gaps or Misjudgement of situation. As mentioned above, the two interpretation genotypes imply that, had the driver been aware of the circumstances, he/she might have acted differently. By means of antecedent genotypes linked through one of the two interpretation genotypes, a completed chart indicates how to support the driver in order to avoid similar crashes, for example using interactive risk-avoidance systems.

With regard to the second issue pertaining to the background to the links, there are no documented references to the literature in DREAM, DREAM 2.1, SNACS or CREAM, on which a majority of the links in the former methods are based. Therefore, in connection with the update resulting in DREAM 3.0, a literature review was conducted in order to investigate the empirical support for the links between the genotypes in the classification scheme. The literature review resulted in the Documentation of references supporting the links in the classification scheme report (for further details see Wallén Warner et al., 2008a).
6 General discussion

Road traffic injuries due to motor-vehicle crashes are a huge public health problem all over the world and there is a need to find new measures to minimise the number of casualties, preferably by reducing the incidence of crashes. It is expected that the vehicle-mounted active safety measures now being introduced will accomplish this objective. The development of these measures has so far been driven by engineers and technological advances. The focus has been on the emergency phase in close proximity to the crash, a phase that is regarded as being beyond driver control and recovery. The development of higher level systems that address the circumstances occurring before this phase requires additional understanding and knowledge of driver behaviour. An obvious way to acquire this knowledge is to make use of multidisciplinary in-depth case studies. Traditionally, the qualitative information in case study files is analysed using a standard statistical approach. An alternative approach would be to convert the case files into causation charts. By aggregating causation charts, common causation patterns might be identified.

6.1 Two purposes behind aggregating case files by means of causation charts

Case files from multidisciplinary in-depth accident investigations have a high level of detail and can thus provide information that is unattainable in other ways, which is the main advantage of such studies (Fleury et al., 1994; OECD, 1988).

Due to the high level of detail, Fleury et al. (1994) concluded that multidisciplinary in-depth accident investigations, and thus case files, can be used in two ways:

1) To formulate hypotheses of general causes and causal relationships, which can subsequently be tested in experimental trials or analyses of representative accident material.
2) To supply detailed information about causes and causal relationships, of which some has been established in previous studies, for example in experimental trials or analyses of representative accident material.

With regard to the first way of using case files, Midtland et al. (1995) recognised that formulating hypotheses on the basis of one case is inadequate and, while two cases indicating the same causes would be somewhat better, several aggregated cases showing common causes have an even greater value. However, due to the large amount of information, case files as such are not easy to aggregate (Fleury and Brenac, 2001; Midtland et al., 1995), making the identification of common causes difficult.

One way of overcoming the difficulties of aggregating case files and obtaining a better verview of their similarities is to convert them into causation charts for subsequent aggregation. With reference to the first way of using case files, an examination of case files can give rise to hypotheses of common causes or causal relationships. However, because of their rich level of detail, it can be difficult to clearly recognise these causes and their relationships. Consequently, the causation charts for the case files in question can be aggregated, and if common causation patterns emerge, they give greater weight to the hypotheses.

Figure 13 illustrates how case files are converted into causation charts for subsequent aggregation, where one purpose is to facilitate this first way of using case files by elucidating
common causes and causal relationships that are indicated within them. Another purpose of aggregation can be to facilitate the second way of using case files. Aggregated causation charts can supply additional information about causes and causal relationships, of which some has been established in previous studies.

Figure 13. Two purposes of using aggregated causation charts based on case files from in-depth accident investigations.

From this presentation, it is clear that the aggregation of case files can be facilitated by means of aggregated causation charts. The idea of aggregating causation charts to find common causation patterns was suggested by Leplat and Rasmussen (1987). However, the authors did not present a complete methodology. In order to aggregate causation charts, the charts must be compiled by means of a systematic method. Three parts of such a method have been suggested by Hollnagel (1998) in the analysis of accidents in the processing industry and nuclear power plants.

### 6.2 Relations between models, classification scheme and classification method

Hollnagel (1998) argued that systematic analysis methods require three parts that are intrinsically related:

1. model(s) of a domain or phenomenon that describes the structure of the classification scheme
2. a classification scheme with categories of contributory factors
3. a classification method that describes the classification procedure

The definition of these parts determines the function and characteristics of the overall analysis method, which ultimately decides the appearance of the results and conclusions (Figure 14). This can be seen when examining the traditional analysis methods, the chain-of-events methods, CREAM and DREAM.
In traditional analysis methods, the qualitative information in case studies is coded using classification schemes of individually listed factors, after which the coded information is analysed by means of standard statistical techniques. Several studies that have adopted the traditional analysis method have acknowledged that an accident is the result of a breakdown of the traffic system containing the interacting elements of road-user, vehicle and traffic environment (Baker and Ross, 1960; Sabey and Staughton, 1975; Staughton and Storie, 1977; Treat et al., 1977a). This perspective of accidents clearly resembles a general system accident model (Paper I). However, more specific models are required in order to build appropriate classification schemes. In the aforementioned studies, such specific models may be either absent (e.g. Sabey and Staughton, 1975) or more detailed (e.g. Baker and Horn, 1960; Treat et al., 1977a). As the classification method comprises the assessment and coding of contributory factors by means of lists for each accident case, the interactions and causal relationships between the factors are analysed using correlation when all cases have been collected. However, the coding of individual factors leads to a fragmentation of the causal and interactive relationships that could be revealed by the original qualitative information. This retrospective analysis of interactions is the reason why the three traffic elements are treated as separate components as argued in Paper I.

The common feature of chain-of-events methods is that they explain accidents in terms of multiple events sequenced as a chain over time, where the events are chained together by means of connection rules. Chain-of-events methods are generally based on sequential models (Paper I). However, the majority lack specific models with which to build classification schemes. Consequently, the classification method comprises the rules for how to connect the events in a sequential and hierarchical fashion. If, however, a classification scheme should be combined with such strict connection rules, it would become strictly sequential and hierarchical. Thus, every category in the scheme would always follow or be followed by the same categories, whereby the analysis would also be strictly sequential and hierarchical (Hollnagel, 1998; Paper I).

CREAM and DREAM have an accident perspective in accordance with a general system accident model. The classification schemes are, however, based on more specific models and frameworks, i.e. the Contextual Control Model (COCOM) and Man-Technology-Organisation

<table>
<thead>
<tr>
<th>Data: observations, accident reports</th>
<th>The classification method describes the classification procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis</td>
<td>Classification scheme</td>
</tr>
<tr>
<td>Results and conclusions</td>
<td>Model(s)</td>
</tr>
</tbody>
</table>

The model(s) describe the structure of the classification scheme

![Figure 14. The relations between models, classification scheme and classification method (from Hollnagel 1998)](image)
(MTO) triad that are common to both methods, in addition to the accident model that is only relevant for DREAM (Figure 9). In CREAM and DREAM, the classification method comprises a classification of causes that is guided by the links between the categories in the classification scheme. While the links indicate the different ways in which categories of causes may affect each other, the actual relationships between them are based on available accident information. Thus, in every accident, the assignment of causal relationships is guided as well as limited by the links. In a systematic accident analysis resulting in causation charts, such a guided assignment of links is preferable to not assigning them at all as in traditional analysis methods, or assigning the causal relationships in a predefined sequence, or with no guidance whatsoever (e.g. Fell 1976). Because of the flexible links, the classification scheme in CREAM and DREAM is not organised in a strictly sequential and hierarchical fashion. Therefore, the classification method has to include stop rules that determine when the analysis has come to an end.

6.3 Different causation charts for the risk and emergency phases

While several accident analysis methods make use of accident or causation charts (Sec. 1.5), Malaterre (1990) describes the only attempt to aggregate causation charts, in terms of accident mechanisms. While the aggregation of the accident mechanisms seemed to work satisfactorily, extending the scope of the aggregated charts proved complicated. This was especially true when including the variety of emergency manoeuvres that occurred in the emergency phase, which clearly indicates that it is a difficult task to represent the whole pre-crash phase in one and the same causation chart. One way to solve this problem could be to divide the pre-crash phase into two; an emergency phase and a risk phase. Because of their different characteristics, each phase could be represented by its own type of chart that is best suited for the phase in question.

The course of events in the emergency phase is highly dependent on the success or failure of driver actions (see e.g. Malaterre et al., 1988). Therefore, the emergency phase can be represented by a method focusing on observable events that interact in a mechanistic fashion. There are several available chain-of-events methods, for example multilinear event sequence methods (Benner, 1975; Hendrick and Benner, 1987) and the Variation diagram method (Leplat 1987), that can be used to compile causation charts with these characteristics.

The risk phase contains contributing factors that cannot be labelled as events, for example factors related to human cognition, software failures and latent failures in general. Due to the focus on events in chain-of-events methods, they are inappropriate for the inclusion of factors that are not directly involved in an observable event (Leplat 1987; Leveson 1995, 2004; Hollnagel, 1998; Paper I). In addition, the causal relationships between factors in the risk phase are less mechanistic than between observable events. Therefore, the causal relationships that occur between factors in the risk phase have to be inferred from the available accident information.

Two traffic accident analysis methods described in Sec. 1.5 do not focus on events; the previously mentioned compilation and aggregation of accident mechanisms (Malaterre 1990) and the Accident causal system (Fell 1976). However, neither of these methods has been sufficiently developed and the accident mechanisms represented only a small part of the pre-crash phase, i.e. the “accident phase”. In spite of the facts that Malaterre (1990) tried to extend the aggregated accident mechanisms by combining them with the factors in the antecedent “driving phase”, the author concluded that aggregating causation charts with greater scope required further development. The Accident causal system (Fell 1976) has a greater scope than the accident mechanisms (see Figure 4b). However, Fell’s (1976) system was only proposed on a conceptual level. Although Fell (1976) outlined an accident model,
the author did not include a classification scheme or classification method for testing and evaluation; nor did he propose causation charts or the aggregation of them.

In conclusion, several event-based methods can be used to compile causation charts for the emergency phase. However, few methods can do the same for the risk phase. DREAM has a classification that focuses on inferred causes that are clearly distinct from observable events. In addition, the DREAM classification scheme incorporates a flexible link scheme that is effective for guiding the assignment of the inferred causal relationships in the risk phase. Therefore, DREAM is well suited for the compilation of causation charts for the risk phase of an accident that occurs prior to the emergency phase. See Figure 15 for an illustration of the aforementioned methods where they are related to the emergency and risk phases.

### 6.4 Experiences gained from aggregating DREAM charts

As previously discussed, there are two purposes behind the aggregation of causation charts. The first is, by means of aggregated causation charts, elucidate common causes and causal relationships that are indicated in case files. The second is, by means of aggregated causation charts, to obtain additional information about causes and causal relationships, of which some has been established in previous studies (Figure 13). So far, however, the only reported attempt to aggregate causation charts is that of Malaterre (1990).

Regarding the aggregation of DREAM charts in Papers II and III, each paper mainly followed each of the two purposes behind the aggregating causation charts. Paper II that focused on the 38 single-vehicle crashes was in line with the first purpose. The case files were examined with the intention of distinguishing crashes with similar circumstances at the moment immediately before the crash. Four types of loss of vehicle control were identified and were similar for some crashes, but different for others. The four types of loss-of-vehicle control were termed scenarios, and the DREAM charts related to each scenario were aggregated.

Paper III that focused on the 26 intersection crashes involving 52 drivers was in line with the second purpose of aggregating causation charts. On the basis of previous studies of databases containing police-reported crashes, the most common violations related to intersections were identified, i.e. running a red light or a sign, or failing to give the right of way. On the basis of these common violations, six risk situations were defined; four risk
situations for the drivers without the right of way and two for the drivers with the right of way. The DREAM charts associated with each risk situation were aggregated.

It is obviously possible to aggregate the DREAM charts in other ways than those chosen in Papers II and III, and is the case for any kind of causation chart. Unlike case files, the advantage of causation charts is that they can be aggregated in several ways and for different purposes. In this regard, the aggregation of causation charts is no different from traditional analysis methods, in which individually coded factors are aggregated, for example in staple diagrams. Individually coded factors are often aggregated in several ways in order to test different research questions. The only, although important, difference between aggregating causation charts and individually coded factors is the appearance of the results. Aggregated causation charts can indicate several interlinked factors at the same time, while individually coded aggregated factors are also aggregated individually. Thus, just as aggregated charts, may or may not indicate that one causation pattern is more frequent than another; aggregated individual factors in a staple diagram may or may not indicate that one factor is more frequent than another.

In this context it should also be emphasised that it is not recommended to simply aggregate all available charts with the intention of identifying common causation patterns. Charts aggregated in this way are likely to become chaotic, unreadable and not very explanatory. The aggregation should have a purpose and be based on a research question, for example an accident circumstance that is of current interest. Aggregating all charts without a purpose would be just as useless as aggregating all individually coded factors in traditional analysis methods, arriving at the conclusion that human error is responsible for e.g. 70 % of the investigated accidents.

6.4.1 Indicating common causation patterns
When aggregating charts the main interest is to identify common causation patterns. To date, common patterns have been indicated in rather simple ways. Malaterre (1990) highlighted them by including the number of common accident mechanisms in the aggregated sequences (see Figures 5 and 6). In the aggregated DREAM charts in Papers II and III, the number of common genotypes and phenotypes were included in the same way, in addition to indicating them by means of line thickness, which also marked the common links while the number of links was excluded.

The occurrence of common patterns could also be highlighted by displaying percentage levels, where percentage is calculated by dividing the number of each common phenotype, genotype or link, by the total number of charts. By displaying percent levels in the aggregated charts, patterns that are clearer and more significant in comparison with others could be identified quickly. Paper III (p. 7) states that a predetermined formal definition of what constitutes a common pattern would be valuable for future analyses of aggregated charts. One formal definition could be to decide on a percentage level below which patterns are irrelevant. See Figure 16, where the aggregated charts for risk situation V (Paper III, Figure 5) are supplemented by three percentage levels.
Figure 16. The aggregated charts for risk situation V (Paper III, Figure 5) are supplemented by three levels of percentage (i.e. the number of each common phenotype, genotype or link divided by the total number of charts x 100). The total number of charts is 14.

Displaying the percentage of common genotypes and links is an illustrative way to indicate common patterns. However, it should be borne in mind that the DREAM charts are based on detailed case files. Therefore, it is also valuable to analyse the appearance of the patterns in detail in order to see what is concealed under the indicated percentage levels. What is not visible in Figure 16 is, for example, the kind of sight obstructions or distractions, and that three of the drivers were subjected to both (compare numbers 14, 17, 25 in Figure 5, Paper III). A subject for further research is to find a way of illustrating the common causation patterns while retaining a sufficient level of detail.

6.4.2 The critical position of the phenotype

In Paper III, the phenotype was positioned differently for drivers without the right of way in traffic light controlled intersections, which in turn influenced the appearance of the DREAM chart (Figure 12). The position of the phenotype also turned out to be problematic in Paper IV for Driver A in scenario 3 despite improved phenotype definitions (Paper IV, p. 11). However, the position of the phenotype did not influence the appearance of the rest of the causation chart, although it may become an issue also when using DREAM 3.0.

Because the phenotype position can influence the factors included in the pre-crash phase, from the phenotype and backwards, the phenotype functions as a kind of “start rule”. A start rule is necessary in all causation chart compilation methods. Without a clearly defined start rule that ensures consistent positioning, there is a definite risk of inconsistency in the compilation of causation charts. Charts compiled from different start points hinder the possibility of identifying accurate and common causation when the charts are aggregated.

6.4.3 DREAM charts are not only for aggregation

While Papers II and III focused on aggregating case files by means of DREAM charts, the charts can also be used for other purposes. Because one DREAM-chart is compiled for each
driver in a crash, the charts can facilitate the understanding of the crash when two drivers have e.g. misunderstood each other’s intentions. In Paper III for example, there were five crashes (2,7,8,21,22) in which all drivers had seen each other. Because the primary purpose of Paper III was to aggregate the DREAM chart in order to find common causation patterns, the failed interaction between the five drivers was not commented on. It is, however, possible to analyse the failed interaction between the drivers involved in the five crashes. If the numbers 2,7,8,21,22 are followed within, and compared between, Figures 4 and 6 in Paper III, then it can be seen that these drivers misinterpreted each other’s intentions, or that the drivers without the right of way misjudged the speed of the vehicle with the right of way. In the remainder of the crashes, at least one of the drivers had failed to see the other vehicle. In 15 crashes, none of the drivers had seen the other vehicle before entering the intersection.

6.5 Data collection in multidisciplinary in-depth accident investigations

Regardless of classification scheme or method for compiling causation charts, the results can never be better than the data on which the classification or chart is based. With reference to data collection pertaining to the risk phase, an interview with the driver(s) is the most important source of information. When interviewing drivers, it is important to consider the time of the interview and the interview method employed.

The time of interview is important, as it has been found that a story may change in a matter of days or even hours. Treat et al. (1977a) suggested that this is an effect of the road user’s own attempt to reconstruct the accident. The longer the interval between the accident and the report, the higher is the risk that a road user adds information that would be consistent with the recalled events and his/her self-perception as a driver. Due to similar experiences, Baker (1960), Englund et al. (1978), and Treat et al. (1977a) agreed on the importance of interviewing the drivers as soon as possible after the accident.

The fact that the time of the interview influenced the driver responses was also noted in the in-depth investigations on which Papers II and III are based. Therefore, the time of the interview was given high priority. A first brief interview was, if possible, conducted on-scene. In many cases, however, it was not possible conduct even a brief interview on-scene, because the drivers were, as a standard procedure, quickly removed by ambulance to hospital for a check-up, even if they had no visible injuries. The drivers who were interviewed on-scene were often stressed by the accident experience and the presence of the police, rescue services and ambulance. Consequently, the time and opportunity for conducting an in-depth interview at the scene of the accident were limited. Therefore, a telephone interview was conducted as soon as possible after the accident.

The fact that the interviews were conducted over the telephone may appear to be a limitation. However, research has demonstrated that the information obtained by means of telephone interview is comparable with that gleaned in the course of face-to-face interviews (Miller, 1995; Sturges and Hanrahan, 2004). Consequently, telephone interviews are not a problem as such. However, it may be a disadvantage that the driver is not at the site of the accident during the interview, and is thus unable to indicate were, for example, the other vehicle was first observed prior the crash. A few attempts were made to return with the drivers to the site of the accident, but they were unwilling to do so within an acceptable period of time. Consequently, despite the potential shortcomings of telephone interviews, they were deemed necessary in order to conduct the interview as soon as possible. In addition, telephone interviews were considered sufficient due to the fact that the current versions of DREAM do not incorporate the exact time or position of, for example, a missed observation. The influence of the place in which the driver is interviewed as well as the incorporation of time intervals for the DREAM genotypes are interesting subjects for future research.
Regarding the interview method, the drivers involved in the crashes analysed in Papers II and III, were invited to narrate their experience of the accident. A semi-structured interview protocol that corresponded to the structure of DREAM was used as a checklist. The interviews took the form of a conversation and leading questions were avoided as far as possible. Research in the area of interview methods (Gubrium and Holstein, 2002) has revealed that this is more advantageous than e.g. interrogations and questionnaires, since the interviewee then feels more comfortable and less hesitant about answering honestly. This interview method also makes the respondent feel more confident, which is especially important in traffic accident research since a driver who has been involved in an accident may feel guilty and decline participation. In the majority of cases, the drivers were honest and often in the process of trying to understand the accident themselves. Only a few drivers declined to participate or were obviously not telling the truth.

While the driver interview is very important, it was not the only source of accident data. The information from the interview was as far as possible compared with that from other sources, for example other involved drivers, witnesses to the accident, and technical evidence. Furthermore, all accident data cannot be obtained from driver interviews. For example, in the investigation of the single-vehicle crashes (Paper II), it was found that the vehicle dynamics and state of the tyres were involved in the loss-of-vehicle control in the crashes related to scenarios 2 and 3 (Figures 2b and 2c, Paper II). This information was not provided by the drivers, but inferred from the skid marks and the impacts on the tarmac and in the ditches, as well as an examination of the tyres. However, these circumstances might not have been identified without one team member’s expertise in the area of vehicle dynamics.

Multidisciplinary teams are necessary for the investigation of accidents that focus on causation. One investigator cannot possess all the required knowledge of road-user behaviour, vehicles, and road environment, and is unable to collect every item of all accident data with a high level of quality, especially when conducting on-scene investigations. However, multidisciplinary accident investigations focusing on causation have been criticised due to the risk of subjective judgements. Grayson and Hakkert (1987) for example argued that this was a limitation “Since the information collected using a multidisciplinary approach is dependent on the cooperation of many individuals, various interpretations of results have occurred” (p. 41-42).

While a multidisciplinary team is necessary for the collection of a sufficient degree and quality of data, it is obvious that subjective judgements must be prevented from influencing the conclusions of the accident investigations. When it comes to classifying causes, the elimination of subjective judgements is even more important. This is relevant whether the contributing factors are classified individually, or that a case file is converted into causation charts with interlinked factors. It is therefore important to ensure that the operational definitions of the factors in the classification scheme are sufficiently clear, so that different investigators can classify causes or compile charts in a consistent and reliable fashion.

### 6.6 Reliability in the classification of qualitative material

“Reliability is the extent to which a measuring procedure yields the same results on repeated trials. The notion relevant to content analysis is that a measure is not valuable if it can be conducted only once or only by one particular person.” (Neuendorf, 2002, p. 112). “Content analysis” is a methodology that is basically identical to the traditional analysis methods applied on case studies. In content analysis, qualitative material such as text and video
material is first coded using listed variables, after which the coded material is analysed with statistical methods.

A majority of the intercoder assessment techniques have been developed in content analysis, in which intercoder agreement is assessed by letting coders test the classification scheme on samples of the qualitative material in question. Typically, two to three coders are engaged in each test (Krippendorff, 2004; Lombard et al., 2002), and the rule-of-thumb regarding sample size is 30 (Lombard et al., 2002). This way of assessing intercoder agreement differs from that used in Paper IV, in which priority was given to engaging several accident investigators due to the known risk of subjective assessments when several investigators are involved (Grayson and Hakkert, 1987). Thus, a total of seven investigators/coders from four countries compiled seven DREAM charts (i.e., samples) for each driver involved in four types of accidents. Consequently, while the number of coders was larger than is customary in content analysis, the number of samples was smaller, and additional tests with supplementary types of accidents are required in order to confirm the obtained level of agreement. However, the accident samples in Paper IV were chosen and described so as to test the known difficulties when compiling DREAM charts, e.g., the positioning of the phenotype at intersection crashes.

Although a majority of intercoder assessment methods is developed within the field of content analysis, there is little consensus on a single best measure (Lombard et al., 2002). However, two measures stand out as the most common: percent agreement and Cohen’s $\kappa^1$ (Cohen, 1960) (Lombard et al., 2002; Spence, 2004). Percent agreement was the measure used in Paper IV. However, percent agreement fails to account for agreements that would occur by pure chance (Lombard et al., 2002; Spence, 2004). Therefore Lombard et al. (2002) and Perreault and Leigh (1989) recommend using an additional measure that also takes chance into account, e.g., Cohen’s $\kappa$.

However, the measures developed in the field of content analysis that take chance into account presuppose two conditions: firstly, that variables are assessed in a list format, and secondly, that all variables in the classification scheme are used on the material in question.

The first condition is clearly not fulfilled in the classification scheme of DREAM due to the stepwise assignment of genotypes guided by the flexible links. Spence (2004) suggested a calculation method based on Cohen’s $\kappa$ for agreement in such uncommon coding schemes which entails multiple steps in coding. However, the method suggested by Spence (2004) was only applicable on classification schemes where the categories in each step were completely separated from all other categories in that step. This is not the case for the classification scheme in DREAM in which the links suggest genotypes from several categories in each step.

Regarding the second condition in content analysis, the employed classification schemes are developed to identify a specific aspect in qualitative material, for instance the level of aggression in action movies. Consequently, all variables are expected to be utilised throughout the coding process. This is not the case for DREAM, where all genotypes are not always used in all cases (See for example stop rule 3, Sec 3.3.1).

In addition to the lack of consensus in appropriate measures of intercoder reliability, there is a lack of consensus in determining what constitutes an acceptable level of reliability (Lombard et al., 2002). In addition, while percent agreement is liberal, i.e. overestimates the level of agreement, Cohen’s $\kappa$ is overly conservative (Perreault and Leigh, 1989), i.e. it underestimates the level of agreement. With regard to conservative measures, Neuendorf

1 Cohen (1960) calculates kappa ($\kappa$) as follows: $\kappa = (F_0 - F_C)/(N - F_C)$, where $N$ is the total number of judgements made by each coder, $F_0$ is the number of judgements on which the coders agree, and $F_C$ is the number of judgements for which agreement is expected by chance.
(2002) concludes that “coefficients of .90 or greater would be acceptable to all, 0.80 or greater would be acceptable in most situations, and below that, there exists great disagreement” (p. 145). For the liberal measure of percent agreement, the acceptable level of reliability is usually set higher. According to Smith (2000) and Krippendorff (2004), the percentage of agreement between coders should be 85% or higher in order to be satisfactory.

The results in Paper IV, indicated that the intercoder agreement for genotypes ranged from 74% to 94% with an average of 83%, while it for phenotypes ranged from 57% to 100% with an average of 78%, and may therefore be seen as somewhat low. However, when comparing percent agreements from various studies, it is important to take into account at what point in the development/research process the intercoder agreement was tested. The measure of intercoder agreement may be inflated if coders are encouraged to discuss coding procedures with each other before the assessment of agreement (Krippendorff (2004); Paper IV). Krippendorff (2004) argues that “The only publishable reliability is the one measured before the reconciliation of disagreements” (p. 219), however this requirement is rarely followed. Consequently, the obtained level of agreement in Paper IV was considered acceptable given the fact that the coders were trained individually on distance and never had contact with each other. Higher levels of agreement are expected to be reached with enhanced training.

In Paper IV, the percent agreement was calculated for the phenotypes and genotypes, but not for the links between them. It would naturally be desirable to account for the reliability of the links. However, the calculation of percent agreement for the links is less evident than that for the genotypes and phenotypes. A general problem is that the position of a minority choice of a genotype results in a different number of minority links. A minority choice of one genotype that is positioned between two other genotypes gives rise to two minority links. If, on the other hand, a minority genotype is positioned to the far left in a DREAM chart (see Paper IV, Figures 5-11), that genotype results in a single minority link. The same goes for minority choices of phenotypes to the far right. Consequently, in the case of estimating percent agreement of the links, a minority choice of a genotype in the middle of the chart give rise to in twice the disagreement in comparison to a genotype to the far left, or a phenotype to the far right.

The appearance of the qualitative material may influence the intercoder agreement due to the coders’ subjective interpretations. In content analysis it is known that the number of neutral evaluations can decrease significantly as the amount of context increase, (Krippendorff, 2004; Neuendorf, 2002; Potter and Levine-Donnerstein, 1999). In story-like text-material “...even carefully trained coders can easily be led in different directions, making reliability difficult to achieve.” (Krippendorff, 2004, p. 109). An accident case file is clearly material that is rich in context and story-like. In such material, Potter and Levine-Donnerstein (1999) argued that “the greatest threat to reliability lies with the coders’ interpretative schemes. If coders have very different schema, there can be very little consistency in coding.” (p. 271). They further maintained that the key to consistency of coding is cuing all coders to use the same schema, for example by using expert or norm standards during coding.

Treat et al. (1977a) defined a driver norm prior the assessment of human causes according to which a “normal” driver is an alert driver exercising the “expected” defensive driving techniques. This driver norm was also adopted by Carsten et al. (1989). While it is unclear whether Treat et al. (1977a) used the driver norm to increase consistency, the agreements between investigators reached an adequate level for the Human direct causes. On the other hand, one of the authors, David Shinar, acknowledged that the driver norm introduced a bias in the assessment, and that it may be a reason for the high proportion of
human factors found in their study (Treat et al., 1977b, p. 177). Consequently, although predefined norms might increase consistency, they should be used with great caution.

With reference to Paper IV, the accident cases were written as clearly as possible to reduce the risk of ambiguous interpretations. Moreover, there is no predefined driver norm in DREAM, although it is strongly emphasised that the responsibility of the accident is of no interest when compiling DREAM charts. Sabey and Staughton (1975) seems to have had the opposite approach, i.e. “In assessing human factors the investigators attempted to allot responsibility for [the] accident to drivers or pedestrians,…” (p. 5).

6.7 Further research: identifying countermeasures by means of causation charts

As early as 1960, Baker and Ross (1960) argued that all the factors identified as contributing to an accident must be present in order for that accident to occur. Regarding accident prevention, some factors may be considered important - not because they contribute more to the accident, but because they are more easily controlled, for instance with countermeasures.

With reference to identifying countermeasures by means of causation charts, Leplat and Rasmussen (1987) suggested two principles. The first was to eliminate a contributory factor altogether, and the second to break the causal relationship between the factors by informing the driver. The first principle is straightforward, albeit sometimes inapplicable. For example, a tree that obstructs the view in an intersection cannot suddenly be removed. In such situations, the only alternative is to inform the driver of, for instance, a vehicle hidden behind the tree.

These two principles cannot be applied in the same ways when factors are coded and aggregated individually by means of e.g. staple diagrams. Here, the only applicable principle is the first one, i.e., the elimination of a factor. What the elimination would lead to is however less evident since factors are separated from each other. Correlation analyses can overcome this latter issue to some extent, but such correlation analyses on factors once all cases have been collected is an unnecessary detour, in addition to the fact that correlation analyses are applicable only on small numbers of factors at a time.

Consequently, the advantage of causation charts is that several contributing factors are kept together by means of the links, even when the causation charts are aggregated. Aggregated charts would allow the identification of countermeasures that have a recurrent effect, while countermeasures that are only relevant to a single case can be rejected (Leplat and Rasmussen, 1987). Nevertheless, before countermeasures are identified on these premises, the charts to be aggregated have to be compiled in a consistent way with a systematic method. The present thesis has shown that DREAM is a highly potential method for the compilation of causation charts for the risk phase of an accident that occurs prior to the emergency phase.
7 Conclusions

Multidisciplinary in-depth case studies of motor-vehicle crashes can provide detailed causation data that is otherwise unattainable. This data allows for the formulation of hypotheses of general causes and causal relationships for further research. By converting case files into causation charts that are aggregated, indicated common causation patterns would give greater weight to the hypotheses. This approach requires that the aggregated charts first be compiled by means of a systematic analysis method, for which there are three necessary parts; a model, a classification scheme and a classification method.

The general Sequential, Epidemiological, Energy transfer and Systemic accident models were found to be inadequate to form the basis for a causation analysis method because they focus on events immediately prior to the crash and comprise either sequential, static, or absent modelling of interaction. In addition, they share a structural problem, in which the road-users, vehicles and traffic environment are treated as separate components. As the resulting analysis method comprises the separation of components in order to identify the failing one(s), the interaction between contributory factors is neglected.

It is possible to identify common causation patterns in aggregated DREAM charts, provided that the charts are consistent. While the common patterns were clear in the study of single-vehicle crashes, those in the study of intersection crashes generally exhibited greater variation. In the latter, three investigators had prepared the charts, which in combination with some indistinct operational definitions in the DREAM classification scheme resulted in inconsistent charts. The definition of the phenotype was found to be important as it functions as a kind of “start rule” in the compilation of DREAM charts. Charts compiled from different start points can hinder the possibility of identifying accurate and common causation patterns when the charts are aggregated.

The intercoder agreement assessment of the revised version DREAM 3.0 indicated that the agreement for genotypes ranged from 74% to 94% with an average of 83%, while it for phenotypes ranged from 57% to 100% with an average of 78%. This an acceptable level of agreement that is expected to increase with enhanced training. Furthermore, assessing intercoder agreement can play an important role in identifying weaknesses in the classification scheme, the training of coders and the presentation of accident information.

The development of vehicle-mounted active safety measures has so far focused on the emergency phase in close proximity to the crash, a phase that is regarded as being beyond driver control and recovery. The development of higher level systems that address the circumstances occurring before this phase requires additional understanding and knowledge of driver behaviour. An obvious way to acquire this knowledge is to make use of case files from multidisciplinary in-depth accident investigations. The focus of the present thesis has been to evaluate whether case files can be aggregated with the help of causation charts. The results show that DREAM is a highly promising method for the compilation of causation charts for the risk phase of an accident that occurs prior to the emergency phase. Future studies are expected to benefit from aggregating DREAM charts when formulating hypotheses of general causes and causal relationships as a subject for further research, as well as identify alternative countermeasure strategies.
8 References


Baker, J.S., Horn, L.R., 1960. An Inventory of Factors Suggested as Contributing to Traffic Accident Causes, Traffic Institute, Northwestern University, Evanston, Illinois.


Haddon, W., 1975. Reducing the damage of motor vehicle use. Technology review 77 (8).
IATSS, 2005. Statistics 2005: Road Accidents in Japan, International Association of Traffic and Safety Sciences,


PAPER I

Accident models for modern road traffic: changing times create new demands

Huang, Y., Ljung, M., Sandin, J., Hollnagel, E.

Understanding the causation of single-vehicle crashes: a methodology for in-depth on-scene multidisciplinary case studies

Sandin, J., Ljung, M.

PAPER III

An analysis of common patterns in aggregated causation charts from intersection crashes

Sandin, J.

Submitted to Accident Analysis and Prevention, 2008
PAPER IV

The intercoder agreement when using the Driving Reliability and Error Analysis Method in road traffic accident investigations

Wallén Warner, H., Sandin, J.

Submitted to Accident Analysis and Prevention, 2008